Educational Technology & Society
An International Journal

Aims and Scope

Educational Technology & Society is a quarterly journal published in January, April, July and October. Educational Technology & Society seeks academic articles on the issues affecting the developers of educational systems and educators who implement and manage such systems. The articles should discuss the perspectives of both communities and their relation to each other:

- Educators aim to use technology to enhance individual learning as well as to achieve widespread education and expect the technology to blend with their individual approach to instruction. However, most educators are not fully aware of the benefits that may be obtained by proactively harnessing the available technologies and how they might be able to influence further developments through systematic feedback and suggestions.
- Educational system developers and artificial intelligence (AI) researchers are sometimes unaware of the needs and requirements of typical teachers, with a possible exception of those in the computer science domain. In transferring the notion of a 'user' from the human-computer interaction studies and assigning it to the 'student', the educator's role as the 'implementer/manager/user' of the technology has been forgotten.

The aim of the journal is to help them better understand each other's role in the overall process of education and how they may support each other. The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to Educational Technology & Society and three months thereafter.

The scope of the journal is broad. Following list of topics is considered to be within the scope of the journal:

- Architectures for Educational Technology Systems,
- Computer-Mediated Communication,
- Cooperative/ Collaborative Learning and Environments,
- Cultural Issues in Educational System development,
- Didactic/ Pedagogical Issues and Teaching/Learning Strategies,
- Distance Education/Learning,
- Distance Learning Systems,
- Distributed Learning Environments,
- Educational Multimedia,
- Evaluation,
- Human-Computer Interface (HCI) Issues,
- Hypermedia Systems/ Applications,
- Intelligent Learning/ Tutoring Environments,
- Interactive Learning Environments,
- Learning by Doing,
- Methodologies for Development of Educational Technology Systems,
- Multimedia Systems/ Applications,
- Network-Based Learning Environments,
- Online Education,
- Simulations for Learning,
- Web Based Instruction/ Training

Editors

Kinshuk, Athabasca University, Canada; Demetrios G Sampson, University of Piraeus & CERTH, Greece; Nian-Shing Chen, National Sun Yat-sen University, Taiwan.

Editors’ Advisors

Ashok Patel, CAL Research & Software Engineering Centre, UK; Reinhard Oppermann, Fraunhofer Institut Angewandte Informationstechnik, Germany

Editorial Assistant

Barbara Adamski, Athabasca University, Canada; Chalarampos Alfragakis, University of Piraeus & CERTH, Greece

Technical Manager

Panagiotis Zervas, University of Piraeus & CERTH, Greece

Associate editors

Vladimir A Fomichov, K. E. Tsiolkovsky Russian State Tech Univ, Russia; Olga S Fomichova, Studio "Culture, Ecology, and Foreign Languages", Russia; Piet Kommers, University of Twente, The Netherlands; Chul-Hwan Lee, Incheon National University of Education, Korea; Brent Muirhead, University of Phoenix Online, USA; Erkki Sutinen, University of Joensuu, Finland; Vladimir Uskov, Bradley University, USA.

Assistant Editors

Yuan-Hsuan (Karen) Lee, National Chiao Tung University, Taiwan; Wei-Chieh Fang, National Sun Yat-sen University, Taiwan.

Advisory board

Ignacio Aedo, Universidad Carlos III de Madrid, Spain; Mohamed Ally, Athabasca University, Canada; Luis Anido-Rifón, University of Vigo, Spain; Gautam Biswas, Vanderbilt University, USA; Rosa Maria Bottino, Consiglio Nazionale delle Ricerche, Italy; Mark Bulen, University of British Columbia, Canada; Talc-Wai Chan, National Central University, Taiwan; Kuo-En Chang, National Taiwan Normal University, Taiwan; Ni Chang, National University of Science South Bend, USA; Yam San Chee, Nanyang Technological University, Singapore; Sherry Chen, Brunel University, UK; Bridget Cooper, University of Sunderland, UK; Darina Dicheva, Winston-Salem State University, USA; Jon Dron, Athabasca University, Canada; Michael Eisenberg, University of Colorado, Boulder, USA; Robert Farrell, IBM Research, USA; Brian Garner, Deakin University, Australia; Tiong Goh, Victoria University of Wellington, New Zealand; Mark D. Gross, Carnegie Mellon University, USA; Roger Hartley, Leeds University, UK; J R Isaac, National Institute of Information Technology, India; Mohamed Jenni, University of Tunis, Tunisia; Mike Joy, University of Warwick, United Kingdom; Athanasis Karoulis, Hellenic Open University, Greece; Paul Kirschner, Open University of the Netherlands, The Netherlands; William Klemm, Texas A&M University, USA; Rob Koper, Open University of the Netherlands, The Netherlands; Jimmy Ho Man Lee, The Chinese University of Hong Kong, Hong Kong; Ruddy Lelouche, Universite Laval, Canada; Tzu-Chien Liu, National Central University, Taiwan; Rory McGreal, Athabasca University, Canada; David Merrill, Brigham Young University - Hawaii, USA; Marcelo Milrad, Växjö University, Sweden; Riichiro Mizoguchi, Osaka University, Japan; Permanand Mohan, The University of the West Indies, Trinidad and Tobago; Kiyoshi Nakabayashi, National Institute of Multimedia Education, Japan; Hiroaki Ogata, Tokushima University, Japan; Toshio Okamoto, The University of Electro-Communications, Japan; Jose A. Pino, University of Chile, Chile; Thomas C. Reeves, The University of Georgia, USA; Norbert M. Seel, Albert-Ludwigs-University of Freiburg, Germany; Timothy K. Shih, Tamkang University, Taiwan; Yoshiaki Shindo, Nippon Institute of Technology, Japan; Kevin Singley, IBM Research, USA; J. Michael Spector, Florida State University, USA; Slavi Stoyanov, Open University, The Netherlands; Timothy Teo, Nanyang Technological University, Singapore; Chin-Chung Tsai, National Taiwan University of Science and Technology, Taiwan; Jie Chi Yang, National Central University, Taiwan; Stephen J.H. Yang, National Central University, Taiwan; Yu-Mei Wang, University of Alabama at Birmingham, USA.

Executive peer-reviewers

http://www.ifets.infoANTED
Supporting Organizations
Centre for Research and Technology Hellas, Greece
Athabasca University, Canada

Subscription Prices and Ordering Information
For subscription information, please contact the editors at kinshuk@ieee.org.

Advertisements
Educational Technology & Society accepts advertisement of products and services of direct interest and usefulness to the readers of the journal, those involved in education and educational technology. Contact the editors at kinshuk@ieee.org.

Abstracting and Indexing

Guidelines for authors
Submissions are invited in the following categories:

• Peer reviewed publications: Full length articles (4000 - 7000 words)
• Book reviews
• Software reviews
• Website reviews

All peer review publications will be refereed in double-blind review process by at least two international reviewers with expertise in the relevant subject area. Book, Software and Website Reviews will not be reviewed, but the editors reserve the right to refuse or edit review.

For detailed information on how to format your submissions, please see:
http://www.ifets.info/guide.php

Submission procedure
Authors, submitting articles for a particular special issue, should send their submissions directly to the appropriate Guest Editor. Guest Editors will advise the authors regarding submission procedure for the final version.

All submissions should be in electronic form. The editors will acknowledge the receipt of submission as soon as possible.

The preferred formats for submission are Word document and RTF, but editors will try their best for other formats too. For figures, GIF and JPEG (JPG) are the preferred formats. Authors must supply separate figures in one of these formats besides embedding in text.

Please provide following details with each submission: • Author(s) full name(s) including title(s), • Name of corresponding author, • Job title(s), • Organisation(s), • Full contact details of ALL authors including email address, postal address, telephone and fax numbers.

The submissions should be uploaded at http://www.ifets.info/ets_journal/upload.php. In case of difficulties, please contact kinshuk@ieee.org (Subject: Submission for Educational Technology & Society journal).
## Table of contents

### Special Issue Articles

- Guest Editorial: Powering Up: Insights from Distinguished Mobile and Ubiquitous Learning Projects across the World  
  Gwo-Jen Hwang and Lung-Hsiang Wong  
  1–3

- Mindtool-Assisted In-Field Learning (MAIL): An Advanced Ubiquitous Learning Project in Taiwan  
  Gwo-Jen Hwang, Pi-Hsia Hung, Nian-Shing Chen and Gi-Zen Liu  
  4–16

- Mobile Phones for Spain’s University Entrance Examination Language Test  
  Jesús García Laborda, Teresa Magal Royo, Mary Frances Litzler and José Luis Giménez López  
  17–30

- Effects of Mobile Instant Messaging on Collaborative Learning Processes and Outcomes: The Case of South Korea  
  Hyewon Kim, MiYoung Lee and Minjeong Kim  
  31–42

- Mobile Inquiry Learning in Sweden: Development Insights on Interoperability, Extensibility and Sustainability of the LETS GO Software System  
  Bahtijar Vogel, Arianit Kurti, Marcelo Milrad, Emil Johansson and Maximilian Müller  
  43–57

- Informal Participation in Science in the UK: Identification, Location and Mobility with iSpot  
  Eileen Scanlon, Will Woods and Doug Clow  
  58–71

- Implementing Mobile Learning Curricula in Schools: A Programme of Research from Innovation to Scaling  
  Chee-Kit Looi and Lung-Hsiang Wong  
  72–84

- Ubiquitous Learning Project Using Life-logging Technology in Japan  
  Hiroaki Ogata, Bin Hou, Mengmeng Li, Noriko Uosaki, Kosuke Mouri and Songran Liu  
  85–100

- Context-Aware Mobile Role Playing Game for Learning – A Case of Canada and Taiwan  
  Chris Lu, Mariga Chang, Kinshuk, Echo Huang and Ching-Wen Chen  
  101–114

- Potentials of Mobile Technology for K-12 Education: An Investigation of iPod touch Use for English Language Learners in the United States  
  Min Liu, Cesar C. Navarrete and Jennifer Wivagg  
  115–126

- The Impact of a Principle-based Pedagogical Design on Inquiry-based Learning in a Seamless Learning Environment in Hong Kong  
  Siu Cheung Kong and Yanjie Song  
  127–141

### Full Length Articles

- Spatial Visualization Learning in Engineering: Traditional Methods vs. a Web-Based Tool  
  Carlos Melgosa Pedrosa, Basilio Ramos Barbero and Arturo Román Miguel  
  142–157

- Developing Digital Courseware for a Virtual Nano-Biotechnology Laboratory: A Design-based Research Approach  
  Hsiu-Ping Yueh, Tzy-Ling Chen, Weijane Lin and Horn-Jiunn Sheen  
  158–168

- Using Instructional Pervasive Game for School Children’s Cultural Learning  
  Cheng-Ping Chen, Ju-Ling Shih and Yi-Chun Ma  
  169–182

- Development and Validation of the Online Instructor Satisfaction Measure (OISM)  
  Doris U. Bolliger, Fethi A. Inan and Oksana Wasilik  
  183–195

- Experiences and Challenges of International Students in Technology-Rich Learning Environments  
  Laurence Habib, Monica Johannsen and Leikny Øgrim  
  196–206

- A Blended Mobile Learning Environment for Museum Learning  
  Huei-Tse Hou, Sheng-Yi Wu, Peng-Chun Lin, Yao-Ting Sung, Jhe-Wei Lin and Kuo-En Chang  
  207–218
A Study of the Design and Implementation of the ASR-based iCASL System with Corrective Feedback to Facilitate English Learning
Yi-Hsuan Wang and Shelley Shwu-Ching Young

219–233

Just-in-Time or Plenty-of-Time Teaching? Different Electronic Feedback Devices and Their Effect on Student Engagement
Jerry Chih-Yuan Sun, Brandon Martinez and Helena Seli

234–244

The Relationships among Chinese Practicing Teachers’ Epistemic Beliefs, Pedagogical Beliefs and Their Beliefs about the Use of ICT
Feng Deng, Ching Sing Chai, Chin-Chung Tsai and Min-Hsien Lee

245–256

Developing Learners’ Second Language Communicative Competence through Active Learning: Clickers or Communicative Approach?
Alaba Olaoluwakotansibe Agbatogun

257–269

An Automatic Caption Filtering and Partial Hiding Approach to Improving the English Listening Comprehension of EFL Students
Ching-Kun Hsu, Gwo-Jen Hwang and Chih-Kai Chang

270–283

Assessment of Charisma as a Factor in Effective Teaching
Yun-Chen Huang and Shu-Hui Lin

284–295

An Investigation of the Effects of Different Types of Activities during Pauses in a Segmented Instructional Animation
Jongpil Cheon, Sungwon Chung, Steven M. Crooks, Jaeki Song and Jeakyeong Kim

296–306

A Review of Research on Technology-Assisted School Science Laboratories
Chia-Yu Wang, Hsin-Kai Wu, Silvia Wen-Yu Lee, Fu-Kwun Hwang, Hsin-Yi Chang, Ying-Tien Wu, Guo-Li Chiou, Sufen Chen, Jyh-Chong Liang, Jing-Wen Lin, Hao-Chang Lo and Chin-Chung Tsai

307–320

SEEK-AT-WD: A Social-Semantic Infrastructure to Sustain Educational ICT Tool Descriptions in the Web of Data
Adolfo Ruiz-Calleja, Guillermo Vega-Gorgojo, Juan I. Asensio-Pérez, Eduardo Gómez-Sánchez, Miguel L. Bote-Lorenzo and Carlos Alario-Hoyos

321–332

Automatic Generation and Ranking of Questions for Critical Review
Ming Liu, Rafael A. Calvo and Vasile Rus

333–346

Implementation of a Model-Tracking-Based Learning Diagnosis System to Promote Elementary Students’ Learning in Mathematics
Yian-Shu Chu, Haw-Ching Yang, Shian-Shyong Tseng and Che-Ching Yang

347–357

Book Reviews

Online Learning and Community Cohesion
Reviewer: Dermod Madden

358–359
Guest Editorial: Powering Up: Insights from Distinguished Mobile and Ubiquitous Learning Projects across the World

Gwo-Jen Hwang¹ and Lung-Hsiang Wong²

¹Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan // ²Learning Sciences Lab, National Institute of Education, Nanyang Technological University, Singapore // gjhwang.academic@gmail.com // lunghsiang.wong@nie.edu.sg

Recent progress in mobile and wireless communication technologies has led to new development of technology-enhanced learning, enabling students to learn in the way that encompasses formal and informal learning across locations and time with supports or guidance from learning systems (Hwang, Wu, Zhuang, & Huang, 2013). The field of mobile learning and ubiquitous learning exemplifies such a trend in developing innovative learning approaches (Frohberg, Göth, & Schwabe, 2009; Wong & Looi, 2011; Wu, Hwang, & Chai, 2013). This trend is probably also reflected by the evolution of the definitions or expositions of mobile learning – from “e-learning using mobile devices and wireless transmission” (Hoppe, Joiner, Milrad, & Sharples, 2003, p. 255) (i.e., e-learning through mobile devices) to “any sort of learning that happens when the learner is not a fixed, predetermined location, or learning that happens when the learner takes advantage of the learning opportunities offered by mobile technologies” (O'Malley et al., 2003, p. 9) (i.e., the mobility of the learning devices) to “increasing a learner’s capability to move their own learning environment as they move” (Barbosa, Geyer, & Barbosa, 2005, p. 1) (i.e., the mobility of learners). In a related vein, ubiquitous learning is explicated as an a learning approach that where the ubiquitous technology is leveraged to support the learners in the right way, in the right place, and at the right time, based on the personal and environmental contexts in the real world (Hwang, Tsai, & Yang, 2008).

In the past decade, various issues concerning mobile and ubiquitous learning have been widely discussed. In the meantime, researchers have reported the effectiveness of adopting mobile and ubiquitous learning approach in various learning contexts (e.g., Kukulski-Hulme, Sharples, Milrad, Arnedillo-Sánchez, & Vavoula, 2009; Milrad et al., 2013; Shih, Chuang, & Hwang, 2010). Recognizing such an emerging trend, the educational authorities of many countries have identified the development of mobile and ubiquitous learning as one of the strategic thrusts in their national educational policy. Consequently, more international, national, regional or institutional-scale mobile and ubiquitous learning initiatives have been embarked on across the globe in recent years (e.g., Buckner & Kim, in press; Cochrane & Bateman, 2010). The common aim is to seek efficient and effective ways of harnessing mobile and wireless communication technologies to create scalable and sustainable learning environments to nurture a new breed of learners with 21st century skills.

In spite of articles reporting on short-term, episodic empirical studies, this special issue seeks papers that trace, summarize and reflect upon individual research programs that may have spanned through several research cycles or consist of multiple sub-projects. Each of the accepted papers covers (the evolution of, if applicable) the background objectives, design rationales of the learning systems/environments, pedagogies and/or learning scenarios, empirical studies and the findings of their projects or studies. In addition, the discussion/conclusion sections of the papers are placing greater emphasis on informing fellow researchers, educators or policy makers about the nuances of translating and sustaining the reported innovative solutions.

This special issue features ten of such papers from ten different countries or economies, which would collectively offer a global perspective in the opportunities and challenges in bridging the research and practice in mobile and ubiquitous learning. From Taiwan, Hwang, Hung, Chen and Liu report a four-year national research project known as “Mindtool-Assisted In-field Learning” (MAIL), with a series of ubiquitous technology-assisted learning and assessment models being developed and evaluated, which has eventually informed and been incorporated into a government initiative of nationwide scaling up of mobile and ubiquitous strategies. From Spain, Laborda, Royo, Litzler and López address two intertwined projects for development and applications of a mobile language testing platform in a university setting. In Korea, Kim, Lee and Kim investigated the effects of mobile instant messaging on collaborative learning, and draw implications to its practical applications. From Sweden, Vogel, Kurti, Milrad, Johansson and Müller present the overall lifecycle and evolution of a mobile learning system developed in relation to the “Learning Ecology through Science with Global Outcomes” (LET’S GO) research project, thus provide deeper insights into the importance of properly addressing the interoperability and extensibility issues in order to develop sustainable solutions for future learning practice. In United Kingdom, Scanlon, Clow and Woods ventured into
informal participation in science by developing and deploying the iSpot system through design-based research; the system has over 31,000 registered users (learners) from within and beyond the UK, according to the paper. From Singapore, Looi and Wong report a two-year one-mobile-device-per-student program conducted in a primary school, which provides a good example of doing research that addresses multi-term, multi-pronged, multi-level and systemic aspects of school-based innovations for benefiting schools, deriving and refine scientifically and empirically theoretical frameworks, and designing principles, resources and strategies for learning. From Japan, Ogata et al. present a four-year project for developing a ubiquitous learning log system, which is able to record students' daily life learning experiences and benefit them in language learning via a log sharing mechanism. From Canada and Taiwan, Lu, Chang, Kinshuk, Huang, and Chen present a context-aware mobile role playing game, which is one of the outcomes of a 5-year program aiming to provide learners with a ubiquitous learning environment that facilitates personalized learning. In United States, Liu, Navarrete and Wivagg present the effectiveness and challenges of using iPods in English language learning at elementary and middle school based on the experiences and findings of conducting a two-year project. Finally, from Hong Kong, Kong and Song present a framework for principle-based pedagogical designs for inquiry-based learning in a seamless learning environment and demonstrate how the approach benefit the students in their knowledge gains and inquiry skills.

From these projects, it is found that the studies and applications of mobile and ubiquitous learning has been shifted from merely using mobile and wireless communication technologies in educational settings to the lead-in of various learning strategies or tools, such as inquiry-based learning, collaborative learning and Mindtools. In addition, some countries, such as Taiwan and UK, have started to popularize mobile and ubiquitous learning approaches to nationally wide or regionally wide scales, implying the rapid growth of mobile and ubiquitous learning applications in recent years. As indicated by Tsai and Hwang (2013), it can be foreseen that mobile and ubiquitous learning will be one of the main trends of technology-enhanced learning in the coming years.

References


Mindtool-Assisted In-Field Learning (MAIL): An Advanced Ubiquitous Learning Project in Taiwan

Gwo-Jen Hwang*, Pi-Hsia Hung, Nian-Shing Chen and Gi-Zen Liu

Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan // Graduate Institute of Measurement and Statistics, National University of Tainan, Tainan, Taiwan // Department of Information Management, National Sun Yat-Sen University, Kaohsiung, Taiwan // Foreign Languages & Literature Department, National Cheng Kung University, Tainan, Taiwan // gihwang.academic@gmail.com // hungps@mail.nutn.edu.tw // nianshing@gmail.com // gizen@mail.ncku.edu.tw
*Corresponding author

ABSTRACT

Scholars have identified that learning in an authentic environment with quality contextual and procedural supports can engage students in thorough observations and knowledge construction. Moreover, the target is that students are able to experience and make sense of all of the learning activities in the real-world environment with meaningful supports, such that their learning motivation can be promoted, knowledge can be sensibly constructed, and skills can be fully developed. To develop potential tutoring strategies and learning activity models using mobile, wireless, and sensing information and communication technologies (ICT) in a real-world learning environment, a four-year national e-learning research project entitled “Mindtool-Assisted In-field Learning (MAIL)” has been funded by the National Science Council of Taiwan since 2008 in an effort to lead the development and innovation of Learning Technology. The integrated project aimed to develop Mindtool-assisted knowledge construction models, assessment models, guidance models, and reflection strategies for cutting-edge context-aware ubiquitous learning. Moreover, a series of learning activities has been conducted to examine the effectiveness of the proposed learning strategies and models. Each year, more than 1,500 students have participated in the in-field learning activities with the designed approaches. Based on the results of a series of experiments, it was found that the students’ learning performance as well as their in-field inquiry ability was significantly improved, showing the effectiveness of the Mindtool-assisted ubiquitous learning approach and the success of the MAIL project. In this paper, the background, objectives, theoretical foundations, systems, research issues, applications, and findings of the MAIL project are presented. Finally, the scaling-up plan for applying these research-proven learning models to all levels of educational settings in Taiwan is also addressed.

Keywords

Mobile learning, Ubiquitous learning, Mindtools, Concept maps, Repertory grid

Background and objectives

Learning Technology (LT), as a trans-disciplinary, professional field of leading human developments and innovations in various situations, disciplines, settings, and industries with advanced technological uses, is always in need of creative applications of hard and soft technologies in order to bring about positive changes (Jonassen, 2004; Liu, 2008). Many educators have identified the importance of situating students in real-world contexts for developing and acquiring knowledge and skills (Brown, Collins, & Duguid, 1989; Lave, 1991; Wong & Looi, 2011). In the meantime, researchers have also pointed out the importance of providing personalized learning supports or knowledge sharing facilities during in-field activities (Sharples, Milrad, Arnedillo-Sánchez, & Vavoula, 2009; So, Seow, & Looi, 2009). The popularity of mobile and wireless information and communication technologies (ICT) has provided good opportunities which match this emerging trend in LT; moreover, the advancement of sensing technology has further enabled learning systems to detect real-world information with various types of information-generating e-readers and e-tags. With the help of these technologies, students are able to learn anytime, anywhere. That is, they are encouraged to learn in various real-world environments with supports from and access to the digitalized world (Hwang, Tsai, Chu, Kinshuk, & Chen, 2012; Looi et al., 2009; Wong, 2012); moreover, dynamic learning systems are developed for the user to engage in more active interactions with other learners as well as the learning system itself (Ogata, Li, Hou, Uosaki, El-Bishouty, & Yano, 2011; Okamoto & Tseng, 2008). Generally speaking, this kind of learning strategy has been called “context-aware ubiquitous learning,” and is a state-of-the-art, particular form of ubiquitous learning (u-learning) as defined by Hwang, Tsai and Yang (2008).

Recently, context-aware u-learning has become a popular issue and research topic in the area of e-learning (Ogata & Yano, 2004; Sollervall, Otero, Milrad, Vogel, & Johansson, 2012; Sylvén, Beale, Sharples, Aholen, & Lonsdale, 2005). Researchers have attempted to conduct context-aware u-learning activities for various courses; however, it
has been found that, without effective learning strategies or tools, students’ learning performance could be disappointing (Chu, Hwang, & Tsai, 2010; Chen & Li, 2009; Liu, Peng, Wu, & Lin, 2009). Several studies have pointed out that u-learning scenarios could be too complex for most students without some proper guidance or supports, because the students need to make use of both real-world and digitalized world learning resources at the same time (Shih, Hwang, Chu, & Chuang, 2011). Therefore, it has become an important and challenging issue to provide effective learning supports in mobile or ubiquitous learning activities.

Among the various learning strategies and tools, Mindtools have been recognized as an effective way of assisting students to learn in complicated learning contexts with all kinds of ICT. Educators have indicated that “technologies should not support learning by attempting to instruct the learners, but rather should be used as knowledge construction tools that students learn with, not from” (Jonassen, Carr, & Yueh, 1998, p. 1). Mindtools are cognition tools that are able to assist students to think and learn in a meaningful and constructive way through stimulating them to expand their cognitive ability in interpreting, analyzing, synthesizing and organizing their knowledge. Jonassen (1999) defined Mindtools as “a way of using a computer application program to engage learners in constructive, higher-order critical thinking about the subjects they are studying” (p. 9). With the assistance of Mindtools, students’ knowledge can be constructed to reflect what they have learned and realized, instead of merely memorizing or recalling content taught by their teachers.

Mindtools and their logical learning design have been widely developed and used with various computer-based application programs, which include database systems, spreadsheets, expert systems, semantic nets (e.g., concept maps), video conference systems, multimedia and hypermedia editing tools, programming tools, and Microworld environments (Jonassen, 1999). In the past two decades, scholars all over the world have paid much attention to using Mindtools in various practical applications related to in-class learning, blended learning, and totally online learning; nevertheless, using Mindtools in u-learning activities remains an important but challenging issue (Lee, Lee, & Leu, 2009).

To develop Mindtool-supported u-learning approaches and to investigate their effectiveness, a four-year national e-learning project was initiated in Taiwan in 2008. The aim of the project was to develop Mindtool-assisted u-learning environments and strategies. Numerous experiments were conducted to evaluate the effectiveness of applying the Mindtool-assisted u-learning approaches to various in-field activities in terms of students’ learning achievement, motivation, attitudes, cognitive load and technology acceptance. Moreover, the students’ in-field observation and question-raising abilities were measured as well.

**Mindtool-assisted ubiquitous learning approaches**

To facilitate in-field learning within context-aware u-learning environments, two kinds of Mindtools were developed to support u-learning activities in the integrated, collaborative project; that is, the grid-based approach which originated from a knowledge elicitation method for developing expert systems and the concept mapping approach that has been widely adopted in in-class learning, blended learning and totally online learning environments.

**Grid-based Mindtools for ubiquitous learning**

An expert system is a computer system that simulates expert-level reasoning based on the knowledge elicited from domain experts. The process and know-how of acquiring and organizing knowledge from domain experts for building knowledge bases of expert systems is called knowledge engineering (Feigenbaum, 1977). Jonassen (1999) indicated that such a process of collecting and organizing domain knowledge for constructing knowledge bases could engage students in critical thinking; that is, an effective way of employing expert systems as Mindtools is surely to engage students in collecting and organizing knowledge related to the course/learning content they aim to learn following a knowledge acquisition approach.

Among various knowledge acquisition approaches, the repertory grid method originating from the Personal Construct Theory proposed by Kelly (1955) has been widely adopted and discussed (Aranda-Mena & Gameson, 2012; Canning & Holmes, 2006; Boose & Gaines, 1989). A repertory grid can be viewed as a matrix whose columns are element labels and whose rows are construct labels. Elements could be decisions to be made, objects to be
identified, or concepts to be learned, while constructs are traits for featuring the similarities or differences between the elements. A construct consists of a trait (e.g., "Long") and the opposite of that trait (e.g., "Short"). Meanwhile, a five-scale rating mechanism is usually used to represent the relationships between the elements and the constructs, where “1” represents that the element is inclined to have the trait and “5” represents that the element is inclined to have the extreme opposite characteristic of that trait.

Referring to the “Expert systems as Mindtools” conception proposed by Jonassen (1999) and the repertory grid method, a repertory grid-oriented Mindtool was developed in the MAIL project for supporting in-field ubiquitous learning in several ways. In the earlier stage of this project, the repertory grid-based ubiquitous learning system was used as a guiding system for helping students observe learning targets in the field, collect data based on their observations, and develop repertory grids for organizing the collected data. Since 2008, a series of learning activities has been conducted with the tool to help students identify and classify a set of learning targets (e.g., plants on a school campus, butterflies in ecology gardens, or rocks in laboratories) via guiding them to observe the learning targets and organize what they have found in a repertory grid using mobile devices (Chu, Hwang, & Tsai, 2010; Wu, Hwang, Su, & Huang, 2012).

Before the learning activities, teachers were asked to develop an objective repertory grid (i.e., a repertory grid with correct ratings for each <element, construct> pair) to guide the students to make observations in the field and develop their own repertory grids. In such a learning-guiding approach, the elements and constructs were provided by the teachers; therefore, the students only needed to fill in the rating for each <element, construct> entry based on their observations in the field. Users were encouraged to consider the objective repertory grid in Table 1, in which the elements were "Lalang Grass," "Agirated-leaf croton," "Cuphea," "Indian almond," "Money tree," "Crown of thorns" and "Pink ixora," and the constructs were "leaf shape," "leaf point," "leaf edge," and "number of leaf vein branches." For example, the value of the <Lalang Grass, Leaf-shape> entry is 1, indicating that the leaf shape of Lalang Grass is "Long and thin." On the contrary, the value of the <Indian almond, Leaf-shape> entry is 4, implying that the leaf shape of Indian Almond tends to be "flat and round."

<table>
<thead>
<tr>
<th>Trait</th>
<th>Lalang Grass</th>
<th>Arigated-leaf croton</th>
<th>Cuphea</th>
<th>Indian almond</th>
<th>Money tree</th>
<th>Opposite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf-shape long and thin</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>Leaf-shape flat and round</td>
</tr>
<tr>
<td>Perfectly smooth leaf edge</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>The leaf edge has deep indents</td>
</tr>
<tr>
<td>The leaf vein has few branches</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>The leaf vein has many branches</td>
</tr>
</tbody>
</table>

During the learning activities, the students were asked to complete their own repertory grids by filling in the rating for each <element, construct> pair based on their observations in the field. If the students failed to give the correct ratings in comparison with those given by the teachers in the objective repertory grid, the learning system would guide them to observe a comparative learning target with the “incorrect main-feature” and would ask them to compare it with the original learning target. For example, if a student observed the plant (learning target) “Lalang Grass” and described its “Leaf point” as “round with a blunt tip” (by giving rating "4"), by comparing the student’s answer with the rating given by the teacher (i.e., “tapering to a long point” with rating "2"), the learning system detected that the student’s answer was incorrect. Accordingly, the learning system tried to find a comparative plant with “round with a blunt tip” leaf points from the objective repertory grid. In this case, “Indian almond” met the condition; therefore, the student was guided to observe the “Indian almond,” and compared its “Leaf point” with that of “Lalang Grass,” as shown in Figure 1. Having made the comparison, the student was then asked to answer the question again. If the student still failed to correctly answer the question, the learning system then provided supplementary materials and the teacher’s rating to the student.

In addition to the guidance from the learning system, the students were able to browse their own repertory grids via the mobile device during the field trips, which was helpful to them in identifying and distinguishing the learning targets via comparing the corresponding ratings of the features of the targets in the grid.
The student is asked to observe the leaf shape of "Indian almond" and compare it with the leaf point of "Lalang Grass". The answer "Round with a blunt tip" to the "leaf point" of "Lalang Grass" is incorrect.

Figure 1. An example of guiding the student to the plant with the “incorrect feature”

To engage students in higher-order thinking during the in-field learning activities, the repertory grid-based Mindtool has further been used to support collaborative knowledge construction in context-aware ubiquitous learning activities in the MAIL project. For example, one of the applications conducted in 2010 aimed to engage students collaboratively in a more challenging learning task; that is, the teacher only showed them the learning targets/objects (i.e., plants on the school campus) without providing any other guidance during the in-field learning process, meaning that the students needed to determine the constructs for identifying and differentiating the plants themselves as well as observing the plants and collecting data. To enable the students to share their repertory grids and make reflections after referring to the repertory grids developed by their peers, a knowledge-sharing system was developed. Through this knowledge-sharing system, the students could upload their repertory grids to the system, browse others' repertory grids, receive feedback from teachers, and discuss with their peers. The experimental results showed that the quality (i.e., completeness and correctness) of individual students' knowledge structure (i.e., the constructs they used and the ratings they gave to represent the relationships between elements) was significantly improved after the ubiquitous learning activity.

Concept mapping as Mindtools for ubiquitous learning

Concept mapping is a well-known learning tool for helping students organize and visualize knowledge and learning experiences (Chiou, 2008; Fischer, Bruhn, Grasel, & Mandl, 2002; Hwang, Wu, & Kuo, 2013; Novak & Cañas, 2006). It is also an effective assessment tool for helping teachers evaluate students' cognitive levels and knowledge structures (Ingeç, 2009; Liu, Don, & Tsai, 2005; Peng, Su, Chou, & Tsai, 2009; Trent, Pernell, Mungai, & Chimedza, 1998). In the past decades, many studies have shown the effectiveness of concept mapping in engaging students in meaningful learning, and hence their learning achievements could be improved (Amadieu, Tricot, & Mariné, 2010; Anderson-Inman & Ditson, 1999; Horton et al., 1993; Markham, Mintzes, & Jones, 2006).

In the MAIL project, concept mapping has been employed in context-aware ubiquitous learning activities in several ways. For example, one of the learning activities was conducted for butterfly ecology observations. Before the learning activity, the students were asked to develop a concept map about butterfly ecology based on what they had learned from the textbook. During the in-field observations, the students browsed their concept map when observing the butterflies in the ecology garden. They could modify the concept maps if something different or interesting was observed. Alternatively, they could take notes and modify the concept maps when they went back to the classroom. Figure 2 shows the scenario of the in-field learning activity. The students were situated in the butterfly ecology garden with wireless communication networks. The garden consisted of 25 ecology areas in which particular kinds of butterflies and the related host plants of the butterflies are raised. In this earlier study, an RFID (Radio Frequency Identification) tag was placed in each ecology area and each student held a PDA (Digital Personal Assistant) with an RFID reader. When the students walked into an area, the learning system could detect the information within the tag in that area via the RFID reader and confirm individual students’ locations, such that corresponding learning guidance or support could be provided.
In recent studies of the integrated MAIL project, smartphones and the QR-code sensing technology have been used to support extensive self-directed learning in the field. A series of large-scale and long-term activities has been conducted in Chiku Ecology Park in southern Taiwan, where various species of mangroves grow. Figure 3 shows the plan for one of the learning activities. In this activity, the students were equipped with a smartphone to interact with the learning system as well as a telescope for long-distance observations. The learning system provided online instant feedback (e.g., hints to remind the students that part of their answers to the questions raised by the learning system were incorrect) and learning guidance (e.g., clues to find the correct answers to the questions) to the students via wireless communications. Moreover, an e-library was developed to provide supplementary materials for the field-based activities.

A series of concept mapping tasks was developed to scaffold the students’ knowledge construction during the field trips in a progressive manner. In the first stage, the learning tasks included multiple-choice questions, short-answer questions, and a structured two-level concept map in order to help the students clarify their knowledge about the basic features of individual mangrove species. In the second stage, the learning tasks were designed to encourage the students to describe the advanced features for classifying different species via the use of developing multiple-level concept maps. In the third stage, the learning tasks aimed to guide the students to compare the species and find the relationships among the species based on what they had learned and observed via developing the multiple-level and cross-relationship concept maps.
Figure 4 shows an illustrative example of a student’s emerging concept map while working in the field. The title of the concept map is “the life of Idea leuconoe clara,” which is a species of butterfly found in Taiwan. Before the field trip, the student developed an initial concept map consisting of four stages (i.e., egg, pupa, larva and imago) to describe the life of Idea leuconoe clara. Later, he went to the butterfly garden to complete the learning tasks. When observing the butterfly ecology in the field, the student browsed the initial concept map and found several facts to be added: (1) In stage 1 of the Idea leuconoe clara, he had only described the egg as being “white or lemon yellow;” however, in the field, he noted that the egg was also “translucent.” Therefore, this new feature was added to the concept map, as shown in block A of Figure 4. (2) In stage 2 of the Idea leuconoe clara, he had not described the features of the larva. When learning in the field, he noted two of its features, that is, “red spots on the sides of the body” and “black and white;” therefore, these two features were added to the concept map, as shown in block B of Figure 4. (3) In stage 4 of the Idea leuconoe clara, he had not given examples of food plants. When observing in the field, he found that the Idea leuconoe claras were acquiring honey from magnolias; therefore, the proposition “magnolia is an example of food plants” was added to the concept map, as shown in block C of Figure 4.

Figure 4. Example of a students’ emerging concept map while working in the field

Applications and research items

From August 1, 2008 to September 30, 2012, the research team conducted 73 u-learning activity-based studies to try out the learning system, Mindtools, learning models and strategies, evaluation scales, and the design of the learning content and activities. The in-field learning environments included the Chiku Mangrove Conservation Area, the Chiku Black-faced Spoonbill Conservation Center, the butterfly ecology garden in Cheng-Kung Elementary School, the science parks and museums in several cities across Taiwan, and the campuses of several educational settings in Taiwan. The learning content of the in-field activities not only focused on natural science, but has also been extended to various academic disciplines. So far, there are 38 natural science studies, 19 social science studies, 5 computer science studies, 2 nursing training studies, 7 language learning studies, 1 mathematics study and 1 Art learning study conducted as part of the MALL project.

The total number of participants has increased each year, from about 500 participants in 2008 to nearly 3,000 in 2012, as shown in Figure 5. To date, the total number of participants in this integrated LT project has reached more than 6,000 students. Most of these experiments have been conducted by comparing the learning performance of experimental groups and control groups; moreover, the students’ pre-and post-test scores as well as their perceptions collected based on several measures have been analyzed. It is evident that based on these experiment results, both the
quantity and quality of this integrated, collaborative research project are satisfactory in terms of academic rigor and knowledge innovation.

![Number of participants](image)

**Figure 5.** The number of students participating in the MAIL ubiquitous learning studies in 2008-2012

### Achievements and implications

In the early experiments, we aimed to investigate the students’ perceptions of learning with the ubiquitous learning approach in comparison with their past experiences of learning with the traditional one-to-many in-field instruction. In the meantime, the teachers’ perceptions of conducting u-learning activities were also investigated. For example, one of the experiments was conducted to collect the feedback from 30 elementary school students and 9 teachers after they experienced a u-learning activity in the butterfly ecology garden (Peng et al., 2009). From the questionnaire survey, it was found that, in comparison with the traditional instruction, the students’ learning motivation and interest were significantly promoted with the help of the personalized guidance and feedback provided by the u-learning system in the field. The average rating given by the students was 4.53 in a five-point Likert rating scheme. Moreover, from the interviews, it was found that the teachers highly accepted the u-learning approach owing to several reasons: (1) it provided the students with better access to online resources during the field trip; (2) it enabled the students to make observations and collections with learning guidance without being constrained by time or location; (3) the u-learning activities engaged the students in learner-centered activities seamlessly across locations and contexts; and (4) the u-learning approach was able to provide step-by-step expert advice and record the students’ learning portfolios.

To improve the students’ learning achievements, in the second stage of the MAIL project, we aimed to compare the effectiveness of the Mindtool-assisted in-field learning approaches with that of the conventional tour-based u-learning approach, which guides individual students in the field, providing them with supplementary materials and giving feedback to them based on their observations and input. It was found that, with the assistance of the Mindtools, the students’ learning achievements, as well as their learning attitudes, were significantly improved. Moreover, it was also found that via the sharing of the constructed knowledge (e.g., repertory grids or concept maps), the knowledge structures as well as learning achievements of the students were further improved. For example, in an experiment for conducting a “plant identification” activity at an elementary school campus, the repertory grid method was implemented in the u-learning system to serve as a Mindtool to help the students summarize the features of the plants observed in the field (Chu, Hwang, & Tsai, 2010). From the experimental results, it was found that the Mindtool-integrated approach not only enhanced the learning interest (the average rating changed from 4.85 to 5.31 in a six-point Likert rating scheme), but also improved the learning achievements in comparison with the conventional u-learning approach via ANCOVA analysis ($F = 9.573$, $p = 0.011$ and $d = 1.39$) for the two groups of students. Another experiment was conducted in the butterfly ecology garden with embedded concept mapping in the u-
learning system to help students organize what they had observed in the field and compare the acquired knowledge with their prior knowledge learned from the textbooks (Hwang, Shi, & Chu, 2011). The experimental results showed that, after the learning activity, the students who learned with the Mindtool-based u-learning approach showed a significantly positive change in their attitudes toward learning science (from an average rating of 3.97 to 4.38 in a five-point Likert rating scheme); moreover, their learning achievements were significantly improved in comparison with the achievements of those who learned with the traditional concept maps (with paper and pencil) in the field and the conventional u-learning approach based on the ANCOVA result ($F = 4.257, p < 0.05$).

In the third stage of the MAIL project, Mindtool-integrated u-learning was included in the formal science curriculums of several selected schools in Taiwan. Accordingly, several long-term activities were conducted in field trips to observe the growth of students’ inquiry competences with the u-learning approach, such as problem-posing and problem-solving abilities. For example, in one of the activities conducted in the Chiku Ecology Park, the participating students were forty-nine elementary school students aged 11.5 years old on average. The students experienced the field trips within four months to complete a series of learning tasks. Twenty-five of them who were assigned to the experimental group learned with the u-learning approach. Another twenty-four students who were the control group learned with the traditional in-field instruction; that is, they were guided and instructed by the teacher on the field trip. The students’ inquiry performances were evaluated by the teachers based on several criteria, including the quantity and accuracy of the descriptions of the learning targets for completing the learning tasks, the number and quality of the questions raised and the responses to the peers’ questions during the field trip, and the relevance and correctness of the features and relationships used to describe their findings in the learning diaries. It was found that through the assistance of Mindtools, the students’ inquiry behaviors, such as the quantity and quality of the questions they raised and the depth of their descriptions of their observations in the field, were significantly increased in comparison with traditional in-field learning based on the ANCOVA result ($F = 4.72$ and $p < 0.05$); in the meantime, the students’ learning performances were significantly improved.

Another three-month experiment was conducted to compare the learning performance of 18 gifted students and 30 average students who were 11.5 years old on average. The participants were scheduled to learn with the concept map-based u-learning approach in the Chiku Ecology Park. Within the three months, the two groups of students showed remarkable progress in ecology observations based on the Computerized Ecology Observation Competence Assessment (CEOCA) developed by Hung, Hwang, Lin, Hung and Wu (2010). The CEOCA consisted of three facets, that is, knowledge, observations and conceptual relationships. The test items were presented with real pictures, films or concept maps. In the pre-test, the average performance of all of the participants was close to the norm (0.04 vs. 0.00) of the students of the same age in Taiwan. After the learning activity, the post-test scores showed that the average growth slope of all of the participants was significant ($\mu = 0.27, p < .01$) in comparison with their pre-test scores with effect 0.53; however, there was no significant difference between the two groups. By conducting a follow-up test one month later, a significant difference was found in the CEOCA scores between the two groups. The gifted students revealed positive performance growth, while the performance of the average students decayed after the learning activity, showing the need to provide continuous supports to average students after field trips (Hung, Hwang, Lin, & Su, 2012).

Furthermore, some experimental results also showed that the Mindtool-assisted u-learning approach can help students improve not only their learning achievements, but also their higher-order critical thinking competences. For example, in one of the u-learning activities conducted in the butterfly ecology garden, the students were asked to develop repertory grids based on what they observed on the field trip (Hwang, Chu, Lin, & Tsai, 2011). The experimental results showed that the students who learned with the Mindtool-based u-learning approach showed better learning achievements than those who learned with the conventional u-learning approach. By comparing the students’ answers to the learning sheets before and after participating in the repertory grid-based u-learning activity using a t-test, it was found that the students’ ability of determining the characteristics for differentiating the butterflies and their competence for identifying and differentiating the butterflies had significantly improved with $t = 7.13$ ($p < 0.001$) and $t = 9.23$ ($p < 0.001$), respectively. This implies that their higher order thinking (i.e., analysis and evaluation) performance was improved.

In addition to the lead-in of various Mindtool-based u-learning strategies, it should be noted that some of the participating schools of the MAIL project have already included such Mindtool-assisted u-learning approaches as part of their regular curricula. For example, a nursing school in southern Taiwan not only prepared their own u-learning equipment (i.e., mobile devices, wireless networks and sensing devices) after participating in one of the
experiments of the MAIL project, but also started to use the repertory grid-based u-learning approach as a standard way of teaching some clinical nursing courses.

Another issue raised in the MAIL project was the cognitive load of the students who participated in the u-learning activities (Hwang, Wu, Zhuang, & Huang, 2013). As the students needed to interact with the real-world learning environment as well as the e-learning system simultaneously, there was a concern that their cognitive load might be too great in some cases; therefore, several experiments of the MAIL project measured the students’ cognitive load using the measures developed by Paas (1992) and Sweller, van Merriënboer, & Paas (1998). It was found that, with a proper learning design, the Mindtool-assisted u-learning approach could significantly decrease students’ cognitive load; on the contrary, students were likely to meaningfully expand their cognitive capability after the practice of integrating in-field observation and technology-driven knowledge construction into situated learning. This decrease could be due to the fact that the Mindtools were able to assist the students in organizing the collected data from the field by linking the chunks of information in a well-structured form, which eased their load in interpreting the data (Verhoeven, Schnotz, & Paas, 2009). For example, the repertory grid-based Mindtools can help students organize the observed features of the learning targets in a unified form (i.e., ratings ranging from 1 to 5), which is very helpful to them for comparing the learning targets and identifying the significant features that can be used to distinguish the targets. Consequently, students’ cognitive load could be decreased; in the meantime, their learning achievements could be improved owing to learning in a more efficient and effective way.

From the series of related studies conducted in the MAIL project, it is found that technologies are not the key or solution to cope with in-field learning problems. Without proper learning supports in the field, students might feel helpless, frustrated and aimless, and hence their learning attitudes or motivations could be affected. Moreover, their cognitive load can be high owing to the strategies or tools used to link what they have learned and observed together, and hence their learning achievements could be disappointing. On the other hand, from the experimental results, it is also suggested that grid-based tools are effective in helping students identify and differentiate a set of learning targets, while concept mapping tools are helpful to students in linking and organizing what they have observed in the field and have learned from the textbooks. That is, grid-based tools help students observe the learning targets with a "micro view," while concept mapping enables them to see things with a "global view." For example, if the aim of a context-aware activity of a language course is to help students learn to use vocabulary, phrases and sentence patterns related to the contexts, concept mapping could be useful; similarly, if the aim of a social studies course is to let students have a whole picture of a cultural asset, concept mapping is also a good choice. Nevertheless, if the aim is to foster students’ ability of identifying or differentiating the artifacts from different historical periods, grid-based Mindtools are good candidates.

Therefore, when designing Mindtool-based u-learning activities for different subjects, the following procedure is suggested:

1. Review the nature of the learning content to see if the aim of the subject unit is relevant to identifying and differentiating a set of learning targets based on their features, or organizing the relevant concepts by finding the relationships between them. Accordingly, the Mindtools to be employed in the learning activity can be determined.

2. Design the learning tasks based on the aims of the activity. For the learning activities with grid-based Mindtools, both problem-based and inquiry-based learning tasks are recommended, depending on the level of learner control. For novice or younger learners, problem-based learning with instant feedback would be preferable; for experienced or older ones, inquiry-based learning with supplemental materials in e-libraries or on the web would be better. On the other hand, for the activities which incorporate concept mapping strategies, inquiry-based learning is recommended.

3. Determine the technologies used in the learning activities. It is suggested that at least mobile and wireless communication technologies are required, while sensing technologies are optional. One of the reasons for adopting sensing technologies is to provide students with learning tasks, learning supports or supplementary materials at the right place and at the right time, which not only reduces the load of students in searching for the information, but also makes the learning process more efficient.

4. Determine the way to measure the learning performance of students and provide feedback to them. For the activities using grid-based Mindtools with the problem-based learning approach, automatic scoring and instant
feedback can be provided. For inquiry-based learning, scoring rubrics need to be defined by teachers in advance for measuring students’ findings, dialogs, learning sheets, and learning behaviors; moreover, a knowledge sharing mechanism could be helpful to the students in making reflections.

Conclusions and future work

In this paper, an advanced u-learning project entitled “MAIL” with various dimensions of research design, issues and contributions, has been presented. The integrated project has aimed to develop Mindtool-assisted u-learning environments to improve the in-field learning performance of students. A total of 73 u-learning activity-based studies have been conducted in the past five years to investigate the effectiveness of the Mindtool-assisted u-learning approach in terms of improving students’ learning achievements, learning motivation, learning attitudes, and technology acceptance degrees, among other aspects. The experimental results show that the research-proven approach with multiple practices in various settings is both promising and appealing.

In terms of LT innovation, the findings of the MAIL project provide several new contributions to the field of mobile and ubiquitous learning. It has been demonstrated that simply adopting new technologies for students to learn in a real-world learning environment is not good enough. What is more important is for us to design appropriate pedagogical strategies as Mindtools for providing better support to students in an authentic in-field ubiquitous learning environment with procedural and contextual components. It is expected that the accomplishments of the MAIL project can provide research-proven LT know-how of Mindtool-assisted ubiquitous learning as well as references for those researchers and practitioners who are interested in conducting in-field activities with instant supports from technologies.

The various designs and experiments described in this paper can serve as a good reference model for practitioners and researchers who are interested in this emerging field of LT. This paper also reveals several essential future research topics, which are summarized as follows:

- Track students’ learning activity logs as a way to support learning analytics studies in mobile and ubiquitous learning environments. For example, it would be interesting to analyze the students’ learning patterns and investigate the relationships between the patterns and their learning performance. Moreover, it is important to further examine the effects of the Mindtool-based u-learning on students’ higher order thinking based on the learning logs of students’ in-field learning behaviors.

- Provide instant and personalized learning supports based on the learning logs and profiles of individual students. Although Mindtools are theoretically helpful to students in constructing and organizing knowledge, students might find it difficult to effectively use Mindtools during the in-field learning activities. For example, some students might have difficulty in developing concept maps without appropriate assistance. That is, while learning with Mindtools in the field, students might require instant and personalized supports. Therefore, it is important to provide instant learning supports by analyzing the learning logs and profiles to identify their problems and needs in the field.

- Develop seamless learning environments by integrating front-end in-field learning experiences with the backend support of Learning Management Systems (LMS) by using the cloud technology so as to apply versatile Mindtools in more courses. In addition to the concept map and grid-based Mindtools developed in this paper, other Mindtools reported by Jonassen (2004) could be included for helping students learn in more effective and constructive ways. For example, spreadsheets could be an effective Mindtool for helping students infer the relationships between variables in Mathematics and Physics courses; database management systems could be Mindtools that engage students in analytical tasks; simulation software could be helpful to students in associating abstract theories with real-world scenarios. Therefore, it is worth investigating the possibility and effectiveness of applying those Mindtools to different u-learning activities.

These cutting-edge research topics and issues are worth our efforts to shed more light on this ever-changing, promising field of context-aware ubiquitous learning in Learning Technology. Recently, the Ministry of Education in Taiwan has initiated a large-scale program for applying mobile and ubiquitous strategies and tools in all levels of schools. In 2012, one hundred schools were selected as demonstration sites, and the number of schools participating in the program will be increased each year. In those schools, each student in the selected classes is equipped with a
mobile device. In each city or county, a cloud-based educational service system has been established to support the anywhere and anytime learning. Moreover, a series of training programs has been proposed to train teachers in how to design in-class and in-field activities with the strategies and tools developed based on the experiences and findings of MAIL and some other studies. It is expected that mobile and ubiquitous learning will become a regular form of learning in the coming five years in Taiwan.

Acknowledgments

This study is supported in part by the National Science Council of the Republic of China under contract numbers NSC 100-2631-S-011-002 and NSC 100-2631-S-011-003.

References


Mobile Phones for Spain’s University Entrance Examination Language Test

Jesús García Laborda1, Teresa Magal Royo2, Mary Frances Litzler1 and José Luis Giménez López2

1Departamento de Filología Moderna, Universidad de Alcalá, calle Trinidad, 3, 28801 Alcala de Henares (Madrid), Spain // 2Departamento de Ingeniería Gráfica, Universidad Politécnica de Valencia, Camino de Vera, s/n, 46022 Valencia, Spain // jesus.garcialaborda@uah.es // tmagal@degi.upv.es // mf.english.uah@gmail.com // jojilo@degi.upv.es

*Corresponding author

ABSTRACT

Few tests were delivered using mobile phones a few years ago, but the flexibility and capability of these devices make them valuable tools even for high stakes testing. This paper addresses research done through the PAULEX (2007-2010) and OPENPAU (2012-2014) research projects at the Universidad Politécnica de Valencia and Universidad de Alcalá (Spain) to provide a powerful but low cost delivery system for the foreign language paper of the Spanish College Entrance Examination (henceforth PAU). The first project, PAULEX, intended to create a robust mobile platform for language testing while the second, OPENPAU, examined the specific applications of ubiquitous devices to create more dynamic forms of assessment. This paper focuses on the projects’ design, testing theory, and technical evolution including visual ergonomics. The current results demonstrate the technical and didactic feasibility of mobile-based formal assessment that aligns student needs with the kind of inferences that the mobile based language test should provide academic authorities.

Keywords

Mobile learning, High-stakes testing, College entrance examination, Foreign language, Higher education

Introduction

Mobile phones have been playing an increasingly significant role in education in the last years, and although until recently very few tests were delivered through them, their flexibility and capability to do so have suggested their potential even for high stakes-testing. High-stakes testing can be defined as those tests with important consequences for the test taker such as acceptance to university, a scholarship, or a license to practice a profession, all of which may have a great influence of the testee’s life. Using mobiles beyond their traditional uses such as podcasts, mp3 applications, and even learning apps seems to be a real challenge at this point, yet they have already been used for language testing as in PhonePass, previously called SET- 10 (http://www.7act.net/7ACT_files/set10.pdf) test. The validity of the test has repeatedly been supported (Downey, Farhady, Present-Thomas, Suzuki & Van Moere, 2008) but little evidence has been provided of its operativeness in real educational contexts. As a consequence, the potential opportunities for mobiles for language testing are still open (Valk, Rashid, & Elder, 2010) but sound projects need to be implemented. This is the kind of research that the Ministry of Education started to support in 2007. By that year, the Spanish and regional educational authorities responsible for the high stakes University Entrance Examination (“Prueba de Acceso a la Universidad”, henceforth PAU) had determined the need to design a new test with greater validity than the current paper-based test, which only included the traditional tasks of reading and writing along with grammar questions. The new test had to include listening and speaking activities. However, budget cuts reduced the chances of implementing a new exam that could include speaking and listening activities unless a low-cost possibility could be found. With a view toward designing a modern test in the hopes of saving the possibility of including the two skills, the Universidad Politécnica de Valencia obtained funds for the development of an online testing system (PAULEX, “PAU en Lenguas Extranjeras”) project from 2007-2010 and the OPENPAU (“PAU abierta”) later on between 2012-2014 whose results were described by García Laborda (2012). After testing over 150 students online it was concluded that implementation of the computer test would save human resources and be economically feasible in a period of two to three years. As a subproject, the PAULEX project addressed the use of mobile phones (García Laborda & Giménez López, 2010), which is the main focus of this paper.

Literature review and theoretical approach

When addressing this sub-project, the research team felt that the use of mobiles phones, like any other delivery system, could not challenge three main testing features: validity, reliability, and practicality (Bachman & Palmer, 1996, among others). Validity here means that if a student gets a score of X on the test, it means that he should be
able to study using that foreign language at university; reliability provides information on the “the precision of the test measurement” (Salmani-Nodoushan, 2009, p. 1); and practicality implies that the test can be implemented in real life. In addition, the test construct in which mobiles are to be applied has taken into account current theories in language testing and Communicative Competence (Canale & Swain, 1980; Canale, 1984; Bachman & Palmer, 1996). Thus, the questions considered were (1) why use mobile phones for language testing? (2) How can the basic testing features be assured? (3) What learning theories are implicated in their use?

For most part, the collection of evidence in both projects was based on Weir’s validation framework and the Evidence Centered Design (Mislevy, Steinberg, & Almond, 2002; Mislevy & Haertel, 2006). Weir (2005) feels that the reliability of a test depends mostly on its conditions for validation. For him, it is necessary to have warrants that there will be two main types of validity in implementing a test: context validity and theory-based validity. Both are interrelated and need to be considered interdependently. Context validity is divided into three parts: task, setting and administration, and task demand (similar to test construct), while theory-based validity includes executive processes and executive resources. Given this framework, the theoretical application for the mobile application of the PAU took into account the aspects included in figure 1.

---

**Figure 1. Delivery framework (based on Weir, 2005)**
As indicated in figure 1, the theoretical aspects that needed to be considered in the implementation of the mobile-based test were grouped into three main phases: design, delivery, and consequential. The limitations and scope of this paper only allow for a discussion of the most important aspects of the design phase. Within this framework, this paper will mostly focus on contextual validity because this is where Weir places the delivery system factor. However, there is no question that the implications of using certain delivery systems—whether pen-and-paper or mobile—are present in all the aspects presented in figure 1. This approach is based on performance-oriented tasks, which intend to resemble communicative acts of the language. In designing the test process in the test, two main options were included: a cognitivist and a social constructivist intervention. While the cognitivist approach supports the notion that students bring knowledge at the time of testing and this is represented by observable behaviors, the social constructivist approaches in language testing are very much related to the development of the Zone of Proximal Development (ZPD) (for further discussion see Poehner, 2008) through the examiner’s intervention and moderation. The Zone of Proximal Development is defined as “...the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978, p. 86). Based on the development of the ZPD, Poehner (2008) also suggests that Dynamic Assessment includes both approaches. Both PAULEX and OPENPAU projects follow the principles of Dynamic Assessment (DA). This approach considers that students have a current level of language knowledge (the one they show without test moderation) and a potential level (the one that they show when their output is computer or human-based moderated and improved through this interaction). According to DA, both can be included in a two-part assessment if the first is moderated (socio cognitive approach) and the second just serves to obtain current language evidence (cognitivist approach) without any tester’s intervention. This can be seen the process shown on figure 2.

According to these principles and given their experimental nature, the PAULEX and OPENPAU projects placed more emphasis on achieving a sound design based on experimental evidence than on potential achievement scores through the use of the test. In practice, evidence was collected and recorded through the use of mobiles. There were five main benefits that justified the decision: (1) the lower cost of mobile based hardware; (2) immediacy of rating and results; (3) ease of recording during oral interviews (hence, data available for further revision of the test and research); (4) the candidates’ familiarity with the delivery means; (5) possibilities for students to rehearse; and (6) ease of rating and administration. Additionally, accessibility for schools and/or official testing centers would enable the optimization of space.

Evidence obtained from the tests, which was moderated, was processed through the Evidence Centered Design (figure 1) (Mislevy, Steinberg, & Almond, 2000; Mislevy & Haertel, 2006) after the first interview, and implied the design of adequate tasks that considered all the linguistic requirements (as seen in figure 1) and could be delivered through mobile phones. According to the cognitivist approach, used in the second testing session, tasks had to be automatically delivered and recorded without moderation to provide current real data. Then the responses were rated (and the scores validated) and with a view toward having an impact on decisions for teaching and high stakes decisions.

![Figure 2. Development of test process](image)

Overall, the research team felt that the use of mobile phones was strongly founded but they recognized that the advantages and disadvantages needed to be weighed. Table 1 presents the pros and cons of their use:
Table 1. Use of mobile phones according to test characteristics

<table>
<thead>
<tr>
<th>Test characteristics</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place and time circumstances</td>
<td>Convenience of location and time because they require no external human presence</td>
<td>Security and technical assurance of full functionality required; otherwise test is at stake</td>
</tr>
<tr>
<td>Test rubric and process for responses</td>
<td>Tasks are consistent and delivered exactly in the same way to all testees</td>
<td>Testees may have different levels of understanding of the instructions according to proficiency levels</td>
</tr>
<tr>
<td>Test delivery</td>
<td>Current developments in mobile phones increasingly enable the inclusion of audio and image, thus enhancing contextualization and richness of responses</td>
<td>Standardization of mobile phones used for the test is needed; otherwise significant differences in responses can be found even by the same speaker.</td>
</tr>
<tr>
<td>Construct, rating, and scoring</td>
<td>Automated rating validates the equanimity in multiple choice items; separation of the rater from the testee enables rating protocols to be followed without the influence of contact with the testee. Thus assessments are more objective.</td>
<td>Human protocols do not assure complete equanimity (Baldwin, Fowles, &amp; Livingston, 2008).</td>
</tr>
</tbody>
</table>

Mobile phones in high stakes testing: The PAULEX project (2007-2010)

Since the implementation of the originally planned computer-based language testing platform was costly - albeit assumable in the long term-, one of the suggestions for the researchers was the use of mobile phones for the Speaking test only until the online platform could be used. However, while mobiles were originally thought to support student training, almost from the beginning the project management felt that they could also have a very positive effect on learning and they could encourage after-test washback effects. The main reasons to implement mobile phones were that the hardware was less costly than for computers, their use could be more accessible as they can easily be delivered and collected to and from each school, and their use could facilitate rapid assessments by testing units (which would resemble calling centers in their functioning and organization). These testing units could potentially organize and deliver a large number of tests in a limited time. The tests could be delivered automatically; the students’ responses could be recorded and assessed later by human raters.

From the beginning of this three-year project, it was clear that a well-trimmed double design project was needed for the delivery, ergonomics, and content inclusion. Figure 1 describes the organization of the PAULEX project.

![Figure 2. Organization diagram of PAULEX project](image)
As can be seen in figure 2, two branches were organized: one devoted to the linguistic and validation aspects, and the other focused on the technological design of the online and mobile platforms. From the beginning it was clear that most of the significant difficulties were associated with the test design since the technology group had already been involved in similar projects before. Because the validation process was central to the project, the mobile application was designed and tested considering a variety of students and also bearing in mind that the PAU project served to obtain inferences of whether students would be able to use English for university work. Furthermore, the mobile technology branch considered that not all students have the same ability in using mobile technology so the technological specifications were relevant and accessible to students with special needs.

The development of the mobile phone subproject within the larger PLEVALEX project was intended to provide information on three aspects: (1) student adaptability to the new environment, (2) content and test validity for the listening/speaking tasks, and (3) delivery reliability. As mentioned above, mobile phones have been thought to foster learning more than to be used to assess students. Learning would take place by providing them with test samples that could be used anywhere and at any time. In this way mobile phones would bring to the fore the required testing skills in combination with similar listening and speaking tasks along with affective considerations but in a more interactive and usable manner. This process would also provide opportunities for authentic learning and the elimination of test fear would probably favor motivation. In this sense, the mobile phone sub-project sought to engage students in terms of motivation, high stakes test practice, and language learning. The results for this project were obtained through triangulating linguistic achievements, field notes, and a usability analysis carried out through a 20-item questionnaire computer delivered to all the students who took part in the research (García Laborda, Giménez López, & Magal Royo, 2011).

Validation method

We used five types of validation analysis for the PAULEX project. First, we did a Delphi analysis (Custer, Scarcella, & Stewart, 1999) to foresee potential issues in the mobile phone test with experts and then a reduced number of regular users (3). Second, we observed the intended scores of the mobile phone users and compared them with those obtained with the online platform (Sariola, 2003). Next, we analyzed the video recordings from the pilot studies. After that, inspection techniques were followed to do a usability analysis (Nielsen & Mack, 1994). Finally, the students’ attitudes toward mobile phone use for the test were analyzed (as seen above).

Usability analysis

As discussed above, the first test was conducted as a pilot test with a small sample of students, to detect faults in the design of the application and to debug and test the viability. After making some corrections such as adapting some aspects of the content and navigability, a second review of the application was made using potential users. In this second test the number of the sample was expanded to 144 individuals in the last year of high school (aged from 17-18 years), all of whom lived and studied in the area of La Oliva-Gandia (Valencia, Spain) (see table 2).

<table>
<thead>
<tr>
<th>Table 2. Students’ school of origin</th>
<th>Frequency</th>
<th>Valid percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 IES Tirant lo Blanc</td>
<td>20</td>
<td>13,9</td>
</tr>
<tr>
<td>2 IES Monduver</td>
<td>22</td>
<td>15,3</td>
</tr>
<tr>
<td>3 IES Veles e Vents</td>
<td>33</td>
<td>22,9</td>
</tr>
<tr>
<td>4 IES Maria Enriquez</td>
<td>29</td>
<td>20,1</td>
</tr>
<tr>
<td>5 IES Ausias March</td>
<td>40</td>
<td>27,8</td>
</tr>
<tr>
<td>Total</td>
<td>144</td>
<td>100,0</td>
</tr>
</tbody>
</table>

To evaluate the usability of the interface, a likert scale questionnaire ranging from 1 to 4 (to avoid indecisions) was used.
Results of the second test

Once collected, the data were processed using the SPSS statistical program. The first part of the test, which related to knowledge of the environment, focused on aspects to justify routine use and availability of phones, adaptation to the environment of the test items, and utility-satisfaction.

<table>
<thead>
<tr>
<th>Table 3. Students’ attitudes toward the mobile-based tool operability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Responses</strong></td>
</tr>
<tr>
<td>Totally disagree</td>
</tr>
<tr>
<td>Disagree</td>
</tr>
<tr>
<td>Agree</td>
</tr>
<tr>
<td>Totally Agree</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4. Students’ attitudes toward the mobile-based tool usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Responses</strong></td>
</tr>
<tr>
<td>Totally disagree</td>
</tr>
<tr>
<td>Disagree</td>
</tr>
<tr>
<td>Agree</td>
</tr>
<tr>
<td>Totally Agree</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5. Students’ attitudes toward the mobile-based tool time facilitator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Responses</strong></td>
</tr>
<tr>
<td>Totally disagree</td>
</tr>
<tr>
<td>Disagree</td>
</tr>
<tr>
<td>Agree</td>
</tr>
<tr>
<td>Totally Agree</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Once the descriptive data had been surveyed and the group statistics had been examined, it was determined that the results were satisfactory as a whole. The results are above 1.5 on average (on a 0-3 scale).

<table>
<thead>
<tr>
<th>Table 6. Students’ attitudes toward other factors in relation to mobile phone use for language testing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Responses</strong></td>
</tr>
<tr>
<td>A mobile-based task design helps me to perform better</td>
</tr>
<tr>
<td>A mobile-delivered test is useful</td>
</tr>
<tr>
<td>Mobiles help to save time in taking this test</td>
</tr>
<tr>
<td>Mobiles are adequate to cope with my needs for this test</td>
</tr>
<tr>
<td>I learned to use the application quickly</td>
</tr>
<tr>
<td>It is easy to remember how to use the application</td>
</tr>
<tr>
<td>I became familiar with the application easily</td>
</tr>
<tr>
<td>I think this is a good application</td>
</tr>
<tr>
<td>It is user friendly</td>
</tr>
<tr>
<td>The application works as I expected</td>
</tr>
<tr>
<td>I would recommend its use to other students</td>
</tr>
</tbody>
</table>
As observed, in general the students valued the use of mobile phones for language testing very positively. The results of these tests have led us to continue with the research, which is still currently being developed.

Advancing toward solutions for the PAU: The OPENPAU project (2012-2014)

In the years between the PAULEX project and the beginning of the OPENPAU project, Spain started experiencing one of the worst financial crises in its history. In that context, the research team of the PAULEX project observed that mobile phones would be a valuable asset in testing oral skills (speaking and listening) and reading efficiently at a low cost. However, the team also considered that mobiles would be inappropriate for writing due to the intrinsic difficulty of keyboard use (García Laborda, Giménez López, & Magal Royo, 2011; Park, 2011). The design principles for developing a mobile phone application were the following (also Keskin & Metcalf (2011):

- Use of video communication with the examiner or a video delivery system if videos are used (possibly the most likely situation),
- Creation of a podcast library for student test preparation,
- Adequate real or deployed time access and adequate connectivity,
- Augmented reality possibilities,
- Mobile Blackboard or a similar platform for test preparation.

The first trials on a large scale are expected to begin by September 2013. The technology is designed to incorporate these conditions. The following sections address this concern. Thus far only the Delphi analysis and a very small sample of research have been complete.

Significant results of the PAULEX and OPENPAU project for m-testing technology

The results hereby presented are mostly related to the observations and research undertaken between the end of the PAULEX project and the beginning of the OPENPAU project. However, the ideas are based on the results from PAULEX and the triangulation of the Delphi method and focus groups. From their reactions and opinions, we concluded that the different kinds of interfaces for mobile devices favor tasks such as speaking and listening, even multiple choice tasks but are rather limited for reading and especially for writing. In the case of online exams with different kinds of tasks, adjustments must be made to navigation and content so that users can feel more comfortable when viewing and inputting information or data. The adaptation of user interfaces for certain tasks in a limited period of time requires prior understanding of certain determining factors such as the physical, functional, and formal accessibility of the application. For example, an interface with a hierarchical menu on a mobile phone is useful for beginning users because the appropriate options can be selected through the presentation of a series of menus. Hierarchical menus require relatively higher numbers of key clicks but this is acceptable for novel users who need help using unfamiliar navigation systems and thus leads to diminishing differences due to technology knowledge and serves to validate the use of mobiles as delivery system (Weir, 2005).

Interface design

The most recent interfaces developed have been designed with specific criteria for users taking the university entrance exam. The fundamental criteria studied in this period were accessibility, ergonomics, and the functionality or usability of the application.

Accessibility

Both projects followed the criteria for technical accessibility for interface design proposed by the World Wide Web Consortium, W3C, so that they reached the largest possible number possible of students as end users including those having visual or auditory impairments. This was achieved mainly by following the Web Content Accessibility Guidelines, WCAG 2.0 (http://www.w3.org/TR/WCAG20). In terms of accessibility of the contents of the university entrance exam, we considered the type of programming language used for navigation and also established guidelines for information access (Nelly et al., 2009). The first applications developed, PAULEX and OPENPAU, were created.
in environments accessible on Internet with access to the contents delivered online by way of contextual menus for accessing the task management area and student management area. As for the exams created for the students, special attention was paid to visual and functional accessibility of navigation during the final exam tasks. The tests permitted us to determine the potential for mobile phones and the testees’ acceptance mobil phones (Magal-Royo, Fajarnes, Tortajada Montañana, & Defez García, 2007). At the same time, they revealed the importance of two determining factors in the development of future applications: the present rate of technological progress of ubiquitous devices, and the adaptation of contents to the restrictive conditions imposed by them.

Ergonomics

The ergonomic aspects examined for the different applications created for mobile devices focused on the visual ergonomics that enabled students to focus easily and effectively on completing the different tasks. Various studies conducted after the experts’ research revealed the need to establish formal visual guidelines for the content of language learning tasks to enable navigation that is directed and transparent (Weining, Heng, & Guoping, 2007; García Laborda, Magal-Royo, de Siqueira Rocha, & Álvarez, 2010). Ubiquitous devices (mobiles, PDAs, smartphones, netbooks, etc.) can be small in size. That is, they have small screens that limit the space for user interface and the information available on them thus writing and reading tasks may have an additional difficulty due to the fact that a global vision of the read or written text is always desirable. In fact, the information shown must be carefully selected and presented so that it facilitates user interaction not only with the device but mostly with the task content. The major problem with the large variety of screens on ubiquitous devices is the direct impact on information access and visualization because no normalized standards have been established so far (Chae & Kim, 2004; Piolat, Roussey, & Thunin, 1997). This problem was of the most significant ones in test validity. In theory, if the test is implemented, the Ministry should provide all the testees with the same mobile phone to avoid biased or unfair testing conditions.

Limited data input mechanisms

These devices are also limited in terms of data entry procedures because of their reduced size. The methods used most often for mobile devices nowadays are the keypad, which has more than one function associated to each key, and the touch screen. Both methods require a high degree of attention on the part of users and can lead to errors, a situation that limits how they can be used.

Thanks to improvements to user interaction mechanisms now found on ubiquitous devices, different channels can be used for data input, which can be simultaneous, synchronized, or combined for certain tasks. In the case of the mobile phones, which have different kinds of exercises (oral, comprehension, writing, etc.), data entry mechanisms can slow down or directly affect completion of the exam, for example, on the reading test, which can involve reading a long text or typing using a virtual physical (Giménez López, Magal Royo, García Laborda, Garde Calvo, & Prefasi Gomar, 2009) (see figure 3). The approach to design this interfaces was taken from the socio-constructivist theory of language that uses images to trigger the testee’s output and the visuals to support and enrich the production. This can be also the case when the videoclips are interrelated in semi-interactive conversation through short questions or even in connection to other user to make dialogues between two testees in which the potential knowledge is visible after reconstructing one’s production (Poehner, 2008). At the same time, a robust recording system and clear rubrics support the cognitivist approach that can be best seen in the long responses for descriptions or the multiple choice responses in which the students need to show evidence of knowledge without external mediation or support (what has been called current knowledge).

Usability

In terms of the usability and functioning of the applications adapted to mobile phones created for the PAULEX project, the results show that the students considered it to be useful because it enabled them to save time while taking an exam of this kind (see figure 3). It was also determined that the students learned to use the application for mobile phones faster and independently due to their familiarity through daily use of mobiles, which enabled them to adjust quickly to it and to its guided interactivity. Analysis of the data related to level of satisfaction with the use of the
application was very high when there was a sensation of predictability that leads to a fast understanding of the method and learning how to do specific tasks on the mobile phone.

Figure 3. PAULEX application on mobiles

The overall conclusion of these first trials was that the students felt comfortable with the format (bearing in mind the limitations of the devices. The oral tasks with video presentations were evaluated with the same degree of confidence and reliability as the analogous activities on the web platform for personal computers.

Figure 4. Usability test of the PAULEX application

Proposals for the design of an m-testing platform

The proposals in this section were also applied to the OPENPAU project and any future project of a language testing m-platform and are strongly based on the findings from the PAULEX project. The OPENPAU project has incorporated the application to the HTC Desire model mobile device whose base technology allows multimodal use of different forms of data input and output (see figure 3). To do so, a study was carried out in advance to determine the initial conditions needed for completion of the tasks on an English language skills exam. These included the
The general format of the application contained the following visual sections:

• The program header area. The application name and official logo of the program participants appear in this section of the screen.

• The user data area. This area is fundamental for the final coding of the exam and student for initial correction and any future corrections, as well as any official reviews required by law at the national level.

• Area for viewing progression through the exam. This area has numbers indicating the different tasks that must be completed on the exam. This section will enable the students to know their progression throughout the exam from the point at which they enter their application access code until they send the completed exam. It starts with the reading of the student's data before the actual completion of the exam and provides information throughout completion of the exam including selection of the interaction mode, and completion of the different tasks on the exam, etc.

• Test area. This area shows the questions or exercises to be completed on each of the tasks. The content will vary depending on the functional and/or content characteristics of the exercises.

• Help area. This section will show general as well as specific information about how to complete the exam including the maximum score assigned to each section.

Project results and conclusions

As observed in the PAULEX project, in situations of high stakes tests with a large number of students, mobiles have some advantages that may put them ahead of other testing systems in terms of budget, accessibility, familiarity, and sound quality. Additionally, although the results in the PAULEX project were limited, the validation methods provided information about the ergonomics, usability, integration, and motivation of the application. According to the data obtained, it was observed that prospective research should include the following aspects:
• Task adaptation to new types of mobile phones;
• Multiplatform systems;
• User satisfaction;
• External validity as compared to other delivery systems and other tests including similar pen and paper versions;
• Technical advances in software design;
• Pedagogical benefits;
• Delivery reliability;
• Functionality.

The students were eager to use mobile phones for language teaching and learning, but they mostly wanted to use them for speaking and listening. Still, the multiple choice items for grammar were also well regarded. However, the students predictably indicated that reading and writing were too difficult to be implemented, with reading rated in a better position than writing (García Laborda, Giménez López, & Magal-Royo, 2011). The PAULEX project also showed that mobiles were excellent for test preparation and an even more encouraging finding is that they offer great opportunities for the real test itself because the students would accept using them for real testing tasks. All three teachers indirectly involved in piloting their use supported mobiles and liked the sequencing and delivery procedure for questions, but they claimed that they had no software up to that time to implement the teaching at a large scale. They also found that, although the testing system could, in fact, be valuable to assess oral skills in the PAU in the long term, phones with bigger screens were desirable. At the same time, they doubted that the Ministry of Education would spend large sums of money on the terminals. However, they believed that the listening and speaking sections could be done online while the rest of the test could be done with pen and paper in order to lower the cost. Additionally, they mentioned that one set of mobile phones could serve more than one high school and maybe more than one year given adequate hygienic measures. Finally, they mentioned the convenience for raters since they could work from a distance either on synchronous or asynchronous testing.

Our experience also determined that technologies for developing user interfaces should focus on the requirement to offer simple interaction modes that are highly natural and adapted to future terminals and communication networks (Oviatt & Cohen, 2000). It is in this area in which technologies face their biggest challenge: attempting to integrate different modes of communication (visual, oral, auditory, gestural, etc.) in order to offer new more powerful methods of interaction with the user, grouped under the name of natural or multimodal interaction, thus overcoming the limitations of interfaces available today (Oviatt, 1999). The ultimate objective of natural interaction is to enable users to be able to use all the communication resources available to them, combining multiple modes of interaction and, therefore, creating a multimodal environment for information access (voice, audio, graphics, video, keypad, electronic pencil, pointer, mouse, etc.) (Oviatt & Larson, 2003). In this sense, the OPENPAU project is currently being driven by practical concerns. The current research is now exploring the potential for implementation and the pedagogical implications while extending the domains of the project to make it a multiplatform one. The study has shown the feasibility of using mobiles for the intended purposes and that the cost could probably be lower than the traditional face-to-face interviews while also permitting a better distribution of space for delivery and adequate rationalization of testing times. Most of the students might also engage in this testing means more easily than in a face-to-face interaction with the examiner. With the development of the OPENPAU application for ubiquitous devices, it has been found that technology has now progressed sufficiently to propose the offering of exams using multimodal access. The incorporation of new modes of interaction such as voice recognition for navigation, the use of touch screens, or synchronized use of the keypad will enable users more comfortable access in accordance with their needs and, thus, solve problems related to accessibility to the media (Magal-Royo, Giménez-López, Pairy, García Laborda, & Gonzalez-Del Rio, 2011; Magal-Royo & Giménez López, 2012).

Progress in the use and research of mobile phones for language learning is receiving increased attention and their use in Mobile Assisted Language Learning (MALL) is an area of steady growth. Despite the advantages this area offers users in terms of the flexibility and ubiquitous nature of the device and environment, as well as advances in mobile applications and Internet access, it must still deal with the need to seek efficient adequate interfaces for user needs for information access and transfer. In the specific case of task completion or specific processes, it is important to evaluate the impact of functional environments that enable users to find comfort and accessibility in the information provided in order to favor this mode of learning.
Future lines of work

The potential of technologies adapted for multimodal interaction in language testing offers huge possibilities for development of innovative applications. In that sense, devices will enable users to select between using one mode or another exclusively (for example, using an online dictionary or making a voice call), to the possibilities of changing between modes of interaction in the same session (sequential multimodality, as in consulting the dictionary on occasions on a mobile during a test), to true freedom in combining and changing modes (simultaneous multimodality: talking, keying, dialing, viewing, etc.) on terminals or ubiquitous devices that enable simultaneous access to voice and data channels, and thus offers opportunities for new items that resemble more what speakers do with the language and how they use it.

In reference to the project impact in the Spanish educational system, it is believed that an inexpensive system to assess speaking skills may have two potential benefits: first, it will enable testing of this skill at a low cost; and second, the impact on the classroom of implementing speaking skills may lead to a great educational improvement in foreign languages. Thus, as a whole, the expected effect of the project if used in the near future is immense and certainly very significant for the educational system.

To conclude, while the use of mobile phones for high stakes testing may be feasible, it is necessary to obtain a commitment from all the stakeholders including the students and the administration authorities. Since the oral test is a social, professional, and educational demand, delivering the oral section of the PAU through mobile phones would require adequate facilities from all the high schools, a better understanding of technology from teachers and new ways to plan and prepare for the test on the part of students. Researchers should also seek ways to overcome the difficulties associated with hearing impairment or other restrictions. While mobiles could be a great asset in education, it is necessary to recognize that not all teachers may be equally prepared to face such as a technological change or eager to change their ways of teaching to cater to the students’ needs by facilitating them with the necessary strategies for taking the test. Thus, practitioners should also receive the necessary instructions and courses to facilitate their adaptation to the new context. Nevertheless, it is believed that this change would not be any more traumatic then others that they have seen in recent years. The ongoing work in the PAULEX project is expected to continue to address these issues. The information obtained so far, while initial, provides enough evidence for the potential of this innovation in both the national and international contexts in areas such as educational planning, course design, test delivery, specifications, and information and communication technologies development. It also takes the use of mobile phones far beyond their traditional perspective of mere supportive elements of courses or learning to enhance their role as high stakes testing facilitators.

Acknowledgements

The authors would like to thank the Ministry of Economy and Competitiveness for funding the research project (with co-financing by ERDF) within the framework of the National R + D + I (2011-2014) "Guidance, proposals and teaching for English section in the entrance examination to the University" (Reference FFI2011-22442). The researchers would also thank the Spanish Ministry of Education, Culture, and Sports because without the grant for the Senior Researchers Mobility this paper would probably have not been possible.

References


Effects of Mobile Instant Messaging on Collaborative Learning Processes and Outcomes: The Case of South Korea

Hyewon Kim1*, MiYoung Lee2* and Minjeong Kim3

1Center for Teaching and Learning, Dankook University, South Korea // 2School of Education, Virginia Commonwealth University, USA // 3Department of Teaching Education, Dankook University, South Korea // khw5780@dankook.ac.kr // mylee@vcu.edu// minjeong69@dankook.ac.kr

*Corresponding authors

ABSTRACT

The purpose of this paper was to investigate the effects of mobile instant messaging on collaborative learning processes and outcomes. The collaborative processes were measured in terms of different types of interactions. We measured the outcomes of the collaborations through both the students’ taskwork and their teamwork. The collaborative learning processes and outcomes in the Mobile Instant Messaging group (Mobile IM) were also compared with the Personal Computer-based Instant Messaging group (PC IM) and the Bulletin Board System group (BBS). A total of 48 students participated in this study, and the main results show that more cognitive and metacognitive interactions were found in the BBS group while social and affective interactions were the major types of interactions in the Mobile IM group and the PC IM group. As a result of the collaborative learning outcomes, the Mobile IM group shows better teamwork than the other two groups. However, better taskwork was found in the BBS group and the PC IM group rather than the Mobile IM group. Finally, the researchers discuss the implications of this study from the perspective of the educational potential of mobile learning.

Keywords

Mobile-based collaborative learning, Mobile instant messaging, Collaborative learning processes, Collaborative learning outcomes

Introduction

Many researchers have claimed that mobile learning will greatly influence the future of teaching and learning in collaborative learning contexts (El-Hussein & Cronje, 2010; Huang, Yang, Huang, & Hsiao, 2010; Ryu & Parsons, 2012). The main reason behind many researchers’ enthusiasm about mobile based collaborative learning stems from its spontaneous, portable, personalized, ubiquitous and situated characteristics (Motiwalla, 2007; Patten, Arnedillo Sanchez, & Tangney, 2006; Rau, Gao, & Wu, 2008; Ryu & Parsons, 2012). Moreover, mobile learning has gradually become stable and mature (Huang, Yang, Huang, & Hsiao, 2010) and has attracted an increased number of learners in recent years.

Educators in South Korea are particularly fascinated by the concept of mobile learning due to its potential to overcome the limitations of traditional education and web-based learning. According to Korea Internet & Security Agency (2011), the infrastructure for mobile learning (e.g., WiFi networks, high-speed internet connection) is well established in South Korea. The Organization for Economic Cooperation and Development (OECD) also recently reported South Korea has the most mobile wireless broadband subscriptions of 34 OECD counties (OECD, 2012). South Korea has 104.2 subscriptions per 100 inhabitants. Additionally, several South Korean universities have distributed free iPhones or smart phones and encouraged students to utilize them to participate in lectures, to access library sources, and to access educational administration system (Lee, 2010). Students’ adoption of mobile technology is not surprising, given recent statistics on Internet usage. The Korea Internet & Security Agency (2012) finds that the internet usage rate for university students is almost 100% (99.9%) and among instant message users, 49.4% use mobile instant messaging services. The rapid diffusion and use of mobile devices suggests students may be receptive to educators’ incorporation of these tools for learning or ubiquitous learning in South Korea (Park, Nam, & Cha, 2012).

However, the true extent of the impact of mobile learning on education is still contested, both theoretically and empirically (Motiwalla, 2007, Ryu & Parsons, 2012). Moreover, previous research is limited to two specific themes – the effectiveness of mobile learning and the design of mobile learning systems (Wu, Wu, Chen, Kao, Lin, & Huang, 2012). Researchers have typically measured the effectiveness of mobile learning using learning outcomes rather than learning processes (Chen, Chang, & Wang, 2008; Hwang & Tsai, 2011). These outcomes comprise motivations, perceptions, attitudes, academic achievement, and satisfaction of students.
In this respect, various research topics that can uncover the potential of mobile learning are warranted to present more practical guidelines in this area. To address this gap in the literature, the present study explores how mobile learning affects collaborative learning processes and outcomes. Specifically, we examine the extent to which students’ cognitive, metacognitive, and social/affective interactions vary in mobile-based collaborative learning environments. We also examine the quality of cognitive messages and the level of team effectiveness in order to measure taskwork and teamwork, respectively.

**Theoretical background**

**Mobile-based collaborative learning in social and situated learning frameworks**

It is important to emphasize that the use of technology in educational settings must be in accordance with educational theories and specific pedagogical considerations (Patten et al., 2006). According to Ryu and Parsons (2012), social and situated learning can be experienced through mobile-based collaborative learning since mobile learning facilitates seamless social interaction in learners by providing them advanced functions such as mobility and instant connectivity. Social learning theory emphasizes that learning occurs within a social context, which means people learn through observing and modeling other learners’ behaviors (Bandura, 1977; Hung, Looi, & Koh, 2004). Mobile-based collaborative learning can maximize the quality and quantity of interactions and observations through its rich communication channels. On the other hand, situated learning theory emphasizes authentic contexts and real learning activities (Lave & Wenger, 1991). Situated learning occurs in educational settings, which provide authentic contexts and activities to promote social interaction and collaboration (Herrington & Oliver, 1995; Lave & Wenger, 1991). Unlike traditional classrooms that decontextualize learners from authentic and practical situations, mobile learning provides a borderless context where learners can reach their goals and needs through real-time interactions. Thus, learners will experience enhanced social and situated learning through mobile learning. Also, mobile learning grounded in social and situated learning will provide learners with more updated learning environments.

**Mobile instant messaging for collaborative learning**

Collaborative learning is defined as ‘a situation in which two or more people learn or attempt to learn something together’ (Dillenbourg, 1999, p. 2). Collaborative learning can be mediated through many different tools, such as discussion boards, blogs, and instant messenger. Like computer-based collaborative learning, mobile-based collaborative learning is mainly text-based, which can enable students to express their opinions and to ask questions without the pressure or feeling of threat that can accompany traditional classrooms (Kitsantas & Chow, 2005; Rau et al., 2008; Ting, 2012). However, Chen & Huang (2010) note that computer-based collaborative learning has a limitation with respect to meeting learners’ educational needs, especially for students who want a more informal and flexible learning environment. In this respect, mobile-based collaborative learning can be more in accordance with their needs by providing ubiquitous and situated learning environments (El-Hussein & Cronje, 2010).

Instant messaging is one of the most widely-used mobile applications for education (Rau et al., 2008). Rau et al. (2008) found that mobile instant messaging supported social bonding between students and instructors. Additionally, Yengin, Karahoca, Karahoca, & Uzunboylu, (2011) investigate the potential of using mobile instant messaging for education, and they found the successful examples such as a quiz tool, an assessment tool and discussion tools in several previous studies (e.g., Attewell, 2005; Stone, Briggs, & Smith, 2002; Markett, Sánchez, Weber, & Tangney, 2006; Bollen, Eimler, & Hoppe, 2004; Holley & Dobson, 2008). Other studies suggested that when used as a discussion tool, mobile instant messaging can promote interactivity and led to more active collaboration (Markett et al., 2006; Bollen et al., 2004; Holley & Dobson, 2008). Despite positive findings from several studies, Ryu & Parsons (2012) and El-Hussein & Cronje (2010) point out that there is still a need to conduct additional research on how mobile instant messaging could facilitate collaborative learning beyond the ‘novelty effect’ of new mobile technology.

**Collaborative learning processes: Cognitive, metacognitive and social/affective interactions**

Mobile-based collaborative learning supports interactions among students as well as instructor-student interactions (Ting, 2012). Students can also enjoy the increased frequency of social interaction through mobile technology in group-based projects (Seppala & Alamaki, 2003). A number of researchers emphasize the quality of cognitive
interaction in learning environments, which is crucial for the success of collaborative learning. However, many researchers note that students' metacognitive and social/affective interactions also play a fundamental role in collaborative learning (Efklides, 2008; Salonen, Vauras, & Efklides, 2005). Metacognition is defined as knowledge about knowledge or the regulation of cognition (Brown, 1987). Metacognitive interaction is regarded as the interactive activities that monitor, evaluate and revise other team member's cognitive processes when they work as a team. They involve the sharing of metacognitive justification, evaluation and feeling (Efklides, 2006). Social/affective interactions are an inevitable part of human communication and play an essential role in collaborative learning (Shen, Wang, & Shen, 2009). Learners express a variety of emotional states (e.g., interest, curiosity and confusion) (Kort, Reilly, & Picard, 2001) as well as social expressions (e.g., greeting, complimenting, and expressing appreciation) (Rourke & Anderson, 2002) when they work together. Furthermore, Panitz (1999) argues that it is important to create an emotional environment that enables students to take initiative in expressing their opinions about any given topic while constructing a shared learning experience.

Interestingly, Ting (2012) suggests that mobile technologies can strengthen learners' interactions and ultimately help learners achieve better collaborative learning outcomes. In addition, Rogers & Price (2006) indicate that mobile technology can change learners' collaborative learning processes, particularly their cognitive, metacognitive and social/affective interactions. However, it is hard to find studies that focus on these specific types of interactions, even though much research has been done on the topic of computer-based collaborative learning (Guan, Tsai & Hwang, 2004; Hara, Bonk, & Angeli, 2000). In addition, Wu et al.'s (2012) meta-analysis on mobile learning using 164 published papers from 2003 to 2010 shows that evaluating the outcomes of mobile learning rather than processes was the most researched topic in the field of mobile learning. Thus, our study, which addresses how these interactions occur in mobile-based collaborative learning environments compared to collaborative learning via desktop computer or BBS, will be valuable to practitioners as well as researchers who are interested in facilitating students' informal or seamless learning by applying mobile technologies to education.

**Collaborative learning outcomes: Taskwork and teamwork**

Unlike individual learning, collaborative learning not only needs task-related skills but it also needs team-related skills that enable team members to work together smoothly and effectively (Eccles & Tenenbaum, 2004). Moreover, a high performance team is characterized as a group of people that is effective in creating a balance between taskwork and teamwork (Johnston, Smith-Jentsch, & Cannon-Bowers, 1997). Mathieu, Heffner, Goodwin, and Salas & Cannon-Bowers (2000) describe taskwork as the skill necessary to accomplish a given task. Taskwork is identified by a learner’s cognitive activity. On the other hand, teamwork is described as the skills needed for effective team functioning such as proper role assignment/responsibility, using efficient communication channels and accurate decision making.

Although many researchers argue that teams develop both taskwork and teamwork through performing their team projects, the evaluation of collaborative learning tends to only focus on their task achievement in terms of how effectively and efficiently they accomplish their given tasks (Mathieu et al, 2000). However, Stott and Walter (1995) indicated that taskwork and teamwork are conceptually independent, but the nature of their functioning is intertwined and affects team performance. Therefore, it is more reasonable to measure both taskwork and teamwork as outcomes of collaborative learning instead of measuring taskwork by itself.

**Research questions**

To examine the extent to which learners’ cognitive, metacognitive and social/affective interactions vary in mobile based collaborative learning as well as the effects of mobile learning on collaborative learning outcomes in terms of taskwork and teamwork, the specific research questions are as follows.

First, are there any differences in collaborative processes in terms of learners’ three types of interactions when they use Mobile Instant Messaging in comparison to Personal Computer-based Instant Messaging and Bulletin Board Systems?

Second, are there any significant differences in collaborative outcomes in terms of learners’ taskwork and teamwork when they use Mobile Instant Messaging in comparison to Personal Computer-based Instant Messaging and Bulletin Board Systems?
Third, are there any differences in learners’ perceptions when they use Mobile Instant Messaging in comparison to Personal Computer-based Instant Messaging and Bulletin Board Systems?

Method

Participants

A total of 48 students in three classes from a large private Korean university participated in the study. All participants were enrolled in an introductory educational technology course which was a required course. Their average age was 21.57 (SD = 13). They participated in the study as part of their regular class activity. The three classes were randomly assigned to one of the following three groups: a mobile instant messaging group (Mobile IM; n = 22), a personal computer-based instant messaging group (PC IM; n = 12), and a bulletin board system group (BBS; n = 14).

Three communication media for discussion

Mobile Instant Messaging (Mobile IM): The Mobile IM group used the KakaoTalk application to conduct their discussion task. It is one of the most popular free mobile messenger applications in South Korea. It provides free text messaging and free calls. The students in the Mobile IM group can share various content and information such as photos, videos, and URL links. Group discussion is possible without the constraints of time and space.

Personal Computer based Instant Messaging (PC IM): The PC IM group used MSN Messenger in their desktop computers. The MSN Messenger is a form of communication over the internet on a PC that offers a quick transmission of text-based messages from sender to receiver. Computer instant messaging basically offers real-time online chat but students need to set a time and to log into the messenger for their group discussion.

Bulletin Board System (BBS): The BBS group used a discussion board system like Blackboard provided by the Learning Management System in a University. Through the BBS, students can do an asynchronous discussion while students can do synchronous discussion through Mobile IM or PC IM. Students in the BBS group are able to revisit their discussion board and post their message whenever they want.

To fairly compare the differences in the three communication media groups, both mobile and computer instant messaging groups are allowed to use only text-based messaging even though they can use voice chatting through their devices. Also, the BBS groups are only allowed to use the discussion board through their personal computer even though they can access it through mobile technology.

Task and procedure

The team task was an ill-structured problem describing a novice teacher who took on a very low achievement class with many troublemakers, and a school principal who directed her to increase student academic achievement within a year. Before students could solve the ill-structured problem, lessons on learning paradigms such as behaviorism, cognitivism, and constructivism were provided to the participants in a regular class. Then, they were randomly assigned to one of three communication media groups. Each group consisted of three or four students and they were asked to discuss a best solution to solve the given problem based on three learning paradigms within a week. All participants were required to discuss the topic using only an assigned communication medium. After the discussion week, participants answered an open-ended perception question which asked them what were the most and the least favorite aspects of the medium that they used for their discussion.

Measures

To examine the effects of Mobile Instant Messaging on collaborative learning, students’ interactions were measured as the learning processes of their collaborations, and taskwork and teamwork were measured as collaborative learning outcomes. The specific methods were described as follows.
Three types of interactions

The content analysis method was used to analyze the types of interactions. As Henri (1992) suggested, an individual theme or idea (thematic unit) was used as the unit of the analysis rather than a word, sentence or paragraph in order to maintain consistency in analyzing students’ discussion messages that occurred in the three different media. For example, the Mobile IM group expressed their opinions in a short phrase or word instead of using a full sentence (e.g., “when?”, “in this case”) while the BBS group usually posted at least one paragraph to state their idea. Therefore, the individual theme or idea was used as the unit of analysis in this study, so the unit of analysis can be any size text from a single word to a paragraph as long as it expresses a theme or idea.

The types of interactions are composed of three categories: cognitive or metacognitive interaction, social or affective interaction, and other interaction. Cognitive or metacognitive interaction is a task-related meaning unit. Social or affective interaction is a non-task-related meaning unit such as personal talks or the expression of feelings. Other interactions are interactions about managing the discussion such as scheduling for the task and setting discussion rules. Two researchers then developed a coding scheme and classified each thematic unit into one or more of the aforementioned categories. The coding scheme is described in Table 1 with samples of thematic units. Inter-rater reliability for the classification of categorical variables was determined by Cohen's Kappa, which measures the agreement between two raters who each classify thematic items into mutually exclusive categories. Cohen’s Kappa for the inter-rater reliability was 0.96 for the agreement of thematic unit and 0.94 for the classification of interaction. The two raters discussed until they reached a consensus, and a total of 1,850 messages were analyzed in this study.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive/metacognitive interaction</td>
<td>Talking about key concepts of learning theories</td>
</tr>
<tr>
<td></td>
<td>Talking about learning theories’ principles</td>
</tr>
<tr>
<td></td>
<td>Talking about implication of learning theories</td>
</tr>
<tr>
<td></td>
<td>Sharing learner’s opinion on learning theories and application</td>
</tr>
<tr>
<td></td>
<td>Speculating some issues on learning theories</td>
</tr>
<tr>
<td></td>
<td>Questioning something about learning theories</td>
</tr>
<tr>
<td></td>
<td>Summarizing what they discussed on learning theories</td>
</tr>
<tr>
<td></td>
<td>Reflecting what they discussed on learning theories</td>
</tr>
<tr>
<td>Social/affective interaction</td>
<td>Praising the other student’s utterances</td>
</tr>
<tr>
<td></td>
<td>Chatting about student’s private lives</td>
</tr>
<tr>
<td></td>
<td>Chatting about non-task-related topics</td>
</tr>
<tr>
<td>Other interaction</td>
<td>Talking about scheduling for the task</td>
</tr>
<tr>
<td></td>
<td>Talking about taking turns</td>
</tr>
<tr>
<td></td>
<td>Talking about setting discussion rules</td>
</tr>
</tbody>
</table>

Teamwork

A survey was used in order to measure teamwork. The survey consisted of five questions about team effectiveness: Efficiency of team management, Observance of team schedule, Conviction of team output quality, Adequacy of team output quantity, and Satisfaction with team output (e.g., “Our team management was efficient,” “Our team members kept our team schedule,” “we think the quality of team output was excellent,” “we think the quantity of team output was appropriate,” and “we are satisfied with our team output.”). Students responded on a five-point Likert scale ranging from “Strongly Agree” to “Strongly Disagree,” depending on how well they thought that the statement described their team effectiveness. The responses were coded in the following manner: strongly agree = 5, agree = 4, not sure = 3, disagree = 2, strongly disagree = 1. The reliability of the survey was .80 for the pilot test and for this study it was .78.

Taskwork

To examine how well learners discussed a given topic, we evaluated the quality of their group discussion. Specifically, we measured their cognitive messages based on four criteria: novelty, importance, relevance, and ambiguity. Among the 10 criteria in Newman, Webb, & Cochrane’s (1996) study, four criteria which measure the
quality of cognitive messages were selected for this study. Two researchers who specialized in educational technology scored each cognitive message as 1 or 0 based on the four criteria described in Table 2. The inter-rater reliability through Cronbach alpha analysis was 0.92.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Descriptions</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novelty</td>
<td>New information, ideas, solutions</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Repeating what has been said</td>
<td>0</td>
</tr>
<tr>
<td>Importance</td>
<td>Important points/issues</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Unimportant, trivial points/issues</td>
<td>0</td>
</tr>
<tr>
<td>Relevance</td>
<td>Relevant statements</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Irrelevant statements, diversions</td>
<td>0</td>
</tr>
<tr>
<td>Ambiguities</td>
<td>Clear, unambiguous statements</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Confused statements</td>
<td>0</td>
</tr>
</tbody>
</table>

Perception on communication media

For a more in-depth understanding of the characteristics of each communication medium, students’ perceptions on medium were measured by a survey which contained one open-ended question. The question asked students what was the most and the least favorite aspects of the medium in their discussion. One piece of paper was given to each student, and they described their thoughts about the given communication medium for 30 minutes.

Data analyses

Content analysis was conducted to examine how the types of interactions were different across the three groups. For the analysis of the comparisons of the three groups in terms of taskwork and teamwork, one-way ANOVAs were conducted. The perception survey data was analyzed qualitatively based on the main themes that students addressed as characteristics of the communication medium they used.

Results

Collaborative process: Types of interactions

This study was designed to discover if there are differences in the types of interactions such as: cognitive or metacognitive interactions, social or affective interactions and other interactions (not included in the two categories) among the three communication media groups. The interactions were analyzed by a content analysis and the percentage of each interaction compared to the total number of messages from each group was discerned. The results are shown in Table 3.

<table>
<thead>
<tr>
<th>Communication Media Groups</th>
<th>Cognitive/metacognitive interaction</th>
<th>Social/affective interaction</th>
<th>Other interaction</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile IM</td>
<td>614(50.00%)</td>
<td>449(36.56%)</td>
<td>165(13.44%)</td>
<td>1,228(100.0%)</td>
</tr>
<tr>
<td>PC IM</td>
<td>205(45.15%)</td>
<td>166(36.56%)</td>
<td>83(18.28%)</td>
<td>454(100.0%)</td>
</tr>
<tr>
<td>BBS</td>
<td>123(73.21%)</td>
<td>32(19.05%)</td>
<td>13(7.74%)</td>
<td>168(100.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>942(50.92%)</td>
<td>647(34.97%)</td>
<td>261(14.11%)</td>
<td>1,850(100.0%)</td>
</tr>
</tbody>
</table>

From the results, it was found that cognitive/metacognitive interaction accounted for approximately 50%, social/affective interaction 37%, and other interaction 13% in the group utilizing Mobile IM in their discussion. A similar tendency was found in the group utilizing PC IM with results that showed that cognitive/metacognitive interaction accounted for approximately 45% of the total number of messages, social/affective interaction 37%, and other interaction 18%. On the other hand, cognitive/metacognitive interaction accounted for more than 73%,
social/affective interaction 19%, and other interaction 8% in the group utilizing BBS for interaction.

In terms of the three types of interactions, the Mobile IM and PC IM groups showed similar results. However, the result reveals that the BBS group had more cognitive/metacognitive interactions and fewer social/affective interactions compared to the Mobile IM and PC IM groups. In addition, other interaction was also lower in the BBS group compared to the other two groups. The results are arranged into a pie chart as follows.

Collaborative outcomes

Teamwork

To identify if there are differences in teamwork scores across the three different communication media groups, the mean of the teamwork score was calculated as shown in Table 4. A one way ANOVA analysis was conducted to see if there were any statistically significant differences in the teamwork scores among the groups. The results are presented in Table 4.

Table 4. Teamwork scores by communication media groups

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile IM</td>
<td>22</td>
<td>4.12</td>
<td>.65</td>
</tr>
<tr>
<td>PC IM</td>
<td>12</td>
<td>3.47</td>
<td>.41</td>
</tr>
<tr>
<td>BBS</td>
<td>14</td>
<td>3.65</td>
<td>.58</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>3.82</td>
<td>.63</td>
</tr>
</tbody>
</table>

From the results above, the Mobile IM group recorded the highest average points with 4.12 in the teamwork score, followed by the BBS group with 3.65, and the PC IM group with 3.47. As shown in Table 5, there were significant differences in teamwork scores between groups according to the type of media. From the results of the Scheffe verification, there was a significant difference between Mobile IM and PC IM at the $p < .01$ level, however, significant differences were not found between the BBS group and the other two groups (Mobile IM and PC IM). That is, the result reveals that the Mobile IM group showed higher teamwork at a statistically significant level compared to the PC IM group.

Table 5. ANOVA analysis of teamwork scores by communication media groups

<table>
<thead>
<tr>
<th></th>
<th>Sum of Square</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3.929</td>
<td>2</td>
<td>1.965</td>
<td>5.803</td>
<td>.006</td>
</tr>
<tr>
<td>Within-groups</td>
<td>15.235</td>
<td>45</td>
<td>.339</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19.164</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Taskwork**

To identify if there was any difference in the taskwork score depending on the type of communication media, the total taskwork score was divided by the total number of cognitive/metacognitive messages in each group of communication media. The results revealed that the mean scores of taskwork were 1.89, 2.59 and 2.62 in the Mobile IM, PC IM and BBS groups respectively. The highest mean of taskwork score was found in the BBS group, and the lowest mean of taskwork score was found in the Mobile IM group.

<table>
<thead>
<tr>
<th>Table 6. Taskwork score by communication media groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Mobile IM</td>
</tr>
<tr>
<td>PC IM</td>
</tr>
<tr>
<td>BBS</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

The ANOVA analysis was conducted to determine if the taskwork score was statistically different across the three communication media groups. The results are shown in the following Table 7.

<table>
<thead>
<tr>
<th>Table 7. ANOVA analysis of taskwork scores by communication media groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of Square</td>
</tr>
<tr>
<td>109.640</td>
</tr>
<tr>
<td>938.589</td>
</tr>
<tr>
<td>1048.229</td>
</tr>
</tbody>
</table>

As shown in Table 7, there were statistically significant differences in taskwork scores between groups. To identify which groups showed the difference, a Scheffe verification was conducted, and its results revealed that there was no difference between the PC IM group and the BBS group, but significant difference was found in the Mobile IM compared to the PC IM and BBS groups at the $p < .001$ level. That is, the taskwork score of the Mobile IM group was significantly lower than the other two groups.

**Perception on communication media**

Based on the results of the student perception survey, the following seven major themes were perceived to be the most or the least favorable aspects of each communication medium for their discussion.

**Theme 1 - Time Constraint:** Mobile IM enabled students to contact their team members whenever they needed them. Most of the students using mobile devices said they liked that it did not have constraining time features of Mobile IM.

“I think our team discussed the topic all day long because we talked whenever we are available. Even though it’s short time... So, I love it. One day, I had a lot of classes but I could read the members’ opinion between the classes and respond them. So, I could keep up with my team members’ discussion.”

However, the students in the PC IM group pointed out that it was difficult to find an available time for all team members to log into the PC IM, even though they did not need to be in the same place.

“It is hard to find the time which all the members can participate because of the differences time schedules. Some of the students wanted to night time and the others didn’t. A student was available on weekend only. It was so hard!!”

**Theme 2 - Limitation of Location:** The students in the Mobile IM group could participate in their group discussion while they were working or moving to another place. Many students in the Mobile IM group pointed out that this was one of the most favorable aspects of the technology.
“I am working at a café as a part time job. Sometimes when there were a few customers, I can respond to the team members’ opinion. It’s the most strong point of mobile IM because if I use BBS or PC IM, I cannot involve the discussion frequently.”

However, some of the students in the PC IM and BBS groups who used PC computers were limited by their location. Students could access chat rooms or discussion boards only where there was a computer available.

“It is hard to get the chance of using computer at the school because the number of the computer in the lab is not enough. So I can access the discussion board at home... sometimes when I came back home lately or I was so tired, I didn’t want to participate the discussion. It bothered me....”

**Theme 3 - Availability for Searching Resources during Discussion:** The PC IM group and BBS group reported that they could search the Internet to find any necessary resources related to their discussion topic while they were communicating. They could refer to any references and write their opinions for as long as they wanted without any sudden interruptions by other students. On the other hand, most of the students of Mobile IM felt it was inconvenient to locate necessary references such as textbooks or articles, especially when they participated in their discussion outside of the home or classroom.

“Sometimes I could not remember the detail of a theory which I learned. I really wanted to back up my opinion.... In that case, I want to search the Internet but it is little bit inconvenient through mobile chat. In a mobile chat, multitasking is exceedingly cumbersome.”

**Theme 4 - Emotional Closeness:** Most of the students in the Mobile IM group claimed that Mobile IM offered a more comfortable and friendly environment where they could talk about private topics as well as their discussion topic using various emotional and social expressions. However BBS and PC IM group students rarely addressed this point as an advantage.

“I usually start to say ‘Hello’ or ‘The weather is great’..... like my team members are besides me. Or sometimes I complain my headache or a lot of papers of the other classes to my team members. I feel free to talk to them about my private stories through mobile chat.”

**Theme 5 - Chance for Careful Thought and Reflection:** Most of the students in the BBS group commented that the BBS enabled them to post their opinions or responses after having enough time to think about the given discussion topic and to review other team members’ postings. Also, Mobile IM students had enough time to review other team members’ messages and provide thoughtful feedback as compared to offline discussion. On the other hand, some of the team members only focused on typing their opinion without reviewing or considering other members’ postings. Therefore, it was difficult for them to have a more convergent discussion in the PC and mobile IM environments. That point was addressed as a disadvantage of mobile and PC IM groups.

“Some members talk so long when we discuss at offline meeting. In that case, I am sure that I forget what he or she was talking about for the first time. However, in the mobile chat, I can review the full message, so I do not forget what I prepared for the comments”

**Theme 6 - Participation of the members:** The students in the BBS group reported that some team members did not frequently visit their discussion board, which resulted in delayed responses and disjointed group discussions.

“A group member didn’t visit the discussion board after his first posting. That was it... I was annoyed that we could not discuss anymore. However, there wasn’t any alternative because we had to discuss through BBS only.”

**Theme 7 - Inconvenience of Using Communication Media:** The most common problem voiced by Mobile IM students was that the relatively small keyboard and screen on their mobile phones constrained them when typing a lengthy opinion or response during their discussion.

“I hate typo but the keyboard of my mobile phone is so small. So, I cannot help mis-typing. Also, the screen is so small. When I read all the discussion, I have to drag the message for a long time. It so irritates me.”
Conclusion and suggestion

Our study contributes by extending the scope of research on mobile learning. Unlike previous research, our study focuses on the effects of mobile learning on collaborative learning processes and outcomes. Social and affective interactions as well as cognitive and metacognitive interactions were also considered as important factors in the collaborative learning processes. Moreover, teamwork that was often ignored as the outcome of collaboration was measured along with taskwork. Based on the results of the study, it is recommended that students use Mobile IM or PC IM in order to facilitate their social and affective interaction at the beginning of their team project when they need to invest in getting to know one another (Lee & Johnson, 2008). Once students have progressed beyond the initial stages of the project, BBS could be the best communication medium to promote students’ cognitive and metacognitive interaction. The results of this study suggest that BBS, PC IM and Mobile IM should be used for different purposes. The BBS and PC IM would be good communication media to improve students’ taskwork while the Mobile IM would be the best choice to facilitate their teamwork. Therefore, understanding the unique characteristics of each communication medium is pivotal to maximize the quality of instruction, and, ultimately, students’ performance.

Future studies are suggested in the following three directions. First, it would be interesting to examine the affective and social aspects of learning as the result of collaborative learning outcomes. In this study, we measured learners’ taskwork and teamwork by focusing primarily on their cognitive development and team effectiveness. However, it will be necessary to examine how much their motivation and attitudes are improved after using Mobile IM for their collaborative learning. Second, we measured learners’ interactions and outcomes that occurred in a one week discussion, but future study is needed to conduct the measurement at least three times to see the change in learner interaction patterns and how their teamwork and taskwork develop over time. According to Fiore, Salas, Cuevas, and Bowers (2003), a team as a cognitive community goes through three coordination phases consisting of pre-process, in-process, and post-process. Depending on each phase of the processes, learners’ interactions and their focus vary. Therefore, it would be a good research topic to examine the change in learner interaction patterns and the development of teamwork and taskwork along with the three coordination phases. Third, from a more practical standpoint, future study also needs to focus on the design of mobile based collaborative learning environments with consideration of the results of this study and provides specific guidelines for effectively and efficiently launching mobile-based collaborative learning in online and offline classrooms. Specifically, it would be interesting to design and develop online instructions using a combination of mobile IM and other online communication tools depending on types of team activities and expected interactions for students’ informal or seamless learning.

References


41


Mobile Inquiry Learning in Sweden: Development Insights on Interoperability, Extensibility and Sustainability of the LETS GO Software System

Bahtijar Vogel*, Arianit Kurti, Marcelo Milrad, Emil Johansson and Maximilian Müller

Department of Media Technology, Linnaeus University // bahtijar.vogel@lnu.se // arianit.kurti@lnu.se // marcelo.milrad@lnu.se // emil.johansson@lnu.se // maximilian.muller@lnu.se

*Corresponding author

ABSTRACT
This paper presents the overall lifecycle and evolution of a software system we have developed in relation to the Learning Ecology through Science with Global Outcomes (LETS GO) research project. One of the aims of the project is to support “open inquiry learning” using mobile science collaboratories that provide open software tools and resources, and participation frameworks for learner project collaboration, mobile data and media capture, publishing, analysis, and reflection. The primary focus of this paper is to report on our technical development, insights and knowledge gained during the past four years. Technical implementations and the prototypes developed in this project have been tested across several educational trials conducted in Sweden and abroad with more than 400 learners. Insights and knowledge gained from these activities verify that learners’ requirements were adequately addressed while satisfying their needs. The outcomes and results of our efforts provided us with a better understanding with regard to which software engineering processes and approaches can be used to address and support the complex requirements that emerge in novel mobile learning scenarios. Thus, the results discussed in this paper provide deeper insights into the importance of properly addressing issues related to interoperability and extensibility in order to develop software solutions to support mobile learning that are sustainable and endurable over time.

Keywords
Mobile learning, Inquiry-based learning, Software lifecycle, User-centered development, Interoperability, Extensibility, Sustainability

Introduction
Web technologies are enabling Internet applications and services to become easily integrated in interactive systems (Holmberg, Wuenzsche, & Tempero, 2006). Thus, the web is gradually becoming a “central computer” that helps to connect diverse computing and data resources and people (Liang, Croitoru, & Tao, 2005; Giusto, Iera, Morabito, & Atzori, 2010). The evolution of these web developments combined with sensor and interactive technologies provide new possibilities for the implementation and deployment of software applications to support a wide variety of human activities.

Mobile and web technologies and applications provide new possibilities for augmenting learning activities. These are Technology-Enhanced Learning (TEL) activities that can be spatially distributed and can incorporate different physical and environmental sensor data (Wu, Yang, Hwang, & Chu, 2008). There are different mobile, web and sensor-based technologies that provide new perspectives on how learning activities can be embedded in different settings and across contexts (Chang, Wang, & Lin, 2009). One innovative aspect of these new learning landscapes is the combination of learning activities to be conducted across different educational contexts such as schools, nature and science centres/museums, parks, and field trips (Kukulska-Hulme, Sharples, Milrad, Arnedillo-Sanchez, & Vavoula, 2009). In these technology rich and dynamic learning environments learners make use of a wide range of devices and applications and the notions of system interoperability and extensibility, become central in order to successfully fulfill the requirements posed by the different educational activities and the learners. Especially, since these aspects directly influence learners’ satisfaction and experiences with regard to the applications and systems they use and also directly affect how tools and applications are adopted, appropriated and sustained over time.

In our Learning Ecology with Technologies from Science for Global Outcomes (LETS GO) collaborative international project (2008-2012), we have been developing, implementing, studying and scaling up novel ways for fostering secondary school student learning in teams for ecological and environmental sciences (Spikol, Milrad, Maldonado, & Pea, 2009). During the last 4 years we have been working with the design, development and implementation of web and mobile services that integrate geo-sensing, multimedia communication and interactive visualization techniques in specific ecology learning scenarios. Our goal has been to create mobile science inquiry...
collaboratories (Pea, Milrad, Maldonado, Vogel, Kurti, & Spikol, 2012) with teachers, learners and developers and domain scientist on topics related to water and soil quality, ecosystems and biodiversity.

One of the main objectives of the LETS GO project was to develop a robust software system including a wide range of applications and services to support educational activities that promote collaborative scientific inquiry as students formulate questions and hypotheses, and collect, analyze, discuss and compare data while studying problem topics in environmental sciences. All our software solutions have been conceived having in mind how to support all these processes. In this paper we present and discuss the overall lifecycle and evolution of the software system we have developed during the last four years. Our choice to focus on these specific aspects is guided by the challenge of how to address those problems related to the scalability and interoperability of mobile learning applications. Thus, the main question we are trying to answer in this paper can be formulated as following: How can software engineering processes support the functional requirements posed by current mobile learning scenarios and applications? The insights and knowledge gained during our research efforts are closely related to the issues of interoperability, flexibility and extensibility of the LETS GO software system and were identified during iterative user-centred development cycles. Developing sustainable mobile applications that can cope with the changing demands of dynamic learning environments requires new knowledge and approaches. The results presented and discussed in this paper provide some new perspectives in this direction.

The remaining of the paper follows with a presentation of the motivation behind our research efforts, as well as the initial requirements, to continue after with an overview of the LETS GO project and its related activities. The following section presents the details of our design, technical solution and the evolutionary stages that were carried out as a part of this development to continue with the Lessons Learned section where we reflect upon the activities we have conducted during these four years of development. At the end, we provide our main conclusions and discuss possible lines for future research.

Motivation and initial requirements

The initial requirements elicitations with stakeholders emerged from a workshop with teachers involved in the LETS GO project that took place in the fall 2008. Different activities in this workshop helped to identify the need to integrate geo-location and environmental sensing, visualization, and Web 2.0 mashup technologies, as part of a broader educational scenario. These requirements identified the need to support “open inquiry learning” for having access to diverse sensor data, live mapping tools, interactive data visualization and collaboration tools, and additional learning resources. Another requirements related to usability, include low cost, using open standards, multiple application support, and support for different types of collaboration modes and contexts (Spikol, Milrad, Maldonado, & Pea, 2009).

Trying to match these initial requirements brought up a number of challenges that concern software tools for supporting inquiry-learning activities. A survey of the literature and existing approaches to support inquiry science learning conducted at the beginning of the project indicated that there were no existing software solutions that could cope with all these requirements at the same time (Vogel, Spikol, Kurti, & Milrad, 2010). More recently, Sun & Looi (2013) report on a review of different web-based science learning environments for collaborative inquiry. The analysis of the results indicate that even those systems discussed in their paper do not cope with the kind of requirements we are addressing in our work. Already at the early stage of our project, those aspects related to the issues of interoperability and extensibility of the system to be developed were identified as one of the central challenges in terms of “building new technologies or further developing existing technologies to create novel possibilities for supporting human activities” (Tchounikine, 2011). Some of the processes that learners need to be actively involved during inquiry learning activities are to problematize, demand, discover and refine, and apply new knowledge and skills to solve complex problems (Edelson, Gordin, & Pea, 1999). Therefore, our primary focus in this project was to facilitate the integration of proper tools (both hardware and software) and services for supporting inquiry based learning activities.

According to Knapp and Barrie (2001), field trips are important to effectively learn about environmental science and they should be actively promoted. It is suggested that field trips can be helpful to generate relevancy to classroom learning when connected with the outdoor environment. For students, such an approach may raise the interest in and aspirations for science-related careers (Rudmann, 1994). The data collected in such field trips play an important role
for analysis, and hence should be saved and carried back to the classroom. Presenting and analyzing these data using visualization tools may help to increase learners’ understanding of complex subject matter.

Reflecting upon our current knowledge and experiences from the field of TEL, two important issues can be identified for supporting environmental inquiry science learning including outdoors and in classroom activities:

1. Providing technological support (in terms of portable instruments and sensors for data collection and software) for field trips activities that include collecting data, and
2. Providing technological support for classroom activities that include visualizing, exploring, analyzing, discussing and reflecting upon the data collected in the field.

Hence, the system support for these kinds of activities needed to include functionalities for mobile data collection and web-based tools and applications for interactive visualizations. These aspects were also in line with the theoretical aspects of scientific inquiry thinking that suggest that this kind of system support has the potential to increase learners’ engagement and curiosity (Pea, 2002). Thus, based on these different requirements we have developed a variety of software tools and solutions to address these different challenges. A detailed description of this work can be found at Vogel et al. (2010), Vogel et al. (2011) and Vogel (2012).

Figure 1 below provides an initial overview of the educational settings related to the LETS GO project and presents some of the initial requirements. Moreover, this figure maps the key processes of inquiry learning activities related to data collection and interpretation, exploration and reflection, drawing conclusions and communicating the results (Edelson, Gordin, & Pea, 1999; Linn, & Eylon, 2011).

As illustrated in this sketch, the initial requirements that guided our research efforts can be specified in terms of the following:

- Mobility of the users/learners
- Distributed environments,
- Service-oriented systems,
- The need for reflection on the collected data and activities, and
- Interactive collaborative technologies.

Some of these requirements are also in line with the current key technological trends identified in the field of TEL: mobile and cloud computing, visual data analysis, web technologies and geocoded data, smart objects and open content, as pointed out recently in the literature (Johnson, Levine, & Smith, 2009; Johnson, Levine, Smith, & Stone, 2010; Johnson, Smith, Willis, Levine, & Haywood, 2011). Indeed, the latest Horizon Report (Johnson, Smith, Willis, Levine, & Haywood, 2011) points out that many technologies used in the field of TEL are increasingly becoming
cloud-based and decentralized. From a pure technological perspective, Hoppe (2009) claims that one of the main challenges we are facing involves the need for integration of diverse technological resources in broader educational scenarios. Therefore, these trends once more reaffirm that the issues related to interoperability and extensibility become central for the integration of diverse technological resources for supporting educational activities. In the coming section we provide an overview of the LETS GO activity flow in order to better understand the interplay between the different learning activities and the technological support.

**LETFS GO activities and testing**

During the four years of the project more than 400 students have been involved in different type of learning activities. These activities included classroom lessons, field trips and lab work and included data collection in the field, taking images and notes, as well as data visualization and discussions in the classroom (see figure 2).

Usually, the participants in the different activities were either students from K-12 schools in Växjö, Sweden or undergraduate students (the teacher training program at Linnaeus University). As part of the environmental science curriculum, they investigated topics related to soil quality (woodland ecology) and water quality in the surrounding lakes. None of the participants in all these activities had prior knowledge regarding how to use the technologies we developed. In our latest pilot activity that took place at the Potomac River in the USA (September 2012), teachers from both Sweden and USA were involved.

**Figure 2. Different learning activities and the technologies in use**

Figure 3 below gives an overview of the learning activity flow and how the different phases of the students’ inquiry process were supported. It should be noticed that these activities were designed according to the different stages of inquiry based-learning as suggested by Edelson, Gordin, & Pea (1999) & Linn & Eylon (2011). Furthermore, these different learning activities have been integrated with their regular curricula in Environmental Science courses at the different schools. A typical LETS GO learning activity usually included workshops for the students to get familiarized with specific subject matter and central concepts and ideas associated with the inquiry learning process. These activities usually comprised six to eight lessons over a period of five weeks starting with the introduction of inquiry process where basic concepts of the activity were introduced; students discussed the initial questions given to them about a specific topic (e.g., water quality). This activity was followed by the preparation for investigation and experiments to be conducted using different technologies (probes, data loggers, mobile applications for data collection in the classroom). Additionally, users conduct field experiments at a local environment and collect samples for lab analysis (see Figure 2a). The data collected using the mobile data collection tool were geo-tagged content and sensor data (usually pH, dissolved oxygen, temperature, conductivity, moisture, etc. depending on the type of the activity). The learning activity usually ended with a discussion about their findings from the field and lab work and an overall class discussion and reflection by using the web visualization tool (see Figure 2b), which tailored different geo-tagged sensor data and digital content collected using mobile data collection tool. In average, (depending of their course schedule) the entire activity was conducted over the period of four weeks across five lessons units. The logistics, as well as the time period of these activities are illustrated in figure 3 below. As it is presented, each one of these lessons units generated a set of functional requirements that the system should support, namely: sensor support, mobile data collection and interactive data visualization. Furthermore, since these activities (mainly for K-12 students) were part of the regular curricula, students were asked to submit short reports on the outcomes of their efforts and reflections after each unit, concluding with a final test at the end of the activity.
Throughout these activities, we videotaped different sessions for later analysis and some of the researchers from our group used a systematic observation sheet during field and lab sessions.

During the lifetime of the project, we actively tested all our developments with school students where we combined classroom and field trips activities. The user trials (prototype testing) allowed testing the software application throughout five development iterations on authentic settings and dynamically changing environments, while new requirements continuously emerged in these activities. These iterations include the release of a prototype and its active testing with the users/learners. The initial two prototypes have been of throwaway type with a single iteration stage each. The last prototype has been of evolutionary nature that evolved through three following iterations cycles. Details regarding these prototypes as well as development iterations are presented in the Table 1.

Table 1 below, provides a detailed overview of the LETS GO field activities conducted since May 2009 until our latest activity conducted in September 2012. This table provides an overview about learner generated content, and the records that were stored in our repositories. It provides a summarized view on how many samples/records learners collected during these activities, as well as number of pictures they stored in our database and server resources. Furthermore, it also provides a rough overview about the number of users that used our software system so far, including information about location of the trials, school they belonged, and type of the activity they were engaged.

<table>
<thead>
<tr>
<th>Prototypes</th>
<th>Development</th>
<th>Deployment and Testing</th>
<th>Organisations</th>
<th>Location</th>
<th>Records in Database</th>
<th>Images</th>
<th>Activity</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Prototype</td>
<td>2009-May</td>
<td>Katedralkolan</td>
<td>11</td>
<td>Near Katedralkolan, Vaasa, Sweden</td>
<td>24</td>
<td>48</td>
<td>outdoor/indoors</td>
<td>soil quality</td>
</tr>
<tr>
<td>1st Prototype</td>
<td>2010-June</td>
<td>Teleborg Skolan</td>
<td>12</td>
<td></td>
<td>8</td>
<td>20</td>
<td>indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>2nd Prototype</td>
<td>2010-June</td>
<td></td>
<td>13</td>
<td>Near Katedralkolan</td>
<td>21</td>
<td>52</td>
<td>Indoor/Outdoor</td>
<td>Water quality</td>
</tr>
<tr>
<td>2nd Prototype</td>
<td>2011-May</td>
<td>Tynneron, Vaasa, Sweden</td>
<td>22</td>
<td></td>
<td>24</td>
<td>51</td>
<td>Outdoor</td>
<td>Water quality</td>
</tr>
<tr>
<td>3rd Prototype</td>
<td>2011-October</td>
<td></td>
<td>23</td>
<td></td>
<td>31</td>
<td>55</td>
<td>Soil quality</td>
<td></td>
</tr>
<tr>
<td>3rd Prototype</td>
<td>2012-May</td>
<td>Lars Kallokkolans</td>
<td>24</td>
<td></td>
<td>21</td>
<td>22</td>
<td>Outdoor</td>
<td>Water quality</td>
</tr>
<tr>
<td>3rd Prototype</td>
<td>2012-September</td>
<td>National Geographic Society</td>
<td>25</td>
<td></td>
<td>18</td>
<td>44</td>
<td>Outdoor</td>
<td>Water quality</td>
</tr>
</tbody>
</table>

Table 1. LETS GO Activities and related data
The evolution of our software system

The LETS GO software system has gone through an evolutionary prototyping approach through the past four years in order to become a stable and robust platform for mobile data collection, aggregation and data visualization. During the prototyping efforts, we extensively used two different development principles: evolutionary prototyping and throwaway prototyping (Sharp, Rogers, & Preece, 2007). The iterative user-centred development cycles of the software system are mapped to our developments, following these two principles of prototyping:

Throwaway prototyping—considers creating the basis of a final product, which is eventually thrown away; however, it remains valuable to construct further evolving ideas related to the final product (1st and 2nd prototype).

Evolutionary prototyping—considers the evolution of a prototype toward a robust final product (3rd prototype). We applied both these principles throughout our development process.

These two principles of prototyping led to the development of an application that was later used in testing. The design and implementation of the three prototypes was made possible by an evolutionary process through several iterations (as already introduced in Table 1 above). Two initial prototypes have been throwaway type while the last one has been an evolutionary prototype.

Figure 4 presents the timeline overview of the different iterative development cycles and stages of the three software prototypes. During these development efforts, we heavily utilized service-oriented approaches for supporting TEL activities in the context of inquiry learning. These prototypes of a software system evolved from being proprietary applications towards combining several Internet-based services to process and visualize the geo-temporal data collected using mobile data collection tools and web. In the first development stage, we mainly dealt with the integration challenges of different technological resources. In addition, the first software prototype has been mainly implemented using static forms for mobile and desktop enabled visualization features and did not provide real-time data representation (and thus was throwaway prototype). During the second development stage, we continued with our integration challenges, however, due to the evolvement of requirements, the challenge of interoperability across diverse technological resources arose. Thus, the second prototype resulted in a combination between more dynamic forms for mobile (XForms), desktop and web technologies and included initial cloud services (Vogel, 2011). This combination enabled data collection in both, online and offline modes and real-time data representation. Finally, during the third development stage we continued to address the interoperability issues, thus the last prototype of the visualization completely relied on an Internet-based environment (real time and online). The final effort during this process was to conduct a user testing study (Vogel, Kurti, Milrad, & Kerren, 2011). The three development stages made our software system more robust as a product compared with the earlier implementation. A usability study was conducted for assessing the web-based visualization tool since it aggregated and presented the entire data collected using mobile devices. Assessment and testing of the web-based visualization tool was an important issue, in order to identify usability aspects that resulted in a number of concrete suggestions for the further enhancement and improvement of. These suggestions were later used and translated into requirements for further development.

![Figure 4. Cyclic prototyping efforts across a timeline related to requirements and prototype testing, mapped with the integration and interoperability challenges](image-url)
The system architecture and implementation

During the efforts mentioned in the previous sections, the technical developments have evolved and changed rapidly, such as in terms of the design aspects, technology choices and implementation, as well as software & hardware components. Despite the rapid changes of such technologies these developments have not been reflected in the changes of the architecture of our software system (Vogel, Kurti, Milrad, & Mikkonen, 2012). The components identified in the initial architecture have been proven resilient to these dynamic changes and requirements. The results described in this paper are an effort to try to tackle some of the challenges described earlier and are an evolution of the efforts we conducted in one of our previous work (see Vogel, Spikol, Kurti & Milrad, 2010). During our earlier explorations and developments, we have proposed and implemented a system architecture that consists of five different blocks aiming to provide logical divisions between the different resources of our system. Figure 5 below illustrates the component view of the system architecture and its potential for expandability with other technologies and external systems.

In the architecture presented in Figure 5, resources are organized into different building blocks that integrate sensors, mobile data collection units, from the server side the data aggregation components used as data and content storage, and visual representation components that utilize diverse web APIs using external services block of this architecture (Vogel, 2012).

The design and implementation of our mobile data collection tool followed with the integration of an open source Java based project into our software system (Anokwa, Hartung, Brunette, Borriello, & Lerer, 2009). This project/solution supported the use of a particular open standard called XForm. XForm is a standard based on a W3C recommendation that is used to build web forms for the easy exchange of data across platforms and devices using XML as the data format. In our case, the design of the forms for data collection used in our mobile applications was developed following the requirements identified from stakeholders, as introduced above. The solution we decided to adopt is based on the Open Data Kit (ODK) which supports various types of data and content inputs, including text, audio, pictures, video, visual codes and GPS and makes it possible to annotate the collected sensor data and content with location metadata (Anokwa, Hartung, Brunette, Borriello, & Lerer, 2009). The use of XForms facilitated data interoperability across a diverse range of devices and applications that compose our system. We have therefore developed several mobile forms that rely on the use of open standards, which provided us with flexibility, fast development and easy adaptation and integration of technological resources for different scenarios (user trials). Figure 6 shows the screen shot of LETS GO data collection tool.

In terms of software development, our major efforts were allocated for the implementation of the web-based visualization tool. The latest version of this tool enables the visualization of different types of geo-tagged content and sensor data collected using the mobile application described above. The web visualization tool utilizes APIs that provide multiple visual representations of the data set available in our repository. These representations allow users to actively interact with graphs, maps, images, and data tables. An initial version of this tool was implemented
completely in AJAX. Spite a user friendly interface and positive feedback from the users, we experienced some performance problems related to loading and processing huge amounts of data using AJAX (Vogel, 2011).

These latest drawbacks inspired the latest version of our visualization tool that has been entirely developed using Google Web Toolkit (GWT). Our latest version of web based visualization tool called GreenLab (see Figure 7) addressed a lot of the requirements generated during our testing efforts by using iterative cycles with users. This tool has become more stable that the previous AJAX version and it has a lot of features for mapping the data automatically, filtering the data based on different criteria, etc. The data visualization process in Green Lab is divided into two stages; first Green Lab selects the type of data collected during the activity with the mobile devices; and second, it presents and visualizes the collected data. In the first stage, Green Lab loads each data type and presents them as clickable buttons. Each type can contain several forms (XForms), which will be loaded after the user has selected a specific activity. In the second stage, Green Lab presents the data retrieved from each form. The filtering of data in Green Lab is located in a panel and the filtering mechanism triggers all active visualization to update, presenting only a certain part of the data set. Each checkbox as such listens for user clicks, triggering an event to update the views. The filtering panel view contains three kinds of options. The first filtering view is based on Organizations (namely participating schools) that follows with the Groups that belong to that organization. Selecting an organization will first filter the data on selected option and also present all groups for that organization, which also allows filtering specific groups. The Attributes filtering view is constituted from each attributes that can be filtered accordingly, by also selecting multiple attributes. Moreover, a single attribute can be filtered also by their values. The Dates filtering view can be used to choose From-To dates to set the time period the user is interested to investigate.

Green Lab contains three different visualization views for presenting the data, which are all resizable. The main visualization view starts up by presenting the data as a table, but has the possibility to switch to bar-, columns-, line-,
or an area-chart by navigating to the icons in the header. The second view contains a Map view, which locates and visualizes all the geo-tagged data collected by the mobile application. The last view contains only a Scatter Plot view, which allows filtering of two kinds of data attributes, by comparing them.

The design and implementation stages carried out during the past four years facilitated the identification of the main features in terms of sustainability of the LETS GO system. The three main salient features that we have identified during our design and implementation efforts can be enumerated as follows: (1) Interoperability of the software and hardware components (2) Extensibility of the visual representation forms, and (3) Sustainability of the software solution. A more detailed view of these lines of action based on what we learned during the last year developing the LETS GO system are presented in the following sections.

Lessons learned

Proprietary software solutions are deployed extensively through multiple platforms such as the web, mobile devices and desktop applications. The use of diverse standards brings new challenges when it comes to flexibility, interoperability, customizability and extensibility of different components that are part of software systems. Continuous evolving web and mobile technologies combined with the changes of the environment result on dynamic and complex requirements that become extremely challenging. As indicated earlier, this paper aims to tackle some of these problems by providing deeper development insights into the issues of interoperability, extensibility and sustainability related to the LETS GO software system and its evolution during the last four years. These insights were gained during iterative user-centred development cycles of a software system designed and implemented for fostering collaborative science learning activities.

Interoperability of the software and hardware solutions

The notion of interoperability constitutes one of the most important principles in system integration (Zeng & Quin, 2008) and refers to "the compatibility of two or more systems such that they can exchange information and data and can use the exchanged information and data without any special manipulation" (Taylor & Joudrey, 2003). During the development efforts of the LETS GO software system, we have in practice been able to tackle two out of four interoperability categories as introduced by Sheth (1998), namely:

- Syntactic interoperability: differences in data formatting.
- System interoperability: heterogeneous systems and applications.

Despite considerable research efforts, achieving interoperability of various sensor readings and mobile devices and various web services has remained an open issue. In connection to syntactic interoperability our focus was related toward making use of open standards for data exchange such as XML, XForms and JSON. While from a system interoperability perspective, the focus was on making use of open software tools with cloud-based services for matching our requirements.

The features of the cloud environment and services made our system more flexible. Initially, traditional desktop-based integrated development environments were employed and the development then gradually moved towards a mashup-pattern that combined different service-oriented approaches. One identified issue was that the rapid speed and evolution of software and web technologies affected the development process and the application itself. From heavily using Internet based services, such as cloud environments and due to the problems encountered from such services, we started to deploy all our developments into our local repositories and environments. By migrating the LETS GO system to our local environment we mitigated the risks and uncertainties of the cloud development environments. Furthermore, this created new opportunities to closely address the interoperability issues of the software components comprising the LETS GO system. During the last year of development we specifically addressed interoperability issues in the “Mobile Data Collection” component and “Data Aggregation” component.

The interoperability issues in the “Mobile Data Collection” component dealt with both hardware and software issues. Since the current mobile data collection tool was tailored only to Android smartphones and having in mind the emergence of HTML5 and CSS3 as well as multiple mobile cross-platform frameworks (such as PhoneGap, Titanium etc.), we have started expanding the mobile collection tool toward iOS and Windows Mobile platforms.
The purpose of this effort will be to widen the base of the eligible devices to be used as mobile data collection tool (so we do not pose any limitations for the use of the mobile application in schools) as well as making them fully interoperable with the rest of the LETS GO system.

Interoperability issues in the “Data Aggregation” component have been addressed from the perspective of data export capabilities. The idea for such development was to make the LETS GO system open and interoperable with similar tools developed in other research projects. The Data Export component we have developed can be described as a middleware that prepares the surveys data stored on the ODK Aggregate server. The main application is a Java servlet that processes calls from clients and replies either a list of available “SurveyTypes,” links/URLs to all the stored forms and their submissions, or the actual data of the submissions. These calls can be processed in the form of a Web API, where the data is responded as JSON. This enables that all the data aggregated in the ODK aggregate becomes available for export using this servlet and in a JSON format. In the current version of the DataExportServlet all the data is read and exported as the clients upload it and stored in the database.

Beside these efforts we have also been working on the development of the form rules as guidelines for design of the XForms for mobile data collection in a form of naming conventions. These naming conventions developed enable mapping of the collected data dynamically to the visualization tool by the form designed based on our guidelines. All these developments enabled our LETS GO system to be fully interoperable with different mobile devices as well as new visualization tools and services.

Extensibility of the software and system architecture

The massive use of mobile and web technologies for data collection purposes produces vast amounts of data. This is another challenging task that requires attention while trying to make sense of all the data generated by users. Therefore, as the amount of available data continues to grow, conceptualizing and developing new interactive tools for visualization becomes an important task to tackle these challenges, for, e.g., seeking new ways of presenting and sorting appropriate and relevant data, or managing and analyzing information (Ackerman & Guiz, 2011). Different visual representations can provide different insights to users by enabling them to observe data in context, to analyze these data and to draw different conclusions by using different analytical approaches (Eiselle, & Weiskopf, 2009; Sedig, Liang, & Morey, 2009). The extension of web-based visualization approaches, along with new forms of interactive collaborative technologies, is constantly growing (Sedig, Liang, & Morey, 2009). Lately, TEL researchers have been taking advantage of different interactive visualization techniques and tools (Linn, & Eylon, 2011). Research in this area indicates that visualizations have the potential to improve learning outcomes, especially related to inquiry science learning (Johnson, Levine, Smith, & Stone, 2010; Edelson, Gordin, & Pea, 1999; Linn, & Eylon, 2011; Pea, 2002). Moreover, interactive visualizations support and increase students’ engagement in scientific inquiry (Linn, & Eylon, 2011; Pea, 2002). In the scope of our work, “learning through collaborative visualization” refers to developments of “scientific knowledge that is mediated by scientific visualization tools in a collaborative learning context” (Pea, 2002). The latest version of the LETS GO system allows for integrating new interaction features provide by multitouch enabled devices and gesture based interaction in a way that we can expand the interaction modes in which learners work with the visualizations.

Recent developments of our LETS GO system include the implementations of two prototypes using gesture based interaction supported by the use of the Microsoft Kinect (Vogel, Pettersson, O., Kurti, & Huck, 2012) and touch enabled interactions facilitated by the use of the Samsung SUR-40 tabletop computing surface (Müller, 2012). Both these prototypes are fully functional and make use of the data already stored at the LETS GO repository. The initial benefit seems to be the fact that these two new interactions paradigms promote collaboration among users while reflecting upon collected data. Figure 8 below illustrates the Natural User Interface (NUI) for the Green Lab application.

NUI Green Lab is a visually driven explorative interactive visualization tool. The tool focuses on a graspable presentation of the geo-tagged environmental data, collected during outdoor activities using mobile devices, in form of digital maps, charts, and images. The application provides a multi-user interface facilitating the synchronous collocated collaboration of at least two users. The interaction makes use of multi-touch interactions on the SUR-40 tabletop computing system and in-air gestures facilitated by the Microsoft Kinect depth sensor as direct input methods. Furthermore, the interface consists of freely movable digital items allowing the users to set up dedicated...
workspaces and explore datasets on their own. Taking all this into account, the main goal was to provide users, such as students and teachers, a prototype supporting and extending collaborative science learning activities. Therefore, an initial usability study was performed upon these two extended developments (multi-touch and in-air gesture). The initial analysis revealed that the study participants achieved overall better results in the multi-touch scenario compared to the in-air gesture scenario. This led to the conclusion that the in-air scenario is not suitable for complex productive workflows, while the multi-touch interaction is. A distinct advantage in this case could be identified when it comes to the visualization of big amounts of data where multiple users could actively be engaged. All the participants approved the possibilities of the collaborative application and liked the visually driven data visualization and exploration as an offset to traditional workstations with single-user input. This overall finding highlights the need for an integration of new interaction technologies and scenarios in collaborative interactive data visualization, and especially in scenarios related to environmental science learning.

The latest activity we recently carried out with regard to “extensibility” was connected to National Geographic Society’s (NGS) GIS tool, which has been designed to support geographic investigations and encourage collaboration between young citizens and researchers. Our software system was successfully integrated with NGS’ GIS platform called FieldScope. FieldScope has been designed to support geographic explorations and to promote citizen science practices in real-world issues. One of the main drawbacks of NGS FieldScope was the lack of uploading data onsite where the data was actually collected. Thus, the main idea was to extend NGS FieldScope by using our mobile data collection tool. This illustrates the notion of extensibility of our software system in the sense of how sensor data and observations collected using our mobile application combined with the Export function from the Data Aggregation component were used to visualize these data sets in other tool such as NGS FieldScope. The activities we conducted validate the flexibility for data exchange and integration that our software system offers. Our software system has been conceived and based upon the notion of an open and extensible architecture (Vogel, 2012). Figure 9 below depicts the screenshot of NGS’ FieldScope that visualizes water quality data collected using the LETS GO system.
Sustainability of our solution

During the last year of development, our efforts were focused into making the current system a sustainable one, so it can be widely used even after the end of the project. Having in mind the problems we experience with cloud environments (as introduced previously), especially on changes on Google App Engine regarding authentication and limitation of the services, we decided to migrate the LETS GO system (i.e., aggregation server) into our own local environment. The entire environment is based on open source software and open standards. Furthermore, by having full control over this environment we foresee a long maintenance of this software solution and its application across different domains that require mobile data collection and visualization activities. The next step will be making our current solution even more accessible, open and usable for a long period of time.

The user testing study described in earlier sections provided us with additional requirements (Vogel, Kurti, Milrad & Kerren, 2011) that were implemented during the two additional evolutionary prototyping efforts (4th and 5th development iterations, introduced in Table 1). These two additional development efforts in total make five development iterations that made our software system extensible and sustainable while new requirements continuously emerged in these activities. These entire processes made it possible to verify that user requirements were adequately addressed while satisfying their needs. Figure 10 below provides details about our continued evolutionary prototyping efforts (as a continuation from Figure 4 introduced above) across a timeline related to requirements and prototype testing, mapped with the extensibility and sustainability aspects that were considered during these two last iterations. The fourth development stage addressed the extensibility challenges as described above. The fifth development stage made us think to make our system more sustainable towards providing an open platform comprised of a rich set of tools that offer flexible mobile and web based applications that can be deployed by users to support data collection, visualization and collaboration.

The LETS GO system has evolved over a four years period from a prototyping system to become a sustainable one with the possibilities to rapidly be adapted with new features and extensions by taking into consideration the rapid evolution of software and web technologies. In addition, we want to emphasize that the modular design and capabilities of such sustainable system have been conceived with the intention to reduce total platform replacements, where the replacement of a certain component or service as well as the extension of the system with new functionalities becomes increasingly feasible into our solution.

Conclusions and future work

The development lifecycle presented in this paper enabled us to gain valuable insights related to different aspects of mobile and web engineering while developing a system for supporting inquiry based learning activities. During the four years of development efforts, three software prototypes were implemented utilizing service-oriented approaches...
that include mobile, web and interactive visualization modules. The main challenges we identified during these efforts were related to integration, interoperability, extensibility and sustainability of our software system while fostering collaborative science learning activities in the field of TEL. These efforts have been tested with more than 400 users in connection to several trials that took place during this period.

The LETS GO educational activities and tools enabled students to learn in a variety of ways that encompass indoor and outdoor activities across locations and time with the support of sensor technologies, mobile devices and web-based tools. The experiences and knowledge gained during these years enabled us to develop the LETS GO system to a sustainable and robust platform for mobile data collection, visualization and collaboration. The user trials allowed testing the software applications throughout five development iterations on authentic settings, while new requirements continuously emerged in these activities. Reflecting upon our latest development efforts and the results presented in this paper the main findings of our research are discussed in the lines below.

Collaborative technologies, which were used in the iterations of this project, facilitated the adoption of a learner/user-centred approach. The learner/user-centred approach has been suggested by Bonk and Cunningham (1998), where they emphasized “the need to anchor learning into real-world or authentic contexts that make learning meaningful and purposeful”. Sensor kits, mobile devices, web services and interactive technologies nowadays provide us with a vast amount of opportunities of embedding learning activities into real world settings. In these environments the real challenge is the need of matching the dynamic requirements that are generated during learning activities. Hence the main contributions that this paper addresses are the development insights while integrating the technological resources and support for successful implementation of the educational activities related to environmental science learning in authentic settings.

Based on the insights gained from our research efforts we consider that by utilizing an extensive prototyping approach, the discussion of ideas, designs, requirements and implementation possibilities with users/learners the development becomes more easily manageable and understandable. For testing the technical feasibility and understanding whether the technology and implementation behaved as expected, agile development approaches based on prototyping (by combining throwaway and evolutionary approaches) were utilized. They offered an easy and communicative way to test it on real time activities and in authentic settings. Furthermore, this approach helps to find a balance between the design and implementation stages by considering the rapid evolution of technologies.

The integration of diverse heterogeneous device environments where learning activities take place must be based on solutions that promote data exchange, integration and reuse. In our research, we have identified that using open standards technologies for data exchange, promotes systems interoperability and extensibility with new features. Clear cases of such approaches have been the extensibility with interactive technologies and services such as NUI elements and NGS Fieldscope. The initial benefit of this approach seems to be the fact that these new interactions paradigms promote collaboration among users while reflecting upon collected data during outdoors activities.

Moreover, the rapid technological changes affect the flow of learning processes and educational organizations. For enabling rapid changes to be smoothly reflected in everyday activities in this area, there must be well-defined processes to ensure the continual refinement of the applications developed. Facilitating the communication between research projects/researchers and developers on the one hand and research projects/researchers and educators on the other hand, are key factors for the success of these interventions and their sustainability. This approach would enable implemented technologies and applications be closely integrated into everyday educational practices, thus maximizing the benefits in terms of the long-term goals, costs, time, and to satisfy learners/educational institutions with their system. A systematic view on those aspects and their implication for developing sustainable software solutions to support mobile learning could lead to a number of potential benefits:

- Standard based systems
- Constant interaction with users/learners
- Incremental development
- Reduced time and costs
- Expandability
- Flexible change of technologies
- Higher usability
- Easy maintenance and sustainability
In summary, all these identified research insights and benefits were gained during iterative user-centred development cycles of a software system for fostering collaborative science learning activities. Moreover, they provided solid foundations in terms of the possibilities of tackling the requirements for supporting inquiry learning in a flexible manner. From a system perspective, these requirements are best fulfilled by using service-oriented approaches that facilitate interoperability through utilizing open source and open standards and by following the evolutionary approach of prototyping. These findings are directly related to software engineering processes aiming to address the requirements posed by mobile learning scenarios and applications. The issues of interoperability and extensibility of the software solution are directly connected with the possibilities of dynamic reconfiguration of learning spaces to respond to learners’ contextual needs. A sustainable design of technological support to meet these needs requires to be closely developed and deployed in close iterations with different stakeholders (including also teachers and students). This approach may help teachers and learners to overcome some of the complexity of the learning activities and furthermore it may promote the seamless integration of physical and digital learning resources.

References


Informal Participation in Science in the UK: Identification, Location and Mobility with iSpot

Eileen Scanlon*, Will Woods and Doug Clow

Institute of Educational Technology, The Open University, UK // Eileen.Scanlon@open.ac.uk // Will.Woods@open.ac.uk // Doug.Clow@open.ac.uk

*Corresponding author

ABSTRACT

Informal participation in science is being recognized as an important way of developing science learning both for children and adults. Mobile learning has particular properties that have potential in informal science settings, particularly outside traditional educational settings. Mobile technologies provide new opportunities for learners to engage with science on the move. This paper reviews the impact of participation in informal science settings on some members of the public using the experiences of the iSpot project as a case study. iSpot aims to create and inspire a new generation of nature lovers by getting people to explore, study, enjoy, and protect their local environment. It facilitates an inquiry learning approach to identification of wildlife with support provided by a community developing round the resource. The iSpot project described here provides evidence of the ways in which informal participation in science can be enhanced by the use of technology. We draw on the findings of two case studies within the project - iSpot Mobile and iSpot Local. These demonstrate particular ways in which location-based activity and mobile learning can be developed and have an impact on the informal learning of science.

Keywords
Informal learning, Participation, Science learning, Mobile learning

Introduction

This paper discusses informal science learning in mobile contexts, and the theoretical framing and development processes used in the creation of iSpot (http://www.ispot.org.uk). It analyses two projects related to iSpot - iSpot Local, and iSpot Mobile – which have developed particular approaches to the support of informal participation in science.

Mobile learning

Informal participation in science is being recognized as an important way of developing science learning both for children and adults (see e.g., Bell et al., 2009). Informal learning (see Trinder et al., 2008) has become an important area of interest for education researchers in recent years. Livingstone (2001) has documented the informal learning opportunities used by adults and the issues which arise in studying such settings. Mobile learning has particular properties that have potential for productive activity in informal science settings, particularly outside traditional educational settings. Sharples et al. (2009) define mobile learning as “the processes (personal and public) of coming to know through exploration and conversation across multiple contexts, amongst people and interactive technologies” (p. 5). Sharples provides some examples of this process including MyArtspace where school pupils use mobile phones to support learning on fieldtrips to museums. In particular, mobile technologies provide new opportunities for learners to engage with science learning. Dierking et al. (2003) have a view of learning as a cumulative process involving connections and reinforcement among a variety of learning experiences and describe informal science education as “science learning which is strongly socioculturally mediated and occurs across a wide range of physical contexts” (p. 109).

Here we discuss the impact of participation in informal science settings where mobility is an asset. The National Science Foundation describe informal learning as follows:

“Informal learning happens throughout people's lives in a highly personalized manner based on their particular needs, interests, and past experiences. This type of multi-faceted learning is voluntary, self-directed, and often mediated within a social context... it provides an experiential base and motivation for further activity and subsequent learning.” (NSF, 2006, Section I, Introduction.)
Increasingly it is recognised that mobile technology can play a part in Citizen Science activities (discussed further below). See e.g., Robson (2012) describing the use of mobile phones in CreekWatch.

It is important to emphasise with Sharples et al. (2009) that in mobile learning what is mobile is the learner. This is important for the topic of this paper: mobile learning of science in informal settings. In the iSpot case studies which follow, the learner is always mobile, sometimes accessing a website from a field location, sometimes using a mobile device but always engaged in location-based learning. Mobile learning in science settings has been studied both in formal and informal settings. There are a range of relevant studies in formal learning (e.g., Littleton, Scanlon, & Sharples, 2012; Chen, Kao, Yu, & Sheu, 2004) but fewer in informal settings. Early examples of studies which demonstrate the potential of mobile learning in informal science settings include that of Clough (2009). She describes developing mobile support for nature trails, and researching the use of mobile technology with GPS in the geocaching community. It is a challenge for learning scientists to develop and study learning in such completely informal settings.

**Approach to development, theoretical framing of the design and methodological challenges**

In this section we set out the development of the design of the iSpot project. A core group (Jonathan Silvertown, Martin Harvey, Richard Greenwood and Doug Clow) led the creation of iSpot, the iSpot website, and generated its initial design by informal discussion, based on the expertise they brought to the project, which included field biology, citizen science, online learning, and software development. Some iSpot team members were driven by theories of participatory design. An initial motivation was the exploration of applications of geographically referenced teaching and learning. Next we compare the features and intentions of the work with theoretical perspectives from research on learning.

iSpot supports a community of practice (Lave & Wenger, 1991) where members learn from legitimate peripheral participation (Wenger, 1998) and develop their expertise through a process close to apprenticeship. A central theoretical design principle for work on communities of practice is the support of different modes of participation. Preece and Shneiderman (2009) set out a ‘Reader to Leader’ framework, categorising successive levels of social participation in online communities as reading, contributing, collaborating and leading. Other work suggests that a developmental model is not a good fit with observed activity in online learning sites: rather, different users participate in different ways at different times (as described in the 'Fairy Rings' model see Clow and Makriyannis, 2011).

One way to consider a contribution on the iSpot website is as a shared social object (see e.g., Knorr-Cetina, 2001) which can structure this participation, and scaffold participation in the community of practice. iSpot also reflects the constructivist notion of authentic learning activities (Jonassen, 1999) together with what Scardamalia and Bereiter (2006) describe as knowledge building: The learning activity is not only akin to scientific activity, it initiates learners into the knowledge-creating culture and enables them to actively contribute to scientific knowledge.

The development of a system such as iSpot needs to combine, in a cyclic approach, research, pedagogical design, and technology development. Accounts of socio-cognitive software design (Sharples, Taylor & Vavoula, 2007; McAndrew, Taylor and Clow, 2010) are influential in developing such processes, as are principles of Agile software development (http://agilemanifesto.org). Substantial engagement and envisioning activities with stakeholders were conducted, followed by deployment of the system to gather feedback from users.

There are a number of definitions of design-based research: our approach was in line with Barab and Squires’s description: “Design-based research [...] was introduced with the expectation that researchers would systemically adjust various aspects of the designed context so that each adjustment served as a type of experimentation that allowed the researchers to test and generate theory in naturalistic contexts” (Barab & Squire, 2004, p. 3). In some aspects of our project, particularly the development of the mobile app this included iterative cycles of designing (both pedagogy and technology), running an inquiry, and then evaluation and analysis that fed into the next cycle. Thus some of the key findings of the research become embedded within the system: not just in the design of the software, but in how it is used by the growing and developing community of practice.

So the approach taken in the design of iSpot was the co-design of technology and pedagogy i.e. to design the educational activities and technology together, drawing on a participatory design approach (see e.g., Penuel, Roschelle & Schetman, 2007).
There are methodological and practical challenges associated with developing an understanding of how learning takes place in the communities which use iSpot. The learning episodes which involve a user can be relatively short and informal. An important perspective on learning that comes from the public understanding of science movement is to think more broadly about the impact of engagement. Relatively simple models of learning, such as the deficit model used at first in work on the public understanding of science, were replaced by an investigation of the potential outcomes, including increased awareness and impact on attitudes, as well as engagement and participation. Groups enabled by technology will form round particular interests and issues suggesting a need to assess how expertise can develop in these groups. There is a complexity to examining such learning settings as iSpot. We need to look more broadly at them, in terms of new data and analysis methods (Scanlon, 2012, July).

**Citizen science**

In considering the learning which takes place through participatory science enabled by the use of mobile technology in field settings, it is necessary to look for some different ways of examining those learning settings and the use of mobile technology. Dron and Anderson (2007) describe how online communities enable different types of participation in the form of groups, networks and collectives.

iSpot may be described in Dron and Anderson’s terms as a network. It also can be seen as an example of citizen science. Wiggins and Crowston (2011) provide a typology of citizen science projects where members of the public work in combination with researchers. Hand (2010) and Newman et al. (2010) caution that additional verification may be necessary on projects which involve citizen scientists. Rotman et al. (2012) surveyed volunteers on ecological science projects to find out their motivations for participation, and many cited their desire to increase their scientific knowledge.

**iSpot**

This section describes iSpot, the project at the heart of this paper. This paper uses the iSpot project to examine the ways in which informal participation in science can be enhanced by the use of technology, and in particular ways in which location based activity and mobile learning are developed in the project. iSpot allows an inquiry learning approach to the identification of wildlife with support provided as part of a community of practice. It is important to note however that in what follows we are drawing examples from approaches taken in the particular case studies, rather than describing the whole cycle of development in the iSpot project or all the particular design decisions taken to develop its website.

iSpot (McAndrew et al., 2010; Woods & Scanlon, 2012) aims to create and inspire a new generation of nature lovers by getting people to explore, study, enjoy, and protect their local environment. The iSpot web site (home page shown in Figure 1), launched in June 2009, allows users to post observations of animals and plants on the site, and the iSpot community helps to identify them reliably. As a web-based system was used, this allows users to access and learn 24/7 and at anyplace with Internet access. These observations constitute the 'shared social object'. Support is provided for identification partly by online resources but more fundamentally by the community of practice active on the site. The site connects together informal novice learners with experts in a wide range of natural history fields, including over 100 who are representatives of natural history organisations. Learning the name of an organism you have observed is the first step in learning more about it. Furthermore, the process of recording observations of species - including the name of the species, the location and the time of the observation - is the fundamental unit of activity in biodiversity monitoring and research. Indeed, selected observations from iSpot users are now used as part of formal biodiversity monitoring. Thus iSpot enables learners to engage in Scardamalia and Berieter (2006)'s knowledge building: they contribute to new knowledge, as a community activity.

A key feature of iSpot is its sophisticated but easy-to-use reputation system, which provides an indication of each user's expertise on the site (see Clow and Makriyannis, 2011). Unusually among online reputation systems, as well as providing an indication of “social” reputation on the site, the iSpot reputation system includes elements designed to provide sound indicators of the expertise – or learning – displayed through activity on the site. The reputation system structures and makes manifest expertise, facilitating learners' development within the community of practice.
iSpot findings on participation and learning

The impact of iSpot can be seen through its wide reach. Currently it has over 31,000 registered users who have added more than 200,000 observations with over 340,000 images, identifying more than 6,900 different species. The project has identified two species previously unrecorded in the UK: a bee-fly (Systoechus ctenopterus) and euonymus leaf notcher moth (Pryeria sinica). Further empirical analysis of learning activity on iSpot is underway, but some initial findings are presented here.

Qualitative analysis shows clear examples of users who start as complete novices, but come to fairly sophisticated understanding of identification. There is also quantitative evidence of users learning. For instance, analysis of a sample of 407 users as they progressed through submitting and identifying their first fifty observations within iSpot is strongly suggestive of learning. As shown in Figure 2, users showed improvement in their ability to identify other people’s observations over the period that they submitted observations: As users progress from their first to their 50th observation posted on iSpot, they have a bigger percentage of correct identifications that is they are more likely to identify what they have seen for themselves.

Figure 1. iSpot home page

Figure 2. How people improve in identifications from repeated use of iSpot
The crowdsourced identification model within iSpot, rewarding improvement in ability to identify observations, provides some of evidence that people are learning and improving their understanding of nature through iSpot. However a person may gain reputation through identifying very common species and without expanding their knowledge of other species.

In order to get a better understanding of how and whether people learn from using iSpot we require empirical evidence of improvement in people’s ability to identify a greater variety of observations as their reputation improves. We designed the iSpot intelligent quiz to test this knowledge. The quiz was launched in July 2013, since then around 350 people per week have taken one or more quizzes, so an average of around 50 people per day. The quiz is tailored to the level and subject area that people request when they start a new quiz on iSpot. The reputation level that iSpot provides is a good indicator of the level that people should take but there is no restriction on the level so, for example, a level five expert could take a level 1 quiz and vice versa. The data from the weekly logs shows however the people are averaging about 7 out of ten for quizzes across the skills levels which suggests that people are naturally finding a level which challenges them.

![Figure 3. Screenshot of the iSpot intelligent quiz](image)

The quiz has a number of different types of question that test a range of knowledge within a specific domain, some questions are multiple choice and others are about entering the correct name or type of observation. The data collected so far indicates that people who use iSpot are gaining knowledge about nature.

Face-to-face outreach work has reached over 55,000 beneficiaries, over 10,000 from hard-to-reach groups, whilst over 800 participants have used iSpot at local “bioblitz” events, including schools, local government and voluntary sector organisations.

This account of iSpot provides the framing for the description of two specific projects linked to iSpot that particularly explore mobile and ubiquitous learning: iSpot Local and iSpot mobile.

**iSpot mobile**

The first case study linked to iSpot is iSpot Mobile. The iSpot website was already available to be viewed on mobile phones. However since people are outdoors making observations, there was both a need and an opportunity to use mobile phones with digital cameras to make observations and interact with the iSpot community.
The iSpot mobile design approach

A lightweight contextual design approach to establishing the requirements for the mobile app was taken based on the user-centred design process developed by Beyer et al. (1998), exploring the types of users, the scientific context of nature study, the environment which they would be exploring and the learning outcomes to be achieved. We defined the main purpose of developing the mobile app as allowing users to create and upload observations (a combination of photo, identification, and location) to iSpot using their mobile device and to become part of the iSpot community using tools for sharing information. The secondary purpose was to enable iSpot website functionality on a mobile device in a native format and using the enhanced capabilities of a multi-touch mobile phone. For example the ability to pinch to zoom on images to see greater detail and the ability to use the devices to interact with the iSpot community whilst on the move and to enhance their experience through utilising the geo-location services available within mobile devices.

A core group consisting of Jonathan Silvertown, Martin Harvey, Will Woods and Richard Greenwood produced the initial design. A light-touch user-centred design approach was used for app development, beginning with a storyboarding process using experiences from users of the current iSpot website. Specifically, data was collected from a small selected group of experienced iSpot website ‘volunteer’ users whose practice was monitored through interviews and forum discussion, taken alongside usage data from the website, and feedback from the core group to establish common patterns of use. These were converted into stories to build a coherent functional specification. For example, one user said “I am running an inquiry based learning project and I want to [use the iSpot mobile app to] develop scenarios around ecosystems that I’m observing, for example birds in my garden, population of bugs in my flower bed, fauna in my pond ...”

The user-centred design approach that the team adopted involved gathering feedback about how people engaged with early prototypes of the environment to inform later iterations. Twenty people volunteered. A small number of volunteers from the existing iSpot community were also invited to participate, including iSpot “mentors” (associates who work with iSpot to assist others in identifying observations). A series of usability and accessibility testing cycles were conducted during the course of the app development. The feedback was gathered and interpreted by the project team to help improve the functionality and design of later iterations.

First iteration

The first iteration of the app started in October 2011 and took a total of ten weeks including development, bug fixing and testing. An initial issue was that Android devices have all manner of shapes and sizes of screen and this made the display of the images a challenge. Figure 4 shows a screenshot of the observation list.

Testing took place over a two-week period which included an evaluation conducted by a usability expert. The results indicated that the app was missing some critical functionality and had a number of bugs.

![Figure 4. Screenshot of original iSpot app design](63)
The application was also provided to a group of ten experienced mobile users. For example, one experienced iSpot user suggested a process of checking and validating an observation using the mobile app, producing the following scenario: “What iSpot offers is an authoritative resource for helping people learn identification. The new app could be like having an expert out in the field with you which is, I'm sure you'll agree, the best way to learn identification; in the field not through photographs.” To test these scenarios, users were asked to go out and take observations in naturalistic settings and then gain identifications from the iSpot community and then to provide feedback on this experience.

From the feedback it was clear that people were generally enthusiastic about the functionality of the app but they were less positive about the interface design. For example, here is a quote from notes taken during an interview with one of the testers:

“She thought she had to put something in the scientific name or the common name and did not realise that she could leave these blank (she knew it was a ladybird but there was not the option to say just ladybird so she selected one of the named ladybirds, a 10 spot one, even though she knew it was wrong just to get to the next screen and submit the observation).”

The iSpot service is distinctive from competitors as it references species dictionaries and because observations are identified by the iSpot community, often within a very short time of being observed and uploaded: half of all un-named observations are identified within an hour of appearing on the site. The app therefore provides these unique services to mobile users, allowing them to have observations identified and potentially to identify and agree with the identification of other people’s observations.

The evaluation process established that the service created for iSpot Mobile largely mimicked the iSpot website navigation and the design felt quite sterile. The team concluded that the app should therefore be completely redesigned around a navigation and layout more suitable for a mobile app, increasing the interactivity and social elements.

**Second iteration**

In January 2012, a second iteration of application specification, design and development took place. A mobile interface designer worked alongside the developer to implement a set of improvements to the interface.

This design iteration involved providing a big button menu screen as the ‘home’ screen to get into the main app functionality (Figure 5). The designer created a stylised logo and incorporated design features of the iSpot website to improve the app and make it feel more nature related by using grass and wildlife within the layout.

![Second iteration of iSpot app](image)

*Figure 5. Second iteration of iSpot app*
The redesigned tool the users directly to the observations. We made the observation thumbnail images larger to increase usability and aesthetic appeal (see Figure 6). To avoid removing valuable screen ‘real estate’ on what is a small screen we explored using a dynamic menu which users could click on or swipe to view and which provided all the functions within the application, allowing extensibility using horizontal swipe to access menu choices.

**Figure 6.** The image-centric design adopted for the beta release of the iSpot app

Further testing was conducted with another group of ten mobile proficient users. This interface received positive feedback from users, including the design, with comments such as:
- “Pull down icon menu intuitive once you try it for the first time”
- “Tried taking photo of pot plant and identifying it. Intuitive interface. Easy to add details. Recognised my location. Though somewhat cramped with keyboard. Pleased to see my first observation appear on iSpot.”
- “Overall I have found the app to be extremely stable, easy to navigate and fairly intuitive.”

However, there were still concerns that the navigation was not providing rich interaction and direct engagement, and that this interface design was not scalable, i.e., as functions were added how would they be incorporated into the fixed four button menu?

As a consequence of the positive feedback from both user testing and technical testing the team felt in a position to move towards releasing the beta version of the app to the public. The Android iSpot application “stable beta” was released to the public via the Google Android app store (Google Play) on 8th June 2012.

**Third iteration**

The third iteration of development began in August 2012. This iteration incorporated improvements to the application through the feedback gained from the testing processes, through user feedback from the beta release, and through use of enhanced reference material from Google on designing for the Android Platform (http://developer.android.com/design/index.html). The beta app on the Google Play store received positive feedback from the public.

The third iteration included enhancements to the geo-location services to provide “around here” information about observations within a specific locale, i.e., within a 1 kilometre radius of the current location using the GPS capability of the device. Users can also scroll to move the map location and receive information about observations within a 1 kilometre radius of any location. There are enhancements to the social and community aspects of the application, in particular allowing users to identify other people’s observations as well as comment on them. Finally, there are improvements to the discovery and filtering services, to filter on species type, to allow users to quickly find out information related to a particular observation, and to create their own individual journeys of self-discovery.
The full release, as a consequence of the testing and evaluation, provides a richer and more interactive experience with an improved user interface, including a contextual “active menu” and larger images, as shown in the sample screenshots below (Figure 7).

Figure 7. Sample screens from current iSpot app development showing (1) “active menu” and text overlays on images (2) The slide out navigation panel (3) The post comment and post ID capability

iSpot mobile testing and evaluation

A further round of comprehensive testing was conducted prior to release using the state-of-the-art mobile eye tracking and mobile data capture facilities available within the Open University Jennie Lee Research Labs (http://jennielee.open.ac.uk). The app was judged to be more robust, fully featured and a better user experience. For example comments included:

“[The] ‘Around here’ map showing locations of observations in my immediate vicinity seems clear …and easy to use”

After the testing and feedback, the version 1.0 product was released to the Google app store (Google Play) in December 2012. It is achieving over 1000 installations per month and currently has a user rating of 3.8 out of 5 (27 September 2013).

As learners become more mobile, the mobile apps may become the default way of engaging with iSpot and establishing participatory science learning journeys. The app may prove particularly suitable for individuals or groups engaging in local community bioblitzes. The iSpot team expect to use the app to support local group learning activities of this type in the future.

iSpot local

The iSpot Local project extended the iSpot approach to investigate the potential of using hyper-local events to frame the learning activity, moving it from a largely virtualised activity (on the iSpot website) to a grounded, community, mobile setting – including beyond the reach of electronic networks. This built on and extended the existing community of practice and knowledge building approach.

Bioblitzes

The key mediating event in iSpot Local was the bioblitz, a survey of the wildlife at a particular site at a particular point in time - say an afternoon, or a day. The general public, supported by a team of experts, try to identify and
record as many different organisms as they can within the time. This can generate real scientific data (knowledge building) as well as engaging the public in the scientific process - and the site itself - through active participation and learning within a community of practice. However, in traditional bioblitzes, it can be difficult to manage the data generated by the public, and identifying the species observed is problematic.

iSpot Local addressed these challenges by coupling the wider perspective and observational recording abilities of iSpot with hyper-local engagement with community stakeholders and effective practical management of the bioblitz events. The basic activity of iSpot Local is set out in Figure 8, a cartoon developed to explain the bioblitz to participants.

Six bioblitzes were organised across the South West of England, at a range of sites from schools to more public sites. The IT facilities available on site ranged from a high-speed wifi network and room full of dedicated computers (at a school) to a nature reserve with no power, no network, and negligible mobile phone voice signal. The team used a hybrid, flexible approach to technology to maximise the benefits given the nature of the site, typically using a set of laptops in a marquee to log photographs and observations, which were uploaded to iSpot later if connectivity was limited on site.

An important feature of iSpot Local was the way in which mobile access - mediated, supported and contextualised - enabled the hyper-local (the individual bioblitz) to connect to the worldwide (the international community network of experts and enthusiastic amateurs on iSpot).

![Figure 8. Cartoon produced to help explain the iSpot Local approach](image)

**Development approach**

Engagement with a wide range of stakeholders was critical to the success of the project. The funded project partners were the UK Open University, Ambios Ltd (a small not-for-profit company promoting environmental understanding) and Learning South West (a membership organisation coordinating learning and skills and youth work, with members including local authorities, colleges, private training providers and voluntary sector organisations). In line with the design approach, the project partners engaged extensively with many other stakeholder organisations including adult educators, family learning specialists, local government, volunteers, technical specialists and natural history experts.

The project developed and validated a three-phase model for ensuring effective participation including pre-bioblitz work, the bioblitz itself, and post-bioblitz activities. Thus the participation in the on-site activities was scaffolded and embedded through framing and linking activities, enabling the learner's participation in the community of practice.
In addition, as part of the iterative approach to development, some technical development was carried out to create a module to enable observations from a bioblitz to be embedded within a community website, part of which is shown in Figure 9 below.

![iSpot Local map showing observations at one of the bioblitz sites](image)

**Figure 9. iSpot Local map showing observations at one of the bioblitz sites**

### iSpot local evaluation

Evaluation started at the beginning of the project, with the production of an initial Evaluation, Dissemination and Mainstreaming Action Plan. As described above, the design-based research approach meant that much of the outcomes are embedded within the practice developed as the project continued.

The bioblitz sites were generally open for participants to come and go as they pleased, and the participants were very diverse, from primary school-age children through to retired-age adults in to their 80s, and with previous experience of nature ranging from negligible to expert naturalists.

To engage with this diversity, a range of evaluation methods were employed to supplement analysis of the online activity, including observation, registration cards and evaluation cards during the bioblitz, and follow-up discussions with selected participants and stakeholders (e.g., school teachers).

Participation levels recorded through Registration Cards were high at the events, with significant data uploaded to iSpot. There was also a high level of activity by the wider iSpot community, with tentative identifications arising from an iSpot Local bioblitz rapidly translated into confirmed identifications on the website, and high numbers of others indicating agreement and posting further comments.

In total, 820 people participated directly in the six iSpot Local bioblitzes, making more than 1,800 observations in the course of the bioblitzes. On iSpot, these observations received over 2,000 identifications (some observations had more than one identification), from the bioblitz participants and the wider iSpot community. These identifications in turn received over 3,000 agreements. Most participants (74%) were children under 18 and the rest were adults. The gender ratio was roughly equal (53% female, 47% male). As a result of this engagement and participation, the iSpot reputation system was able to confirm over 1,250 observations as having a “Likely ID” confirmed by sufficient expertise. This is clear evidence of the participants taking part in genuine “knowledge building”: despite the diversity of their initial expertise, they were able to jointly contribute to new knowledge.

The feedback from the participants shows further evidence of the participants' engagement and learning—for example, the feedback from the evaluation cards included a parent reporting that the best bit was “Watching my children get so involved, questioning and learning about the world around us that we don't always stop to appreciate.” Another participant reported that they gained “A better appreciation of just how much wildlife lives alongside us in the school field. Brilliant experts, really approachable.”
Engagement with iSpot Local motivated many participants to engage further with learning about nature—for instance “Examining and cataloguing and drawing the wild flowers in the lane to Granny's house” and ‘Have looked at iSpot site and held our own mini bioblitz in the garden.” The diversity of the participants was reflected in the diversity of outcomes from the activities: at one site, the volunteers engaged in conservation work planned to run repeat events annually to track the effects of their work; at another (a school), follow-up learning events targeted to the children’s interests and the curriculum were developed.

The use of bioblitz events coupled with the iSpot website, in the context of a wider learning community, shows the potential for the iSpot website to support a vision of mobile, ubiquitous and lifelong learning at many levels, harnessing the power and range of a global, broad network of expertise with local concerns and knowledge.

The individual learner, located in a particular environment, was connected to multiple potential sources of learning, ranging from informal personal contact with experts through to technology-mediated access to explicit learning resources and relevant formal education opportunities. This rich environment structured their apprenticeship within a community of practice, and enabled them to engage in knowledge building.

**Conclusions and lessons learned**

The paper has drawn on an evaluation of the iSpot Local and iSpot Mobile projects to consider evidence on the impact of participation, and on which features can be identified as important in the design of such community projects. In particular, explicit attention was paid to how learners can be supported to be members of a community of practice, with participation structured around a shared social object, engaging in knowledge building as active contributors to knowledge, and the iterative, integrated approach to development have all proven valuable.

The overall experience of the iSpot project with its analysis of the improvement in identification as users become more experienced provides some evidence of knowledge development. However it is also possible that the users are becoming more proficient with the system so there is room for further investigation of how learning taking place with the system.

We know from the analytics of people’s progress through iSpot that they appear to be improving their identification knowledge as they become more experienced users of iSpot. The new quiz service within iSpot, also available on mobile, will provide further evidence to assess whether learning is taking place.

iSpot Local and iSpot Mobile are two elements within a comprehensive roadmap for iSpot development, with a full range of objectives to be achieved within the project through to the end of 2014. These include the internationalisation of the service, extensions to support integration with other systems and services (Facebook, mobile, species dictionaries), improvements to the service robustness, personalisation and the ability to support local and regional content to create an adaptive user-centred service and services to further test the learning that is taking place through analytical tools and intelligent quiz to track how users are increasing in their ability to identify and understand nature.

The iSpot Local project and the iSpot Mobile app are examples of mobile learning: the learners access iSpot in a range of contexts, settings and locations as appropriate to their individual situation. To a degree they are also an instantiation of the vision of context-aware ubiquitous learning (Hwang et al, 2008): the location of an observation is a crucial piece of information on iSpot, and it is possible to use location-aware sensors to capture this data automatically from the learner’s context.

Each case study illustrated a different facet of informal participation in science and contribution to knowledge in this area. The iSpot Mobile case study demonstrated the impact of a design based research approach to the development of such systems. The formative feedback also provided us with information on the processes by which the mobile app would facilitate learning. The iSpot Local case study showed how online communities of practice can be extended and connected to physical locations, providing more contextual opportunities for knowledge building, and co-creating new scientific data.
This experience and analysis demonstrates some of the potential for mobile and ubiquitous technologies to support learning in informal contexts but certain issues remain. These new developments include an iSpot site created for Southern Africa, managed by the South African National Biodiversity Institute so is showing evidence of how the initiative can be translated into new settings. Also, iSpot has linked up with Treezilla an ambitious Citizen Science project to map all of Britain’s trees and record vital data about tree disease and the environmental benefits that trees provide, developing a mobile app for use as part of the Open Science Laboratory initiative. The sustainability plan involves moving the infrastructure to managed cloud-hosting over the next twelve month period and for moving support for the application and services to the central IT department over the next two years, to be completed by July 2015. This underwriting of such a research system demonstrates the understanding of the importance of iSpot to the Open University and to the growing community that it supports. The iSpot team led by Jonathan Silvertown works in partnership with nature organisations (currently more than 100), the iSpot community, and other stakeholders, to enhance iSpot and to help further assure its longer-term future.

Acknowledgements

The authors are grateful to all those involved in iSpot, including Jonathan Silvertown, the originator and overall project leader, the rest of the iSpot team, the many groups and organisations supporting iSpot, and all participants in iSpot. The analysis of learning in iSpot and Figure 2 summarizing data were provided by Jonathan Silvertown and Martin Harvey. iSpot was funded by the UK National Lottery through the Big Lottery Fund for England between 2007 and 2012, as part of the Open Air Laboratories project (www.opalexplorenature.com) and the British Ecological Society and the Wolfson and Garson Weston foundations. The iSpot Local project was funded by JISC through the eContent Programme.

References


Implementing Mobile Learning Curricula in Schools: A Programme of Research from Innovation to Scaling

Chee-Kit Looi* and Lung-Hsiang Wong
National Institute of Education, Nanyang Technological University, Singapore // cheekit.looi@nie.edu.sg // lunghsiang.wong@nie.edu.sg

*Corresponding author

ABSTRACT

Many countries, regions and education districts in the world have experimented with models of one-device-per-student as an enabler of new or effective pedagogies supported by mobile technologies. Researchers have also designed innovations or interventions for possible adoption by schools or for informal learning. Of critical interest to the community is the question of how the more successful of these top-down or bottom-up models or innovations can proliferate to more usage, adoption and adaptation across levels of the education system. This paper describes a research programme that demonstrates how to make successful research innovations count in practice and that delineates what types of educational R&D involving scaling need to take place to make the critical link to impacting practice. We do this in the context of one such curricular innovation in a Singapore school that moves through the various phases to where the innovation is becoming an integral part of routine classroom practices.

Keywords
Mobilized curriculum, Translation and scaling up, Seamless learning model, Design-based research, Science education

Introduction

The literature on educational technology research is packed with examples of pilot studies and proofs-of-concepts. It is rarer, in fact, in the literature, to see a project move through the various phases to where the innovation actually has become an integral part of routine classroom practices. In our collaboration work with a primary school in Singapore, we have developed a viable innovation model (the Seamless Learning Model or SLM in short) by working with a class of primary school students over a period of two school years. The innovation involves the transformation of the existing science curriculum into an inquiry-based one which leverages the affordances of mobile technologies. Because SLM demonstrated increased student achievement, the school has decided to scale-up the roll-out of the transformed curriculum to more classes and more subjects in the coming years, thus providing the opportunity to study an innovation as it scales up.

In this article, we trace the journey of this research programme that started with the co-design of a 1:1 (one-mobile-device-per-student) mobilized curricular innovation for a primary school in Singapore, leading to the establishment of efficacy findings by researchers, the decision to scale-up by the school, and the plans for scaling-up. In doing so, we elucidate a multi-term research agenda on an 1:1 mobilized curriculum that studies scaling beyond the initial proof-of-concept to broad and deep usage in the context of a research-based innovation in a school in Singapore. Such an agenda articulates the research questions and posits the scaling research framework and approaches involved. We hope this can provide an existential example of how researchers can do research that addresses the multi-term, multi-pronged, multi-level and systemic aspects of school-based innovations, and that ultimately benefits schools and yet at the same time, derive and refine scientifically and empirically theoretical frameworks, design principles, resources and strategies for learning.

Our research approach: Design-based research and our roles as meso-level mediators

With the goal of working towards scalable and sustainable classroom practices, we took a design-based research (DBR) approach to address complex problems in real classroom contexts in collaboration with practitioners, and to integrate design principles with technological affordances to create solutions to real needs of teaching and learning. The goal of design research is to conduct rigorous and reflective inquiry to test and refine innovative learning environments as well as to refine new learning-design principles (Brown, 1992; Collins, 1992). DBR is iterative as researchers strive to engage in design, work with teachers to enact the design in classroom settings, do research on
the contextualized learning processes, develop or refine theories of learning, engage in iterative re-design, and thereby continue the cycle of design and implementation. It is characterized as being interventionist, iterative, process-oriented, utility-oriented and theory-oriented (van den Akker, Gravemeijer, McKenney, & Nieveen, 2006). It is distributed across teachers and is ongoing, as opposed to a completed trajectory that we as researchers can foresee and oversee. DBR can result in greater understanding of a learning ecology by designing its elements and by anticipating how these elements function together to support learning (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003).

Cognizant of the multiple level constraints that act on teachers adopting new curricular innovations in the classroom, we recognize the complex interplay of multiple dimensions of education reforms. Thus, we approach our programme of research from a systemic change perspective that recognizes the micro, meso, and macro levels of educational systems (Looi, 2011; Looi, So, Toh, & Chen, 2011). The policy imperatives governing Singapore’s educational landscape constitute the macro-level factors, and the contextualized classroom-based work and interactions as micro-level factors. By meso levels, we adopted the view of Jones, Dirchinck-Holmfeld, and Lindstrom (2006) where they define: “meso is an element of a relational perspective in which the levels are not abstract universal properties but descriptive of the relationships between separable elements of a social setting” (p. 37). Meso-level agencies can be perceived as the “recontextualizers” or “constructors of pedagogic discourse who de-locate and re-locate discourse, moving it from its original site to a pedagogic site” (Jephcote & Davies, 2004, p. 549).

Macro-level actors: Policymakers or other actors who set the climate or policies for educational reforms in schools and in learning

Macro-level environment: As seen in national plans or Masterplans where a conducive macro-environment for innovative practices is enabled by governance practices through:
- Setting up the infrastructures
- Creating readiness
- Phasing changes
- Institutionalizing and undergoing creative renewals
- Providing resources

Macro-level emphases: Setting broad educational outcomes
- Scanning trends and directions, and reviewing research at meso-levels and micro-levels that inform pedagogical and technological practices

Meso-level actors: Researchers as re-contextualizers who moved discourse from original to pedagogic site

Meso-level environment: The socio-cultural factors that make up the school’s learning ecology such as the classroom setting situating between individual activities, small groups and larger communities.

Meso-level emphases: Interpreting and operationalizing macro-level emphasis by:
- Effecting the desired epistemological and socio-cultural changes via design research
- Mapping to effective classroom orchestration and implementation that seeks to achieve the desired micro-level interactions and outcomes, via design research
- Considering systemic forces and mediating inter-related tensions to lead to sustainability and scalability

Micro-level actors: Individuals such as students and teachers

Micro-level environment: Interactions or discourse within small group and classroom settings

Micro-level emphases: Informing macro and meso-level emphases by:
- Studying contextualized group or classroom-based interactions in an in-depth manner
- Eliciting feedback from participants

Figure 1. A systemic framework for enabling innovative practices via the alignment of macro, meso and micro levels (adapted from Looi et al., 2011, p. 11)

The socio-cultural factors of the school’s learning ecology constitute the meso-level environment. As researchers from the university, we serve as meso-level actors who work in that environment to recontextualize pedagogic discourse. This re-contextualization process is a “meso-level” mechanism. The orchestration of efforts from all...
actors will contribute explanatory power to the sustainability of an intervention. By approaching this pedagogy-driven reform at the macro, meso and micro levels, we seek the alignment of systemic forces at work to provide a buttress for sustainability. Thus we, as researchers as the meso-level actors, help the school practitioners understand and interpret policy imperatives and actualize them into classroom teaching and learning practices in ways that are informed by research and theories. Figure 1 shows the roles of actors in these three levels in enabling and sustaining innovation.

**The curricular innovation informed by the seamless learning notion**

Through a DBR approach and a lens of researchers serving as meso-level mediators to work with the teachers, we first conducted several research cycles in the three-year research project entitled “Leveraging Mobile Technology for Sustainable Seamless Learning in Singapore School.” The work helped us to develop a viable innovation model (the Seamless Learning Model or SLM in short) by working with a class of primary school students over a period of two school years. The innovation involved the transformation of the existing science curriculum into an inquiry-based one which leverages the affordances of mobile technologies. The project was framed in the broader context of constructing “seamless learning” environments to bridge different learning contexts (such as between formal and informal learning settings, individual and social settings, and learning in physical and digital realms), mediated by mobile devices in 1:1, 24x7 basis (Chan et al., 2006; Wong & Looi, 2011). The basic rationale is that it is not feasible to equip students with all the skills and knowledge they need for lifelong learning solely through formal learning (or any other single learning space); henceforth, student learning should move beyond the acquisition of content knowledge to develop the capacity to learn seamlessly (Chen, Seow, So, Toh, & Looi, 2010). Nevertheless, as such a learning approach is a tall order for students who are more accustomed to the present instructivist-dominated education system. In this regard, we envisaged the design and enactment of long-term seamless learning curriculum where teachers engaged learners in an ongoing enculturation process (Wong, 2013b) in nurturing their habit-of-mind in seamless learning.

We first describe the underlying intent and the guiding principles of this curricular innovation enabled by mobile technologies. These were what shaped the curricular innovation, and they had been iterated and improved progressively through a process of DBR by working with teachers over three years. The process was marked by insitu iterative and collaborative cycles of co-design by researchers and the teacher, leading to enactment by the teacher in one experimental class with data collection. The researchers observed the classes, and provided feedback to the teacher to improve the design and subsequent enactments. The class was a mixed-ability class.

The design of the learning units in the mobilized curricula for science is designed based on these design principles (Zhang et al., 2010):

- Design student-centered learning activities (to promote engagement and self-directed learning)
- Make students’ thinking process visualizable (so that they can be shared and subject to further refinement)
- Incorporate different learning modalities (to personalize learning)
- Design for holistic and authentic learning (make science learning meaningful)
- Facilitate social knowledge building (to promote collaborative learning)
- Ensure that the teacher plays the role of facilitator (to move away from didactic teaching)
- Provide an environment to integrate all learning activities (students have a hub to launch or continue their learning activities)
- Assess formatively (through the learning activities, students can receive feedback for their own ideas from peers or the teacher)
- Extending classroom learning activities beyond school hours and premises (to support the notion of seamless learning).

In the co-design process, researchers worked with teachers to revise and mobilize two years’ worth of the national curriculum for Primary 3 and 4 Science by considering the opportunities afforded by ubiquitous access to mobile devices. Activities were designed which seeks to extend learning activities beyond the classroom. To support the continuous and long-term learning activities, 34 students from the experimental class were each assigned a smartphone with 24x7 access in order to mediate a variety of learning activities such as in-class small-group activities, field trips, data collection and geo-tagging in the neighbourhood, home-based experiments involving parents, online information search and peer discussions, and digital student artifact creation, among others.
The key epistemological design commitments of the curricular innovation are: learning as drawing connections between ideas, and learning as connecting science to everyday lives, across multiple learning spaces. The curricular commitment is seamless learning, and inquiry-based facilitation and learning. The technological commitments include: technology for construction, technology for communication, and technology for searching information anywhere anytime.

Concerning the curricular commitment, in science teaching, the Ministry of Education of Singapore has advocated the use of the BSCS 5E Instructional Model (Bybee, 2002). This 5Es model consists of the following phases: engagement, exploration, explanation, elaboration, and evaluation. Each phase has a specific function and contributes to the teacher’s coherent instruction and to the learners’ formulation of a better understanding of scientific knowledge, attitudes, and skills. The model is used to sequence the learning activities in a science lesson or over a series of science lessons.

The designed curriculum was developed with the use of software apps on the GoKnow™ MLE (Mobile Learning Environment) that runs on a Microsoft Windows Mobile operating system. The GoKnow MLE enables teachers to create differentiated lessons easily via its online learning management system, GoManage, and it enables students to easily personalize their learning experiences (Looi et al., 2009). MLE supports teachers in creating complete, coordinated, curriculum-based lessons that employ multiple media and applications (e.g., text, graphical, spreadsheet, animations, and the like). It is an environment in which students engage in the specified learning activities and create various artefacts. It includes software tools such as:

- KWL (what do I already Know? what do I Want to know? What have I Learned?) to allow students to learn in a self-regulated way,
- Stop Watch that supports timing of events,
- Sketchy™ as an animation/drawing tool, and
- Picomap™ that allows students to create, share, and explore concept maps.

Typically, the lesson is designed to provide opportunities for the student in the Explore, Explain and Elaborate phases of 5E to use the software tools to do their science inquiry. In the Engage phase, the teacher motivates the inquiry by doing some classroom science demonstration or posing some science questions. In the Evaluate phase, the teacher or student peers review the work done on the software tools to detect and correct students’ developing conceptions of the science concepts.

We designed a total of twelve MLE units in the two years of intervention. Note that we did not design the whole curriculum in one go before the intervention commenced. During and after each design-enactment cycle for a MLE unit, we reflected upon the lessons and apply such understanding to inform the design of the next MLE unit. In addition to offering a logical flow for learning the domain knowledge, we had progressively incorporated various types of inquiry/seamless learning activities, from simpler to more demanding ones. This was to facilitate the students’ gradual changes in their habits of mind moving towards learning seamlessly. For example, while the earlier activities involve the students expressing their understanding using the software tools or capturing artifacts outside of the classroom and relating them to the curriculum concepts, the latter activities involve some form of parental involvement in the students’ learning.

Through a process of considering the space of activities enabled by the affordances of the software tools and the smartphones, and the kinds of activities that would be beneficial for primary school science students, we develop a categorization of the 10 major types of smartphone-mediated activities as shown in Table 1. Table 2 summarizes the essential information, including what smartphone-mediated activities were incorporated, of the twelve MLE units.

<table>
<thead>
<tr>
<th>Table 1. Types of mobile-assisted activities incorporated in the MLE curriculum (Wong, 2013b)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activity ID</strong></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>KWL</td>
</tr>
<tr>
<td>Anim</td>
</tr>
<tr>
<td>Ph</td>
</tr>
<tr>
<td>CM</td>
</tr>
<tr>
<td>Dsc</td>
</tr>
<tr>
<td>Trp</td>
</tr>
<tr>
<td>Exp</td>
</tr>
</tbody>
</table>
Activities with parental involvement

<table>
<thead>
<tr>
<th>Par</th>
<th>Videos &amp; other tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web</td>
<td>Web search and media playing</td>
</tr>
<tr>
<td>Col</td>
<td>In-situ multimedia content creation &amp; forum discussion</td>
</tr>
</tbody>
</table>

| In-situ multimedia content creation & forum discussion |
| ColInq (with geo-tagged postings, each served as a discussion thread) |

**Table 2.** List of MLE units designed for the mobilized curriculum (Wong, 2013b)

<table>
<thead>
<tr>
<th>Level &amp; time period</th>
<th>Topic</th>
<th>Anim</th>
<th>KWL</th>
<th>Ph</th>
<th>CM</th>
<th>Dsc</th>
<th>Trp</th>
<th>Exp</th>
<th>Par</th>
<th>Web</th>
<th>Col</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb 2009</td>
<td>Classification for living &amp; non-living things</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb &amp; Mar 2009</td>
<td>Classification of animals</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar &amp; Apr 2009</td>
<td>Plant</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar &amp; Apr 2009</td>
<td>Plants &amp; their parts</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar &amp; Apr 2009</td>
<td>Fungi</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr &amp; May 2009</td>
<td>Materials</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug &amp; Sep 2009</td>
<td>Body systems</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan &amp; Feb 2010</td>
<td>Cycles</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb &amp; Mar 2010</td>
<td>Matter</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 2010</td>
<td>Light &amp; shadow</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr &amp; May 2010</td>
<td>Heat &amp; temperature</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul 2010</td>
<td>Magnet</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. The first seven units constituted the Primary 3 mobilized curriculum; the last five belonged to the Primary 4 curriculum*

**Demonstration of efficacy of innovation**

The curricular innovation involves designing a coherent and sustainable classroom program. The unique research context is that we worked with a school with 9 classes in the cohort of Primary Grade 3 taking the science subject. The experimental class followed the same class schedule and assessment schemes as the rest of the classes. What evidence can we demonstrate in establishing the efficacy of the innovation? Our analysis of the science examination scores after the one year intervention (with the experimental class using the mobilized curriculum to replace the traditional curriculum) shows that amongst the 6 mixed-ability classes in Primary (Grade) 3 in the school, the experimental class performed better than other classes as measured by traditional assessments in the science subject (see, Looi, Zhang, et al., 2011). With mobilized lessons, students were found to learn science in personal, deep, and engaging ways as well as developed positive attitudes towards mobile learning. We feel that this result is a very worthwhile contribution to the field, as much research work on mobile learning focus only on units of at most a few weeks duration or they are add-on activities to some existing curriculum.

The use of the mobilized technologies provides many leverage points for the researchers and teachers to co-design a new curriculum that focuses on inquiry learning. Once designed, the curriculum can be enacted by science teachers, and it is important for the teachers to understand the design principles behind a mobilized curriculum for inquiry learning and how to implement in the way to harness the best learning outcomes for students. We see a shift in the teacher’s attitudes and behaviors towards science teaching, from a style that sees her pre-occupied with just covering the curriculum to one that allows her to watch over and facilitate students’ work on the inquiry activities on their handhelds.
With the mobilized lessons, we observe students engaging in science learning in personal and engaged ways. They demonstrated their understanding of science phenomenon in multimodal ways and did self-directed learning by doing online search and exploration on questions related to the curriculum topics. They engaged in instructional activities that involve their parents, as in our mobilized lesson for the body systems. This lies in contrast to the more “traditional” way of learning, in which students learned science from the didactic instruction of the teacher or from the textbook.

Providing the basis: Explicating and theorizing the SLM model

The pilot has shown that Seamless Learning (SL) is a viable pedagogical model. SLM harnesses the portability and versatility of mobile devices to promote a pedagogical shift from didactic teacher-centered to participatory student-centered learning (Facer et al., 2004). With appropriate learning design, the mobile technology facilitated the transformation of classroom learning activities into a more student-centered, personalized, and social learning process for the students where they need to process and associate their experiences or the information received in the informal contexts with the knowledge that they have acquired or constructed in the classroom, reflect upon any discrepancy, and apply the knowledge to solve real-life.

As design-based researchers, we also seek to work towards theories for how to get students to learn as self-directed seamless learners (Looi et al., 2010; So, Kim, & Looi, 2008; Wong, 2013b). We explore how the theories and methodology of self-regulated learning (SRL), an active area in contemporary educational psychology, are inherently suited to address the issues originating from the defining characteristics of mobile learning: enabling student-centered, personal, and ubiquitous learning (Sha, Looi, Chen, & Zhang, 2012). These characteristics provide some of the conditions for learners to learn anywhere and anytime, and to be motivated and able to do so in a strategic way, namely, self-regulate their own learning. We propose an analytic SRL model of mobile learning as a conceptual framework for designing and analyzing mobile learning, in which the notion of self-regulation as agency is at the core (Figure 2). The rationale behind this model is built on our recognition of the gaps in the current conceptualization of the mechanism and processes of mobile learning, and the inherent relationship between mobile learning and SRL.

At the center of the model is the notion of self-regulation as agency, referring to the learner characteristics that function as internal driving forces initiating and sustaining a self-regulated mobile learning process. The key personal factors include domain knowledge, prior experiences, motivation, and metacognitive awareness, epistemological beliefs, and so on. Self-regulated mobile learning processes can be understood and analyzed by means of SRL theories and methodologies (e.g., self-report survey, trace analysis). Mobile learning processes that are regarded as manifestation/exercises of agency (fundamentally composed of motivation and metacognition) can be understood, analyzed, and assessed from the theories and methodologies of SRL. Second, mobile learning activities are supposed
to be mediated by mobile technologies and devices, which presumably function as social, cognitive, and metacognitive tools. The existing studies in mobile learning largely focus on the social and cognitive functions but ignore the metacognitive function.

Methodological contributions: Studying students on the move

SLM advocates continuous learning by the students inside and outside of the classroom. The curriculum incorporates components where they use their personal mobile devices to do planned activities at home or outside of the classroom, and where they also learn in emergent and unintended ways. The challenge for us researchers is to study “seamless learning” as synergistic and continuous learning across multiple spaces and time scales which involves theoretical and methodological challenges (Toh, So, Seow, Chen, & Looi, 2013). With students constantly on the move and many of their interactions happening in informal settings, researchers face significant challenges for investigating and documenting emergent forms of learning and participation across multiple contexts. The challenges of collecting data for students on-the-move include making sense of the voluminous data collected, and navigating the complexity of ethical issues involved in the intrusive and non-intrusive data collection methods.

To fill the gap in mobile learning research, we innovated on a methodological approach for researching digital kids’ learning supported by mobile technologies. The specific focus on methodological issues is significant with recent trends in mobile learning research’s focus on learning beyond school settings. Building on earlier work on methodological issues (e.g., Hsi, 2007; Martin, 2004; Vavoula & Sharples, 2009), we adopt and adapt the method of cooperative inquiry to study the sense-making endeavors of digital students who are always on-the-move and learning across multiple settings. We do so by working with parents who would cooperate to collect data such as by using video-recording and taking photos of their children doing things on their mobile devices.

Moving beyond the Proof-of-concept: From innovation to scaling

Is the curricular innovation ready for scaling and how do we know to what extent scaling has taken place? One of the most cited literature on scaling is that of Coburn (2003) who defined scale as encompassing four interrelated dimensions:

- **Depth**: Depth refers to deep and consequential change in classroom practice, altering teachers’ beliefs, norms of social interaction, and pedagogical principles as enacted in the curriculum.
- **Sustainability**: Sustainability involves maintaining these consequential changes over substantial periods of time.
- **Spread**: Spread is based on the diffusion of the innovation to large numbers of classrooms and schools which will adopt and adapt the innovations.
- **Shift in reform ownership**: Shift requires districts, schools, and teachers to assume ownership of the innovation, deepening, sustaining, and spreading its impact.

Building on this work, Clarke and Dede (2009) added a fifth dimension, namely, evolution, in which the innovation, as revised by its adapters, is influential in reshaping the thinking of its designers and creating a community of practice that evolves the innovation.

Our 3-year design research study in the primary school has helped achieve some successes in dimensions of scale by Coburn (2003) in four ways:

- **Depth**: The intervention has created positive learning gains for the students of the 2 classes and positive changes in attitudes and knowledge of the 2 teachers (Looi, Zhang et al., 2011).
- **Sustainability**: Clearly changes have occurred in the school with evidences from research analysis (Looi, Zhang, et al., 2011; Zhang et al., 2010) during the two years of intervention, and from interviews with the stakeholders (school leaders and teachers).
- **Spread**: In 2009, we worked with 1 teacher and 1 P3 class. In 2010, we worked with 2 teachers and 2 P4 classes. The school is spreading the mobilized curriculum for science to all P3 classes in 2012, and to all P4 classes in 2012.
- **Shift in reform ownership**: The school has taken over ownership by driving the spread of the mobilized curriculum for science to all P3 classes in 2012.
When the curricular innovation using mobile devices has been developed and studied in the context of one class, and the empirical evaluation of the mobilized curriculum has shown its potential for learning effectiveness, the school leaders decided that it was a worthwhile innovation and, in consultation with the researchers, would like to scale up the innovation.

We envisage the following four levels of scaling in the life-cycle of educational research and development work that lead to practical and policy implications for how to scale up in a school context:

**Level 1:** Developing the intervention as an innovation through a pilot in one or two classrooms

**Level 2:** Grade-level scale—spreading to more classes in a grade level and eventually to a whole grade level

**Level 3:** School-level scale—spreading the innovation to a whole school

**Level 4:** District or country-level scale—in which such work will have policy implications for the educational authorities.

We note that many research interventions do not survive the first level, or they merely go through many research funding cycles to iterate and re-design the intervention, and stop at that. Our interest is those research innovations which have been shown to have demonstrated learning efficacies, and there is purpose and commitment from all the relevant stakeholders to scale-up the intervention. We know it is rare for bottom-up research interventions to move from Level 1 to Level 2, let alone to the other levels. Thus it is critically important to study and learn from the few instances out there that actually move up the levels, and this is towards exemplifying and informing how research can really bridge the research-practice gap by through such scaling studies.

To make SLM an integral part of classroom practices, what would then be useful for the school and for the Ministry of Education is to “ruggedize” the innovation for sustainability to retain substantial efficacy in diverse or even relatively barren contexts. To ruggedize the innovation, robust-design strategies are needed that will enable the innovation to be used in multiple settings (Clarke & Dede, 2009). The school has moved onto the Level 2 and has plans to share the model with other schools.

So far, we have created a successful story of research-informed technology-enabled practices in the school, but we want to study further how to sustain such successful practices over time after an initial influx of resources and other forms of external support. For sustained changes, we would reemphasize the importance of meso-level mechanisms that support teacher capacity building and reinforce school leadership and culture. Barab & Luehmann (2003) discuss issues of sustainability and local adaptation as crucial for scale. They describe the teacher’s role in local adaptation as identifying local needs, critiquing the innovation in the light of those needs, visualizing possible scenarios of implementation, and finally making plans or decisions concerning the implementation. Teachers will be ultimately involved in the adoption, customization, and implementation process, and they are continually remaking and contextualizing the innovation in terms of their local context. Barab and Luehmann (2003) argues that instead of the equation

\[
\text{Designed Curriculum} = \text{Implemented Experience}
\]

it should be

\[
\text{Teachers Perceptions} + \text{Designed Curriculum} + \text{Classroom Culture} = \text{Implemented Experience}
\]

For example, stark differences would exist between the customization and implementation of an innovation by a subject-matter focused teacher who is very concerned about students’ understanding of the content compared to a logistically focused, teacher who is most concerned that the activities run smoothly and in the appropriated time frame. Our current phase of research studies how different teachers would locally adapt the innovation and what the resulting implemented experiences for the students and for the teacher would be.

The critical research questions are:

1. How to adapt and to “ruggedize” the innovation for sustainability to retain substantial efficacy in diverse contexts of more classes in the grade level and more teachers, in which some of the conditions for success are absent or attenuated? This includes designing a more device-independent curriculum intervention for mobile learning. By clarifying the design principles and the design affordances, it is hoped that the new curriculum development model is more generic and less dependent on the devices the students use.

2. What is the impact on depth, sustainability, spread, shift of ownership and evolution when a school takes over ownership and scales up an innovation developed from one class to more classes and eventually to a whole grade level?
3. What are the strategies for curriculum development and professional development models needed to support the spread of a mobilized curriculum to one whole level?

4. What are the strategies for organizational, technological and institutional changes needed for such a sustainable and scalable translation of research into practice?

A key issue in scaling is how to explicate the curriculum for the typical teacher. While during a pilot, one can leave both the content and the instructional strategies “looser” and “less-defined,” that strategy will not work when all teachers read, use and interpret the same curriculum. The curriculum needs to be specified (Cohen & Ball, 1999). The goal is to make the content and the instructional strategies explicit – make the curriculum transparent for the teacher to enact initially and to adapt subsequently.

Another key issue relates to ongoing professional development. Teachers need to be provided with continuous support as they transition to a new curriculum, providing the performance support (Cohen & Ball, 1999). In contrast to a pilot, where the teachers are typically highly motivated and are top-notch teachers, as an innovation goes to scale, all teachers must be brought up to speed. Some of those teachers will be motivated and some less so; some teachers will be highly competent and some less so. Thus, professional development that is ongoing, continuous must be put in place to help all the teachers, especially the weaker ones, understand how to rollout the innovation.

The teachers need more collaborative work sessions so they can help each other with suggestions on instructional strategies and with tweaks to the curriculum. Those additional collaborative work sessions are critically important. If teachers feel isolated, they will not enact the innovation. They need to form their own learning community for mutual support and rely less on the research team. Moreover, the fundamental issue is to shift the teachers’ epistemological and educational beliefs from being a transmissionist and behaviorist to being a socio-constructivist – without the shift, all the curricular mobilization efforts will become a mere formality.

Aligned with the issue of teachers adopting and adapting a new pedagogy, there is a need to make the formative evaluation techniques explicit and to show teachers how to use the formative assessments in order to tailor their instructional practices to the specific needs of the individual students. In order to go to scale, all teachers need to use the same formative assessment techniques and thus these techniques must be made explicit and teachers must be given support as they learn how to administer and use those formative assessments.

The curricular innovation is enabled by mobile technologies. A more sustainable approach is to adapt the curriculum for a more generic mobile technology. This makes the curriculum and the formative assessments work with a broader range of mobile technologies and applications. While initially the innovation used smartphones, the goal is to create materials – curriculum, instructional strategies, and formative assessments that are mobile technology agnostic. Mobile technologies are changing very quickly; thus, we do not want our learning resources to be tied to a specific mobile technology. Moving towards the blending of mobile and cloud computing technologies (Wong, 2012) is a direction in our agenda for scaling up. The new direction will not only offer a feasible solution to the above-stated challenge of changing technologies, but also has the potential to open up new opportunities for developing more advanced affordances to mediate a wider range of seamless learning activities.

What have we learned about scaling so far?

At the school, we have moved from serving 80 students (2 teachers) to serving 320 students (6 teachers). That almost order-of-magnitude jump in who is supported requires R&D if that transition is to be effective. What we learned in the pilot was critically important, e.g., that SLM can be an effective pedagogical model in supporting students as they engage in inquiry-based learning. However, in going to scale, the issue is no longer one of efficacy but one of infrastructure – how do we take an innovation that was hand-crafted and make it more rugged, more robust, more stand-alone, more transparent.

By the end of academic year 2012, all teachers of P3 have enacted the curriculum. What have we scaled up to enable this spread to all teachers and all classes in P3? The scale-up comprises these multiple dimensions:

1. Mobilized curricula (to lead students to self-directed learning and to bridging informal learning spaces)
2. Teacher facilitation skills
3. Teacher readiness
4. Student readiness including hardware and software training
5. Technology infrastructure, e.g. WiFi and 3G Connectivity; availability of mobile devices in 1:1, 24x7 basis

On assimilating the curriculum into the classroom culture (Squire, MaKinster, Barnett, Luehmann, & Barab, 2003), we summarize the challenges the teachers faced in doing a new curricular innovation that builds new skills and competencies beyond what is usually assessed in the standardized assessments used in the school. In the researchers’ interviews with the teachers, the teachers raised these concerns:

- They are not sure if they are conducting the designed curriculum lessons in the right way. Some teachers expressed doubts on the students’ ability for doing self-directed learning.
- Some teachers expressed that they needed help in developing and practicing questioning skills in the classroom.

What support was provided to the teachers to help them to address these challenges? The response to the first concern was for the researchers to provide support to the teachers by giving personalized feedback after each lesson enactment, and by helping the teachers to adapt the lessons for different ability students. The response to the second concern was for teacher sharing facilitated by the researchers in the weekly curricular design meetings to discuss questioning skills via modelling by a teacher or researcher, and by lesson study discussions on a recorded classroom sessions. While researchers provided these initial scaffolding, the plan was to build up the capacity of this group of primary 3 science teachers so that they can sustain the innovation in the coming years with fading from the researchers in the subsequent year.

In levelling up the capacity of the teachers, some new activities were planned. Arrangements were made for teachers to observe their peers conducting the lesson activities. There were some discussions for each teacher to co-teach with another teacher, but this was later not followed up on because of time constraints and time tabling issues (for example, when the teachers are teaching concurrently thus making it difficult to visit other classes for co-teaching). The researchers decided to edit short clips of videos which they found out to exhibit good teaching and learning scenarios to be shared with other teachers.

Scaling to Chinese and English language learning

In previous sections, we reported the outputs of our past research – the explication of the SLM, the curriculum development principles, the research methodology to study learners on the move, the PD model, our experience and understanding in fostering the technological and systemic conditions for the establishment of a sustainable 1:1 seamless learning environment, etc. Apart from the scaling up of the science curriculum, these research findings and deliverables are also serving as the basis for scaling the SLM to other school subjects. To date, two language subjects, namely, Chinese and English, have embarked on their respective journeys of transforming their curriculum into seamless learning environments, again with the eventual aims of cross-school scaling up.

“MyCLOUD” (My Chinese Language ubiquitOUs learning Days, 语飞行云) (Wong, Chai, Chin, Hsieh, & Liu, 2012) is a levelling up effort of the completed study of “Move, Idioms!” (Wong, 2013a; Wong, Chin, Tan, & Liu, 2010). The DBR study aims to develop a holistic and scalable mobile- and cloud-assisted Chinese Language (CL) seamless learning environment that is informed by language learning/acquisition theories for P3-P5 students. The project involves a 2 ½-year school-based intervention (August 2011-November 2013). The intention is to enculturate students in carrying out CL vocabulary learning and communicative writing activities that encompass formal and informal learning settings. Under this approach, students are assigned mobile devices on 1:1, 24x7 basis in order to stimulate and support their language learning both within and beyond the classrooms. In addition, a cloud-based, device-independent MyCLOUD learning platform leveraging mobile and cloud computing technologies has been developed for the purpose. Another aim is to design new classroom practices that will be integrated into the existing formal curriculum as well as foster students’ competency to engender deep CL learning. The learning design principles developed by the researchers specifically foreground the bridging of language input (classroom, textbook-induced learning, and sharing of students’ linguistic artifacts) and output (students’ daily use of the target language through creations of linguistic artifacts such as photo taking / sentence making and social networking), and the bridging of learning contextualization and learning generalizations/reflections. On top of the more generic 10 features/dimensions of seamless learning as proposed by Wong and Looi (2011), the additional language learning-specific SLM dimensions rooted in various language learning theories may become the basis of the our future exposition of a Seamless Language Learning model. In addition, the research team is working towards spreading the curricular innovation to four more schools by 2014.
Aligned with the goal of implementing seamless learning and cultivating self-directed learners, smartphones were adopted in a trial design and implementation of the Primary 3 English language curriculum at Nan Chiau Primary School. This “WE Learn Project” is a scaling up of the seamless learning initiative in Primary 3 Science. By transforming the classroom from the traditional teacher-centered model to a learner-centered one, the project hopes to enhance the learning outcomes of students. In “mobilizing” the English curriculum and making it inquiry oriented, the teachers and curriculum developers drew on the two pedagogical strategies of P4C and Marzano’s.

P4C (Philosophy for Children), (Lipman, 1980) draws on the Socratic method of learning pioneered initially in Plato’s dialogues and focuses on learning how to ask a question and how to respond when asked a question. Marzano’s 6-steps to Better Vocabulary Instruction, (Marzano & Pickering, 2005) help children understand words by building relationships and links amongst the words, by using words in their proper contexts.

In science, in CL and the English language learning, while the particular styles of pedagogy have their differences, the pedagogies are, at their roots, inquiry-oriented, and learn-by-doing pedagogies that reinforce each other (Norris et al, submitted). In all these three subjects, the affordances of the mobile technology enable the students to use multiple modalities and multiple media, and to carry and use the smartphones anywhere and anytime as a learning hub.

At this point of writing (in November 2012), there are now approximately 350 P3 (3rd grade) students using mobile computing devices daily for science. Out of these 350, 120 students are also using the devices for English and Chinese language (MyCLOUD) learning. The school has plans to scale to approximately 700 students in P3 and P4 in science, to 350 students on P3 English, and 350 students in Chinese language learning.

**Conclusion**

Within the educational research community, many research projects focused on designing or establishing the efficacy of innovations that work well within specific contexts. They typically face the conundrum of narrowing the research-practice gap when it comes to changing or transforming practices in schools and other contexts for learning, and to scaling up to meet the needs of a broader audience. Such research projects are also not organized to address the challenge of long-term systemic improvement as research-practice partnerships often go in tandem with short-term funding of grants and program initiatives at foundations and government agencies. This paper describes a multi-year research programme for doing scaling and implementation research. The agenda is contextualized in the example of a mobile learning curricular innovation in a Singapore school as it goes to scale.

In this curricular innovation on mobile learning, we established that SLM can raise student achievement in the context of one class and one teacher. We have developed models for:
1. Transforming a curriculum to harness the affordances of mobile technologies
2. Technology infrastructure and support for a mobilized curriculum
3. Teachers’ professional development through working with 1 teacher for a grade level 3 (P3) class in 2009 and 2 teachers for two grade level 4 (P4) classes in 2010.

Because SLM demonstrated increased student achievement, the school decided to scale-up the roll-out of the transformed curriculum to all the eight P3 classes, by doing the planning in 2011 and doing the scaling in 2012. The next step in the research and implementation trajectory brings us to scaling research which seeks to make research count in practice. The goal is to study the adoption and the adaptation of the curricular innovation as it goes to scale to more classes, more levels and more subjects, and to documenting the benefits of balancing fidelity of implementation with adaptation to dynamic local contexts. The programme of research enables us to articulate the SLM curriculum, the supporting resources, the teacher learning and professional development models with analysis of their impact, efficacies as well as weaknesses. The scaling research will establish a model of scaling that recognizes the range and diversity of teachers’ local needs as well as the necessary adjustments they need to make in order for the innovation to be useable and effective, and how the school can support them to know how to adapt the innovation and yet retain the essence of its efficacy.

We hope that our narration of the ongoing research journey from innovation to practice and to scale can inspire other research initiatives that will address the multi-term, multi-pronged, multi-level and systemic aspects of school-based
innovations, and that yet at the same time, advance theory, frameworks, design principles, resources and strategies for effective and sustainable mobile learning.

Acknowledgements

The entire seamless learning program reported in this paper consists of several projects funded by Interactive Digital Media (IDM) Program (project number: NRF2007-IDM005-MOE-008 LCK); Office of Education Research, National Institute of Education, Nanyang Technological University (project numbers: OER 26/12 LCK, OER 61/12 WLH, OER 16/10 LCK, OER 17/10 LHW); and EduLab Program (project number: NRF2011-EDU002-EL005 (C12)) respectively.

References


Ubiquitous Learning Project Using Life-logging Technology in Japan

Hiroaki Ogata*, Bin Hou2, Mengmeng Li2, Noriko Uosaki, Kosuke Mouri2 and Songran Liu2

1Faculty of Arts and Sciences, Kyushu University, Japan // 2Faculty of Engineering, University of Tokushima, Japan // 3Osaka University, Japan // hiroaki.ogata@gmail.com

*Corresponding author

ABSTRACT
A Ubiquitous Learning Log (ULL) is defined as a digital record of what a learner has learned in daily life using ubiquitous computing technologies. In this paper, a project which developed a system called SCROLL (System for Capturing and Reusing Of Learning Log) is presented. The aim of developing SCROLL is to help learners record, organize, recall and evaluate ULLs. Using SCROLL, learners can not only receive personalized quizzes and answers to the questions, but also navigate and be aware of their past ULLs supported by augmented reality views. In particular, this paper introduces an approach that helps learners record their learning experiences in daily life from life-log photos with the help of SenseCam. To evaluate the effectiveness of this system, a case study of an undergraduate English course is presented to show how it can be used to facilitate seamless learning.

Keywords
Mobile learning, Ubiquitous learning, Learning log, Life logging

Introduction
CSUL (Computer Supported Ubiquitous Learning) or context-aware ubiquitous learning (u-Learning) is defined as a technology-enhanced learning environment supported by ubiquitous computing technologies such as mobile devices, RFID tags, and wireless sensor networks (Ogata & Yano, 2004; Wu, Hwang, & Chai, 2013). It is characterized by its augmented learning environment which presents information on personal mobile devices through the Internet based on the detection of physical objects in surrounding environment using sensing technologies (Hwang, Tsai, Chu, Kinshuk, & Chen, 2012).

The fundamental issues of CSUL are:
1. How to record and share learning experiences that happen anytime and in any place.
2. How to retrieve and reuse them for future learning.

To tackle these issues, LORAMS (Linking of RFID and Movie System) (Ogata, Matsuka, El-Bishouty, & Yano, 2009) was proposed. There are two types of learners in this system. One type consists of providers who record their experiences on video. The other type are learners who, when encountering some problems in their learning, may find the videos uploaded by the first groups of learners useful. The system automatically links between physical objects and the corresponding objects in a video and allows sharing among users. By scanning RFID tags, LORAMS shows users the video segments that include the scanned objects. Although this system is useful in certain environments, it is not currently easy to apply in practice. Therefore, we have begun a more practical research project called “ubiquitous learning log” (ULL) in Japan in order to store intentionally what learners have learned as ubiquitous learning log objects (ULLOs) and consequently reuse them. The project was conducted from October 2009 to March 2013 with the financial support by PRESTO of the Japan Science and Technology Agency (JST).

We define a ubiquitous learning log (ULL) as a digital record of what a learner has learned in daily life using ubiquitous technologies, and propose a model called LORE to show the learning processes from the perspective of the learner’s activity (Ogata et al., 2012). In this paper, we propose a system called SCROLL (System for Capturing and Reusing Of Learning Log) that helps learners log their learning experiences with photos, audios, videos, locations, QR-codes, RFID tags and sensor data, and share their ULLOs with others. Also, a learner can receive personalized quizzes and answers to their questions. This system is implemented both on the web and on the Android smartphone platform. With the help of built-in GPS and cameras in smartphones, learners can navigate and be aware of past ULLOs via the augmented reality view.

Originally, the term “learning log” was used for personalized learning resources for children. The logs were usually visually written notes of learning journals, which could become an integral part of the teaching and learning program.
and which had a major impact on their drive to develop more independent learners. Research findings indicated that
journals were likely to increase meta-cognition and reflective thinking skills through students becoming more aware
of their own thought processes (Hung et al., 2014; Hwang, Wu, Zhuang, & Huang, 2013; Stockwell, 2007; Susan &
White, 1994; Wood Daudelin, 1996). Our approach focuses on how to enrich learning logs and promote retention
and meta-cognition by using mobile, ubiquitous and context-aware technologies.

Life-logs

Life-logs are a notion that can be traced back at least 60 years (Bush, 1945). The idea is to capture everything that
ever happens to us, to record every event we experience and to save every bit of information we ever touch. For
example, SenseCam (Hodges et al., 2006) is a sensor-augmented wearable still camera which is proposed to capture
a log of the wearer’s day by recording a series of images and capturing a log of sensor data. This is a great tool for
recording life logs, as it is a small digital camera that is combined with a number of sensors to help capture a series
of images of the wearer’s whole daily life at the proper time, and can be worn around the neck (Figure 1). Originally
this device was designed as a memory aid.

MyLifeBits (Gemmell, Bell, & Lueder, 2006) stores scanned material (e.g., articles, books) as well as digital data
(e.g., emails, web pages, phone calls, and digital photos taken by SenseCam). The Ubiquitous Memory system
(Kawamura, Fukuhara, Takeda, Kono, & Kidode, 2007) is a life-log system using a video and RFID tags. Another
application, Evernote (www.evernote.com), is a tool to save ideas using mobile devices such as Android and iPhone.
The common idea of these projects is to use life-log data as a memory aid. SCROLL, however, aims to utilize life-
log data for the learning process.

SCROLL

Design

In this paper, a ubiquitous learning log (ULL) is defined as a record of what a learner has learned in daily life using
ubiquitous technologies. ULL is considered as a set of ULOs (Ubiquitous Learning Log Objects). The learning can
also be considered as the extraction of meaningful knowledge from past ULLs that serves as a guide for future
behavior (Wood Daudelin, 1996). Figure 2 shows the learning processes from the perspective of the learner’s activity
model called LORE (Log-Organize-Recall-Evaluate). These four steps are explained as follows:
1. Log what the learner has learned: When learners face problems in daily life, they may learn some knowledge by
themselves, or ask others for help. The system records what is learned during this process as a ULO. We
designed two modes to record learning contents – active and passive. Here is a typical scenario of the active
mode – when a foreign student in Japan walks into a supermarket, there are many foods that he/she does not
know how to say in Japanese. The student can take a photo of this food by Smartphone and ask someone how to
say it in Japanese, then log the learning content as a ULO including the photo, the name of the food in both
Japanese and the learner’s mother language, location, time, etc. However, in the passive mode, a device such as a
life-log camera can take photos of food and record the contextual information such as location and time
automatically and wait for the learner to review the recorded contents before logging them as ULOs.
2. Organize ULL: When a learner tries to add a ULO, the system compares it with other ULOs, categorizes it and
shows similar ULOs if they exist. There are many ways of categorizing ULOs. For example, a foreign student
in Japan learned a new word “tofu” in a supermarket and logged this process as a ULLO. This ULLO can be categorized as “Japanese,” “Food,” “Supermarket,” etc. As such, it is difficult for the system to categorize it. Therefore, in the designed system, users can add their ULLOs into multiple categories and add several tags to each one. After that, they can review the learned contents by category/tag. Similar objects can be found by matching titles, content of photos, locations and categories, and then the knowledge structure can be regulated and organized.

3. Recall ULL: Learners may forget what they have previously learned. Rehearsal and practice in the same or another context in idle moments can help to recall past ULLOs and to shift them from short-term to long-term memory. Therefore, the system provides some quizzes and reminds the learners of their past ULLOs.

4. Evaluate: It is important to recognize what and how learners have learned by analyzing their past ULLs, so that they can improve what and how to learn in the future. Therefore, the system refines and adapts the organization of the ULLOs based on the learners’ evaluation and reflection.

All of the above learning processes are supported by SCROLL.

We designed SCROLL to implement several types of learning, including self-directed and personalized learning, reflective learning, collaborative learning, situated learning, experiential learning, and seamless learning.

**Self-directed and personalized learning**

The first kind of learning is self-directed and personalized learning. We designed SCROLL based on the following two objectives that enable self-direction and personalization:

- The system can be aware of a learner’s current context. Currently, the context includes location and time. For the location information, the system can detect whether learners are near the place where they uploaded a learning log and whether there are location-based learning logs recorded by other learners nearby. If either requirement is met and if the availability of the device is high, the system will present a quiz based on the knowledge gained in that location or notify the user of the surrounding learning logs added by others.

- The system can record the context data when learners use the system as their context history and then detect their learning habits by analyzing their context history. If the system detects learning habits, and the circumstance meets these habits, it will issue a recommendation message to encourage the user to review what he/she has learned.

**Reflective learning**

An important goal of the SCROLL system is to help learners recall what they have learned after they have archived their learning logs. When a learner captures a learning log, in addition to its location-based property mentioned
above, a number of things are designed as retrieval cues for the learner. For instance, according to the picture superiority effect (Defeyter, Russo, & McPartlin, 2009; Shepard, 1967), learning logs with pictures are much more likely to be remembered than those without. In addition, according to the basic research on human learning and memory, practicing retrieval of information (by testing the information) has powerful effects on learning and long-term retention. Moreover, compared with repeated reading, repeated testing enhances learning even more.

For the above two reasons, taking advantage of photos, locations and so on, the quiz function is proposed. Three types of quizzes can be generated automatically by the system: yes/no, text multiple-choice and image multiple-choice.

Usually, learners can examine themselves by taking quizzes (Hwang, Tsai, & Yang, 2008), but two more ways that are instigated by the system are provided. One is that when a learner moves to the place where he/she captured the knowledge, the system can present quizzes about the learned knowledge. The other is that if learners have learning habits, the system will prompt them to review what they have learned using quizzes. These two methods are discussed in detail later in the paper.

Collaborative learning

We designed SCROLL to also encompass collaborative learning. Since learning logs are logs registered arbitrarily by each learner, collaborative learning in SCROLL adopts an asynchronous model. Any learner in this system is able to share ULLOs, and the system will show their shared ULLOs to others. Besides, they can also ask others questions about their shared ULLOs. In reflective learning, shared ULLOs can also be used to generate quizzes in order to help learners learn more objects.

Situated learning and experiential learning

According to Lave & Wenger (1991), situated learning is learning that takes place in the same context in which it is applied. Itin (1999) defined experiential learning as “learning from experience.” We introduce a concept called "task" in SCROLL to implement situated learning and experiential learning. Learning in the same context enhances the learning effect, and past experiences help learners learn effectively. Tasks refer to the activities through which they can acquire knowledge. Tasks are conducted in the circumstances where learning can happen such as in a school, hospital, post office and so on. For instance, if the system recommends the Japanese word “トマト (tomato)” to a learner in a supermarket, learners can talk with the staff in the supermarket using the word “トマト (tomato),” such as asking its price, location, related recipes and so on. It has been proved that by talking with a native Japanese speaker using the recommended word, learners can master the word well (Jonassen & Grabowski, 1993). The activity of asking about the information is a kind of “task.” Learners who save learning logs are responsible for providing what kind of knowledge can be gained by carrying out the task, and one learning log can be used in several tasks. Moreover, the system provides some predefined tasks in different contexts in order to reduce the learners’ burden when they save their learning logs. In addition, the tasks can be created by the learner and designated by the administrator of the system.

The system assigns an appropriate task for a learner according to the difficulty level of the task and the learner’s ability. For example, asking the price of the product is easy for learners, while asking about vegetable recipes is quite difficult for most learners. When learners receive the recommended learning log and the task, they are also asked to provide feedback for the system. For example, they are asked to take photos of the target object if the learning task is to find where the object is. Moreover, if the learning task is to learn about the place of the object, they need to collect and fill in the environmental information on the system. Only by providing feedback can the users prove that they have really gained the knowledge. Moreover, if the learners meet new problems when carrying out the tasks, they can record them in photos, videos, audios or texts and upload them to the system in order to ask for help. Such accumulated data is also meaningful for other learners.
Seamless learning

Recent progress in mobile and wireless technologies offers us a new learning environment, namely “seamless learning” (Wong & Looi, 2011). It allows learners to learn anytime, anywhere, and provides them with multiple ways of learning throughout the day. By seamless learning, we mean learning which occurs with seamless transitions between in-class and out-of-class learning (Hung et al., 2013). The American College Personnel Association (1994) has indicated the importance of linking students' in-class and out-of-class experiences via providing seamless learning environments to achieve academic success.

Based upon the above ideas, we designed the following Seamless Mobile-Assisted Language Learning Support System (hereafter called the SMALL System) (Uosaki et al., 2012) as a sub-project. The main objective of SMALL is to link learners’ out-of-class learning to their in-class learning. Once a learner uploads a newly learned word to SCROLL, our main system, SMALL, runs a search through the previously updated textbook data. If the new word is found in the textbook data, it jumps to the textbook page where this word is used. Another example is that when a user reads an uploaded textbook and clicks a word, then it jumps to the SCROLL system page to show how other learners have learned this word in different contexts in their out-of-class learning. In this way, users’ out-of-class and in-class learning can be intertwined. We learn words from contexts. In order to master words, it is important to come across them used in various situations.

System interface

We implemented SCROLL both on web and smartphone platforms. It consists of the following components:

ULL recorder

This component facilitates the way learners upload their ULLOs to the server whenever and wherever they learn. As shown in Figure 3 (1), in order to add a ULLO, a learner can take a photo, ask questions about it and attach different kinds of meta-data to it, such as its meanings in different languages.

![Figure 3. SCROLL interface on an Android mobile phone](image)
ULL finder

If a learner registers a new ULLO, the system checks whether the same object has already been stored by comparing the name fields of each object using a thesaurus dictionary. Also, a learner can search ULLOs by name, location, text tag and time. Using this function, learners can understand what, where and when they learned before. Figure 3(2) and Figure 4 (left) show the list of the learner’s ULLO, which helps him/her to recall all of the past ULLs. Besides, it allows the learner to be aware of others’ learning objects and to re-log them if deemed useful. This means that a learner can make a copy of another learner’s learning object into his/her own log. Therefore, learners can obtain a considerable amount of knowledge from others even though they have not experienced that knowledge themselves. By sharing ULLOs with other learners and relogging the other learners’ ULLOs, the acquisition of knowledge is enhanced.

ULL reminder

As shown in Figure 3(3) and Figure 4(right), the system generates simple multiple-choice quizzes based on the metadata of the stored ULLOs. For example, the idea of “quiz with image” is to ask a learner to choose an image that describes the word given by the system. The system immediately checks whether the answer is correct or not. These quizzes are generated according to the user’s profile, location, time and the results of past quizzes and helps learners to recall what they have learned (Li et al., in press). The quiz function is designed not only to help learners to reinforce what they have learned, but also to recommend what other learners have learned and to remind them of what they learned in the past according to their current location and their preferred time. In order to achieve these targets, they can take quizzes whenever they want. In addition, they can send their location information to the server continuously. Therefore, the server side can automatically assign quizzes for them based on their location and time information. It notifies them by showing an alert message and vibrating the mobile phone. Whenever they move around an area where they have encountered some objects, the system will send them quizzes regarding those objects. Furthermore, they can set a time schedule to receive the reminder quizzes.

ULL navigator

The ULL navigator provides mobile augmented reality that allows the learner to navigate through the ULLOs. Like Wikitude (“Wikitude,” n.d.) and Sekai-Camera (“Sekai Camera Web,” n.d.), it provides a learner with a live direct view of the physical real-world environment augmented by a real time contextual awareness of the surrounding objects.
When learners are moving around with their mobile phones, the system sends alerts to the phone as soon as they enter the region of ULLOs according to the GPS data. This view is augmented, associated with a visual compass, and overlapped by the nearest objects in the four cardinal directions. Also, it provides the learners with a list of all surrounding objects. When the learner selects one or more of these objects, the Google map will be retrieved and marked with the learner’s current location and the selected objects. Moreover, the system shows a path (route) for the learner to reach the locations of the objects. This assists the learner in acquiring new knowledge by discovering existing ULLOs and recalling the learner’s own ULLOs.

An important component of the system is the ULL recorder. The current ULL recorder requires learners to capture learning contents (e.g., photos) manually. They might find this troublesome and it might disturb their learning activities. Thus, we have tried to find a better way to log learners’ learning content automatically and unconsciously. We found that SenseCam is able to do this. However, the second problem is: Among the very large amount of captured photos by SenseCam, what are the learning contents? After using SenseCam, we found that there are many kinds of context data such as temperature and brightness that can be used to help learners recall the captured objects. Besides, these data can also be used to help us improve our ULL reminders. This is an extra benefit of the passive capture of data.

**PACALL**

**Design**

Until now all the work that we have done has been using the active rather than the passive logging mode. This means that learners must record their learning experiences consciously. Compared to the passive mode, in the active mode we are more likely to miss learning chances since we are not necessarily able to record what we have learned, or sometimes we just forget to record it. Therefore, we introduced passive capture in our project with SenseCam and named the proposed system PACALL (PAssive C Apture for Learning Logs). In the real world, there are so many things that we have learned but we usually miss the chance to review them; that is, we do not know what we know. Similarly, it is certain that we are not able to learn what we have not noticed. Therefore, we considered this in the learning process.

Since this research is based on our previous work (Ogata et al., 2010) in which we used the active mode to register ULLOs, we need to make it clear how the passive mode differs from the active mode. We compare the features of both in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Comparison of the active and passive data capture modes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Passive</strong></td>
</tr>
<tr>
<td>Number of photos taken</td>
</tr>
<tr>
<td>Data quality</td>
</tr>
<tr>
<td>Recording time distribution</td>
</tr>
<tr>
<td>Content completeness</td>
</tr>
<tr>
<td>Consciousness</td>
</tr>
<tr>
<td>Reflection</td>
</tr>
<tr>
<td>Workload</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

In the first three rows of Table 1, the two modes are compared in terms of photos. When we use SenseCam, it takes photos automatically and continuously, while in the active mode, smartphone photos can only be taken at the time we intend to. As a result, more photos are taken in the passive mode than in the active mode. However, because of the storage problem and some other technology limitations, photos taken by SenseCam are lower in quality; they are however of sufficient quality to be used as ULLOs.

In the next two rows of Table 1, the comparison is made in terms of learning contents. When we use a camera or smartphone, many learning contents are logged in our spare time such as at lunchtime. However when we use
SenseCam, because the recording is processed continuously, we can get photos whenever and wherever we are. As such, the photo contents cover our complete lives. The content type in the active mode, however, is richer than that in the passive mode because in this research the learning content captured by SenseCam consists only of photos.

In the last two rows, the comparison is related to the learners. In the passive mode photos are taken unconsciously, while in the active mode the learner must take photos consciously. When learners use SenseCam, they must review the whole learning process, and reflect on what they have seen and what they have learned and what they missed learning. This process will help them to remember the contents.

According to this comparison, we can see that the passive mode has many advantages over the active mode for language learning by photos. The quality of the photos is low, but it is still acceptable for our purpose. However, the biggest disadvantage is the workload. SenseCam takes photos continuously. Consequently, a huge number of photos are produced, and the more photos, the heavier the workload. If this workload is reduced, learners can learn language in the passive mode more easily. This is the key issue of using the passive mode in language learning. In this research, we focus on reducing the workload when reviewing the photos, and propose a system that can filter the photos to help learners review and upload ULLOs easily.

Figure 5 explains this process and shows how to support learning in the passive mode. We classify all the objects surrounding us into four groups – “(I) I know what I know,” “(II) I know what I don’t know,” “(III) I don’t know what I know,” and “(IV) I don’t know what I don’t know.” For example, for non-English speakers, when a learner walks outside and sees a fire hydrant, if he notices it and knows how to say it in English, he is in status (I). If he does not know how to say it in English, he is in status (II). Since he has noticed it, and does not know how to say it, he can learn it from a dictionary or by asking someone else. Then the status will change from (II) to (I), which only happens consciously in active mode.

Another situation is the case when he did not notice it. If he has not noticed the fire hydrant how can he learn it? The answer is that he can’t without some form of assistance. Therefore, we would like to encourage learners to use passive mode to support their learning. There are also two kinds of “don’t know” situation. In one case, the learner already knows how to say it in English (status III). In this case, captured life-log photos can help him notice this fire hydrant and let him review it. In status IV, captured life-log photos let him know there is an object that he does not know, and then he can have a chance to learn this object (B to C). This is a good way to help learners become aware of what they do not know if they do not notice it.

This system is a sub-system of the Ubiquitous Learning Log, named PACALL. It stands for PaSSive CApture for Learning Log. We set the whole capturing process in passive mode. After that, PACALL analyzes all of the captured photos and finds several important ones to help learners determine which are worth recording.
Figure 6 is the flow of analyzing captured photos. It consists of 5 steps: Loading raw data, Filtering bad photos, Finding good photos, Photo recommendation, and Learning analytics. These steps are introduced in detail in the following.

**Loading raw data**

There are three types of raw data in PACALL: life-log photos, sensor data, and GPS data. Life-log photos are currently captured by SenseCam. In the future, we plan to apply this system to photos taken by smartphones or compact digital cameras which are far more commonly used. That will be more convenient and useful. Suppose a learner took a trip and took many photos. Then she can use this system to find photos which contain useful learning contents.

The sensor data are recorded by SenseCam, and the GPS data are created by the portable GPS unit.

**Filtering bad photos**

In this research, a bad photo is defined as a photo that is hardly recognizable or that is a duplicate of other photos. We define three types of bad photos:

- Dark: a dark photo taken with insufficient light.
- Duplicate: the photos are duplicated.
- Defocused: the photos are blurry and cannot be recognized well.

We use image processing to identify these bad photos. Currently, we are using OpenCV to detect dark photos, and LIRE (a plugin for Lucene) to detect duplicated photos.

**Finding good photos**

A good photo is defined as one that contains clear objects. We use OpenCV to find good photos mainly by feature detection.

After filtering bad photos and identifying the good photos, the rest are of mediocre quality. Those photos might contain learning contents, although they are not so clear. The top priority is given to good photos and then mediocre ones come next when shown to learners to choose.
Photo recommendation

Once photo sorting is finished, the next stage is learning assistance. Our challenge is to detect useful information from photos by machine, and recommend photos that contain information. We define four types of recommended photos:

Character photo: a photo that contains characters. These characters are possibly used as learning contents. Here we are using text detection to find these photos.

Face photo: a photo containing a face. Actually, these photos are usually not appropriate for learning content because of privacy issues. However, faces are also information from photos.

Taggable photo: a photo that can be tagged by text. Tags are an important piece of information and can be used as a title of the photo.

ULLO-like photo: If there is a similar photo that was already registered to the SCROLL as a ULLO, it can possibly be used as a ULLO as well.

System interface

In the previous section, we introduced our design of the flow of analyzing photos. In this section, we explain its functionalities (i.e., PACALL Uploader, PACALL Browser and PACALL Recaller) in detail. PACALL Uploader helps learners upload all the photos after capturing. We have made it easy to upload all the captured photos to the server. Because of the limitation of web technology, this process was not easy in the past. However with HTML5, it became possible. When learners want to upload a whole folder, they can select a photo folder and upload all the photos to the server. Also, the file of sensor data and GPS data will be uploaded.

After uploading the raw data (photos, sensor data and GPS data), the system will analyze all the data and show the results.

![Figure 7. Interface of PACALL Browser](image)

When all the photos are uploaded to the server, the learner can reflect on them with help from PACALL. The PACALL Browser has an interface for browsing all the photos, and it tags photos and provides some information to
help the learner find important photos (Figure 7). Currently, we provide three main functions in the browser – PACALL Filter, PACALL Searcher and PACALL Recognizer.

PACALL Filter classifies all the photos into categories such as Manual, Normal, Duplicate, Dark, Face and Recommendation. Here “Manual” means that a photo is taken manually by pressing the manual button of SenseCam. It usually happens when a learner finds something valuable to record. “Duplicate” and “Dark” mean bad photos. “Face” means the photos that contain faces, and “Recommendation” includes Manual, Faces and other good photos that contain information or similar photos that have been uploaded to SCROLL before. Such photos have tags under them such as 3d or 4d meaning that they were uploaded to the system three or four days ago.

When a learner clicks one photo in PACALL Browser, the PACALL Recaller will be opened. The photo and similar photos and sensor data will be shown on this page to help the learner recall the captured content. There is also an “Upload” button on this page. If the learner decides to upload this photo to SCROLL as a ULLO, he/she can click this button, and the photo will be uploaded to the SCROLL system directly and the page will jump to the learning log registration page (Figure 8). Figure 8 shows the interface of ULLO registration in the SCROLL system. On this page, a learner can see the location of the selected photo and other similar photos captured by SenseCam. If there are some similar photos that are already uploaded in SCROLL, they will also be shown on this page. Once “Upload Now” is clicked, the system will ask the user to answer a survey that lets the system know whether he/she knows it and whether he/she noticed this object when it was captured. The data can be used to evaluate our system and help the user analyze his/her learning situation. When an object is uploaded to the system, the SCROLL system will use the “organize,” “recall” and “evaluate” model to help users remember uploaded objects. For example, if a learner uploaded a photo and set the title as 消火栓 in Japanese, but does not know how to say it in Chinese, he/she can send a question along with the uploaded ULLO. SCROLL will send this question to all Chinese users. After receiving answers from them, the user can learn a new Chinese word. In the quiz module of SCROLL, a learner can answer quiz questions that are created automatically from uploaded ULLOs. By answering these questions, the learner’s knowledge will be enhanced.

![Figure 8. PACALL Recaller](image)

**Evaluation**

We have conducted an evaluation experiment for PACALL. This section introduces the method and result of this experiment.
Method

Since this is an initial evaluation experiment, the study group consisted of 4 Japanese university students taking an undergraduate English course. In this experiment, they were asked to upload photos of learned objects along with titles both in Japanese and English. They used three methods of recording the photos. The entire evaluation experiment lasted for 3 weeks and consisted of 3 phases:

Phase 1: SenseCam

During this phase, students were asked to wear SenseCam every day for one week. Every evening, they needed to review all the life-log pictures and choose proper pictures to upload to SCROLL. They were requested to record the time spent.

Phase 2: Tablet PC

In our previous work (Ogata et al., 2012), we compared the learning effectiveness of Tablet PCs and a traditional learning method such as taking notes. It was found that using SCROLL on a Tablet PC was more effective than the traditional learning method. In this experiment, we compare SenseCam with the Tablet PC in terms of log methods, as logging with a Tablet PC is considered as active, while logging with SenseCam is considered as passive.

During this phase, all the students were asked to record and upload the learning log objects every day using a Tablet PC for one week. We used a Samsung Galaxy Tab in this experiment. The operating system of this Tablet PC is Android, and we developed an Android client that can upload the photos to the system conveniently.

Phase 3: SenseCam + PACALL

During this phase, the PACALL system was introduced into the experiment. This phase was almost the same as Phase 1. All the students should wear SenseCam every day for one week. Every evening, they used the PACALL system to classify the life-log pictures and upload the pictures. They were requested to record their time spent. Besides, they were asked to count the number of classifications after all the pictures were uploaded.

Result

Learning chances - Differences in number of ULLOs uploaded per day

We examined the number of uploaded ULLOs among these three learning methods. Figure 9 shows the average number of uploaded ULLOs for each of the four participants.

![Figure 9. The average number of uploaded ULLOs](chart.png)

In this chart, the horizontal axis shows the four subjects, and the vertical axis shows the average number of ULLOs uploaded by each subject. It shows that they uploaded more photos in the SenseCam and SenseCam + PACALL
modes than in the Tablet PC mode. As a result, the passive mode with SenseCam offered them more learning chances than the active mode. Moreover, it is found that PACALL increased the number of uploaded pictures in most cases (except S2). After this experiment, we examined why the number of uploaded pictures did not increase for S2, and then we interviewed him. We learned that at that time he was not so serious about the experiment and just managed to upload 3 pictures a day as a norm. However the result of Figure 11 shows that it took him nearly half the time to review and upload photos in SenseCam + PACALL mode compared with the time spent in SenseCam mode, so it is expected that if he had been more involved and spent more time, the number of uploaded ULLOs would have increased. In normal circumstances in the subjects’ daily lives, the pictures captured by SenseCam were similar whether using PACALL or not, and the numbers of uploaded objects were almost the same. However, from the feedback, it was found that PACALL reduced the workload of reviewing life-log pictures. The learners could choose pictures and upload them more quickly. So the number of uploaded pictures using SenseCam + PACALL is larger than that using SenseCam only.

**Learning quality - Can learners remember uploaded ULLOs that are taken unconsciously in passive mode?**

What is the difference in the learning effect of the active and passive modes? This is the second question that we attempted to answer. Therefore, we gave all the students a test after each phase to see whether they had remembered the uploaded objects. We devised a test consisting of the uploaded pictures and asked them to write down the title of the pictures, then judged their memory, grading them from 0 to 5, where 5 means they remembered clearly and 0 means no recall at all. Figure 10 shows the results.

![Figure 10. Memory level of active and passive modes](image)

In Figure 10, we can see that the memory level of active mode is a little higher than that of passive mode. In active mode, the learner takes pictures consciously. Naturally the impression of the photos is deeper than that of those taken in passive mode. Besides, the number of uploaded photos in passive mode was larger than that of active mode which might be reflected in their memory level. Therefore this result is understandable. Even though their memory level was lower, when we consider the fact that they registered more photos as their learning logs, we interpret that our system gave them more chances to learn in passive mode. Therefore we believe that the system contributes to their learning.

**Workload issue - How much value does the PACALL add to passive learning mode?**

We examined how PACALL contributed to reducing the time spent on the whole procedure. We asked students to report the spent time, and Figure 11 shows the result.

This chart clearly shows that the developed system reduced the time spent reviewing the life-log pictures by nearly a half. Of course, the workload in passive mode is higher than that in active mode, but very few learning contents are missed. In the future, we will focus on how to reduce the workload and help learners to reflect on and find good photos more easily.
Feedback

We received some suggestions and feedback from the students which helped us to understand the usage of PACALL and to improve our system. Some typical feedback is listed as follows.

I think PACALL is easy to use. When I use the SenseCam without PACALL, I must find good photos in the folder from the browser. However when using PACALL, I just select them and click “upload.” The time is shown with the photo in PACALL, which is also helpful for selecting photos. Besides, inappropriate photos are already excluded by this system. It also helps.

It is better to use the Android Tablet PC in conjunction with the PACALL system.

In the passive mode, the learning contents are recorded even when I do not want to learn anything. On the other hand, in the active mode, photos can only be taken when I want to learn something.

I feel very embarrassed when using SenseCam.

The accuracy is not good enough for analyzing blurred photos.

The above comments show that this system is easy to use and the users seem moderately satisfied with the system. In the passive mode the learning contents could be recorded even if learners do not want to learn. In other words, the life-log pictures create more chances to learn vocabulary. However, the learners may feel embarrassed when wearing the SenseCam in public. Moreover, there is a privacy issue which is yet to be resolved. As for the problem of embarrassment, in the future, we believe that the SenseCam will get smaller and look better, and hopefully, learners will not feel so embarrassed wearing it. Besides, the algorithm for classification needs to be improved in the future.

Conclusion

This paper describes a ubiquitous learning log project in Japan called SCROLL. This project was partly supported by PRESTO of the Japan Science and Technology Agency (JST) from 2010 to 2013. It aimed to capture learning experiences in daily life and reuse them for learning and education. Especially, this paper focuses on capturing ubiquitous learning logs using life-log photos, which are automatically taken by SenseCam. We developed a system named PACALL to help learners find learning contents from life-log photos. Also, we took a further step in analyzing life-log photos for the main system. Therefore, PACALL is not only a learning content provider but also a learning content analyzer. Besides, the provided data of PACALL can also be used by the quiz module of SCROLL to determine the proper time for presenting learners with quizzes. We found that there are many useful pieces of information that can be mined from the life-log photos. In the future, we will improve the algorithm of image processing in this system and conduct an evaluation experiment.
Since SCROLL is intended to be used in general domains and for life-long learning, we will apply it to many application domains including foreign language, math, physics, and science education, and conduct a long-term evaluation with a larger sample of subjects. Another area of our future work is learning analytics. We plan to analyze the accumulated data in the learning logs to find learners’ learning patterns and learning habits in order to supply more appropriate learning materials in more appropriate places and at more appropriate times to improve the learning effects of the system.

Acknowledgements

This research work was supported by PRESTO from the Japan Science and Technology Agency, and the Grant-in-Aid for Scientific Research No. 21650225 from the Ministry of Education, Science, Sports, and Culture in Japan.

References


Context-Aware Mobile Role Playing Game for Learning – A Case of Canada and Taiwan

Chris Lu¹, Maiga Chang¹, Kinshuk¹, Echo Huang² and Ching-Wen Chen²

¹School of Computing and Information Systems, Athabasca University, Canada // ²Department of Information Management, National Kaohsiung First University, Taiwan // crischien630@gmail.com // maiga@ms2.hinet.net // kinshuk@athabascau.ca // ecoh@nkfust.edu.tw // chingwen@nkfust.edu.tw

*Corresponding author

ABSTRACT

The research presented in this paper is part of a 5-year renewable national research program in Canada, namely the NSERC/iCORE/Xerox/Markin research chair program that aims to explore possibilities of adaptive mobile learning and to provide learners with a learning environment which facilitates personalized learning at any time and any place. One of the sub-projects of this 5-year national research program is to design and develop context-aware mobile learning services. The research team of the sub-project applied narrative theory to design a location based Context-Aware Mobile Role Playing Game (CAM-RPG) in order to give students feeling of living in the game world and role playing, exploring the game world, completing the quests, and learning things. A pilot study was then conducted to see how the two game features – context-awareness and story generation – influence students’ attitude towards the use of the mobile educational game. The research findings suggest that the story generated in CAM-RPG positively influences users’ attitude towards game use and increases users’ perceived game usefulness. With the research findings, other components and outcomes of sub-projects, such as natural language processing, location-awareness, multiple input forms, social networking, and student modeling, can then be put together as one piece to provide students effective and efficient mobile learning experiences.

Keywords

Context-awareness, Location-based, Narrative theory, Educational game, Mobile game, Role-playing game, Technology acceptance model

Introduction

The exponential growth of wireless technology in recent years, increasing availability of high bandwidth network infrastructures (e.g., the SuperNet in Alberta), advances in mobile technologies and the popularity of handheld devices have opened up new accessibility opportunities for education. This has given rise to a five-year research program, funded by Canadian federal government and Alberta Provincial government in collaboration with various industry partners. The program aims to explore and develop different applications and content delivery systems, extending our understanding of mobile learning to provide rich learning experiences in order to not only improve the existing educational environment but also to widen access to education for the disadvantaged, particularly those living in remote and rural communities, who generally do not have access to learning opportunities and the disabled, who need specialized devices and applications for learning.

The learning environment that is being developed under this research program consists of different servers and databases, and provides several services for students. The location-awareness service is aimed to help mobile students forming face-to-face-learning groups. Moreover, innovative social networking functions are integrated in the learning environment. An adaptive mechanism is also developed that is responsible for providing learners with learning materials that fit their individual learning styles. The context-awareness service identifies the personalized context-aware knowledge structure in an ubiquitous/pervasive learning environment and is aimed to direct individual learners to learn and move in the real world using automatically generated guidance messages. Furthermore, learners are supported by an intelligent and multimodal asynchronous questions & answers (Q&A) knowledge sharing platform.

The program has three stages. The first stage consists of Canadian research team designing and developing the game and Taiwanese team designing and conducting the pilot study to verify the usability of the game. The second stage involves Canadian team improving the game according to the feedback received in the first stage and conducting a pilot study in Canada for both the iterative development process and cultural difference investigation. The last stage of the program involves application of the well-designed final product in a formal class and a comparative experiment involving both Canadian and Taiwanese students. The program is currently at the end of first stage.
In 2010, the research team of context-aware sub-project developed a Context-Aware Mobile Educational Game (CAMEG) (Lu, Chang, Kinshuk, Huang, & Chen, 2010a, 2010b). The game generates a series of learning activities (i.e., a learning activity chain) to enable students to interact with specific real objects (e.g., projector, restroom, pine tree, etc.) and virtual objects (payroll system, business policy, E-Commerce course, etc.) in authentic environments. The series of learning activities is automatically generated for individual students according to their learning history and surrounding context (i.e., learning objects associated with the chosen role that the student wants to play, the chosen learning theme, student's location, etc.).

However, majority of the existing educational games, including mobile games, have not looked at such individual feelings. Focus has primarily been on how to teach specific discipline or curriculum in formal educational and on-the-job training settings (i.e., workplace, school campus, museum and historical site). These games become boring when students are simply asked to conduct certain activities one-by-one repeatedly. Few researchers have talked about how to design the contents of mobile educational games in order to make them attractive for the students. This paper focuses on this aspect with aim to improve effectiveness of the mobile educational games.

The rest of the paper is organized as follows. The next two sections introduce the research background by reviewing relevant literature on educational games. The research model and hypotheses used in this research are then described. This is followed by the description of the pilot design and the collected data. Statistical analysis methods are then used to find the answer to the research question. Finally, the implications of the findings are discussed and conclusions are drawn.

Background and motivation

In the last decade, many researchers have seen mobile learning (m-learn) as a further evolution of e-learning (Georgiev, Georgieva, & Smrikarov, 2004). Unlike computer-based learning (learning at a specific place with desktop computers), m-learning delivers education and training materials to a variety of lightweight devices such as personal digital assistants (PDAs), tablet PCs, smartphones, and mobile phones, which users can comfortably carry and use for learning anywhere, at anytime (Keegan, 2005). Beyond the learning devices, some researchers also think that the context of pedagogy differs between e-learning and m-learning. Especially for environmental sciences, m-learning brings potential benefits for learners' self-learning by realizing real-time and location-based learning materials (Jones, Scanlon, & Clough, 2013; Vogel, Spikol, Kurti, & Milrad, 2010).

Brown and colleagues argue that students can learn specific knowledge more efficiently by interacting with a situated environment (Brown, Collins, & Duguid, 1989). Learners can observe or touch the learning objects and can interact with the m-learning system immediately. Hwang, Yang, Tsai, and Yang (2009) also point out that context-aware learning is an innovative approach for detecting student situations and providing students personalized services and adaptive support. Wu and colleagues argue that context-aware ubiquitous learning enables students to interact with learning objects in the real world with the supports from the digital world (Wu, Hwang, & Tsai, 2013). It is important for a mobile learning system to be context-aware; hence, the research team decided to create an interesting context-aware mobile game for students learning domain knowledge.

Garris, Ahlers and Driskell (2002) applied an instructional model in games that uses game-feature-relevant instructional contents as inputs and makes the game-play a cycle. In this model, the repeatable judgment-behavior-feedback activity is a game cycle. These repeated activities can increase the student's motivation and enjoyment of playing the game, enable students to play the game continuously, and increase students' confidence in the gameplay (Garris, Ahlers, & Driskell, 2002).

Researchers have also identified the importance of story in the games (Connors, 2013; Simon, 2012; Sanders, 2011). Connors (2013) argues that story is fundamental for players remembering their gaming experiences and a game might be less impactful without the story. Simon (2012) argues that players may perceive two games to be exactly same if the games have no story, which would have negative effect on learners’ motivation to come back to play the games. Sanders (2011) argues that story can make players aware of the goal of the game and can keep them exploring the game.
In order to make the Context-Aware Mobile Role Playing Games (CAMEG) interesting for the users and to motivate them to play, narrative elements were taken into consideration in this research. Narrative theory covers the elements that a story needs (Conle, 2003); therefore, it was decided to design the story generation engine based on narrative theory. The narrative elements such as storyline, character, and interaction have been analyzed in the literature and used in the game-based learning system design (Ying, Wu, Chang, & Heh, 2009). Researchers have also integrated various narrative elements and designed different approaches for generating story (Akimoto & Ogata, 2011; Akimoto & Ogata, 2012). The research team therefore applied narrative theory to enhance CAMEG in order to give students feeling of living in the game world and role playing, exploring the game world, completing the quests, and learning things. At the end, the enhanced mobile educational game with stories - Context-Aware Mobile Role-Playing Game (CAMRPG) was developed in 2011 (Lu, Chang, Kinshuk, Huang, & Chen, 2011c).

The research team has subsequently been tackling the following research question: do the two game features – context-awareness and story generation – really influence students’ attitudes towards using such educational mobile role-playing games? A pilot study has been conducted, where a questionnaire (and associated statistical analysis) was employed to gather students’ attitudes toward the game.

**Story decorated context-aware mobile role-playing game**

To develop a lightweight, flexible, and scalable mobile educational game based on the research components designed by various sub-projects of the research chair program, a multi-agent architecture (MAA) has been used (Lu, et al., 2010b, 2011a). A multi-agent system is a software environment containing many agents who are responsible for their own tasks while collaborating with other agents whose responsibilities belong to the pre- and post-requisite tasks. Multi-agent architecture is particularly useful for developing mobile applications because it can divide a complex task into several smaller tasks and can assign these tasks to different agents. Moreover, these agents can work either within same device (e.g., a mobile phone) or on different machines/platforms as a distributed system (Balaji and Srinivasan, 2010).

Figure 1 shows the multi-agent architecture of the mobile educational game developed in this research. More details for the responsibilities of each agent and the collaboration among agents can be found in Lu, et al. (2011a)

Figure 2 shows the screenshots of the game-play of CAMRPG. During the game-play, the Player Agent is the only agent that interacts with the user and enables data exchange between the user and other agents.

---

**Figure 1. Multi-Agent architecture of the proposed mobile educational game**

**Figure 2. Screenshots of CAMRPG game-play**
As shown in Figure 2, in the game, roles and corresponding pre-defined themes are designed for students to have opportunity to choose what learning direction and discipline they really need. For instance, a student who takes Introduction to Management Information System course may want to know more about what enterprise support system is and what benefits a business can gain from it. In such circumstances, the student can choose a particular role and theme s/he wants to play, for instance, a chief information officer (i.e., step 2).

The game then generates learning activities for the individual students according to the chosen role and theme. Before the students are asked to do the learning activities, the game makes stories up automatically and uses the stories to populate the generated learning activities for the individual students (i.e., step 3). After the students finish reading the story, the game shows them the learning activities (i.e., step 4). The students can use the built-in camera to collect the required learning object(s) by taking pictures of the objects’ QR codes (i.e., step 5). Once the game has verified the correctness of the collected learning objects, it delivers each student a piece of text-based learning material about the corresponding learning object (i.e., step 6). In addition to text-based learning contents, the learning contents can also be HTML-based, binary-based image, URL of webpage, media stream and Flash animation.

At the end, the game checks if the students have completed all generated learning activities for the chosen role and theme. If there are other activities left, the game takes the individual students back to step 3 to read another story and asks them to finish another learning activity (as flow B on Figure 2 shows). If no activity is left, the game takes the students back to theme selection screen (as flow A on Figure 2 shows). The students can then either choose another theme or can even take another role.

**Research model and hypotheses**

The research team decided to explore the connection between student's perceived usefulness of the game and the two features (i.e., context-awareness and story generation) step by step, with the following research question "do the two game features – context-awareness and story generation – really influence students' attitudes towards using educational mobile role-playing games?"
A number of models have been proposed in the literature for analyzing user perceptions and acceptance towards technological systems. A well-established and tested model in the literature is the Technology Acceptance Model (TAM), proposed by Fred D. Davis in 1986. This model has become one of the most common instruments used to explain the users' behavioral intention of using an innovative technology. Original TAM has four constructs: the perceived ease of use, the perceived usefulness, the attitude toward using the innovative technology, and the behavioral intention of using the innovative technology.

Some researchers have also examined the acceptance factors for educational games or entertainment games by adding their own variables to the original model to explore the influences of different external variables, for instance, gender, gaming experience, learning opportunities and the unified theory of acceptance and use of technology (UTAUT) (Bourgonjon, Valcke, Soetaert, & Schellens, 2010; Ibrahim, 2011). In the pilot study of this research, two external variables (i.e., the two game features, namely context-awareness and story generation) are proposed for inclusion in the original TAM.

The proposed research model is adopted from the research done by Ibrahim (2011) and Bourgonjon et al. (2010). Different from previous models, this research has four moderators, namely gender, gaming experience, smartphone experience, and context-awareness feature as variables. The reason for taking smartphone experience into consideration is to analyze whether or not the students who do not have experience in using smartphone encounter difficulty in using the game and perceive low ease of use than the students who have experience in using smartphone. Figures 3 and 4 show the macro view (i.e., all considered theories) and micro view (i.e., the detailed constructs) of the proposed research model respectively.

![Figure 3. Macro view of the proposed research model](image1)

![Figure 4. Micro view of the proposed research model](image2)
The hypotheses needed to be verified in the research model are listed in Table 1.

<table>
<thead>
<tr>
<th>Macro view</th>
<th>Micro view</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>H1: Attitude has a positive effect on behavioral intention.</td>
</tr>
</tbody>
</table>
| H2 | H2a: Perceived ease-of-use has a positive effect on attitude toward using CAMRPG.  
|  | H2b: Perceived usefulness has a positive effect on attitude toward using CAMRPG.  
|  | H2c: Perceived ease-of-use has a positive effect on perceived usefulness. |
| H3 | H3a: Context-awareness feature has a positive effect on attitude toward using CAMRPG.  
|  | H3b: Story generation feature has a positive effect on attitude toward using CAMRPG. |
| H4 | H4a: Context-awareness feature has a positive effect on perceived usefulness.  
|  | H4b: Story generation feature has positive effect on perceived usefulness. |

**Pilot design and data collection**

The purpose of the pilot study was to analyze whether a mobile learning system with the two features improves learners’ willingness of using it. Initially, the researchers introduced the game and conducted a demonstration in a Management Information System (MIS) class at a national university in Taiwan. The researchers explicitly told the students that there was no compensation, reward, or recognition for anyone who participated in the study. It was also made clear that there were no consequences for not taking part in the study.

The experiment environment of the pilot study consisted of three laboratories located within one building of the university. Since all participants were taking undergraduate level MIS course at that moment (June, 2011), the MIS course contents and concepts were incorporated into the game and a virtual science park was built in that building. The park consisted of many famous IT businesses and companies that virtually resided in the park, and participants interacted with those organizations in the virtual park while playing the game.

The participants were asked to complete a demographic questionnaire before playing the game. All participants had 20 minutes to play the game in the authentic learning environment using the smartphones prepared by the researchers.

As the participants started to play the game, they received story-enhanced learning activities and looked for the required learning objects in the real world. The learning objects were associated with MIS topics/concepts and were presented in different formats, such as video clips, presentation slides, case studies, and real systems. In the gameplay, participants acted as information technology (IT) experts and received quests from their boss (i.e., a non-player-controlled character). The quests asked them to visit the science park (i.e., the authentic learning environment) and collect some important information for their company. While they were playing, they would learn about these learning objects actively through presentations and demonstrations instead of sitting passively in a classroom and receiving lectures from the course instructor.

After the game-play, they were asked to fill out the technology acceptance model questionnaire. The questionnaire had thirty one five-point Likert-scale items (ranging from 5 for "strongly agree" to 1 for "strongly disagree") to address four main constructs of Technology Acceptance Model (i.e., perceived ease of use, perceived usefulness, attitude toward using, and behavioral intention of using), and two examined constructs (i.e., context-awareness and story generation).

**Reliability analysis**

The questionnaire was adopted from previous research results, and its validity and reliability have been proven by Lu, Chang, Kinshuk, Huang, and Chen (2011b). The data collected in this research was analyzed before using it to examine/verify the hypotheses. Some participants did not show up at the scheduled time, hence, the corresponding
responses of the questionnaire were removed. In addition, responses of two more participants were removed because they had extreme values for all questions and had conflicting answers for the flip-flop items. The final valid sample therefore included 62 students, consisting of 34 male and 28 female students.

Table 2 lists the results of reliability analysis. The Cronbach's alpha for the overall questionnaire is 0.826, indicating that the questionnaire (and its items) can be seen as reliable because its internal consistency is good enough (i.e., exceeds 0.75) (Hair, Anderson, Tatham, & Black, 1998).

The results showed that all constructs, except the behavior construct, had good measure of reliability. The three items of the behavior construct were reviewed and it was concluded that these items might not explain the construct well in this research because of the different subjective situations this research has. The three items had no correlation with the other constructs either. Therefore, the behavior construct was removed.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item number</th>
<th>Overall Cronbach's alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived ease of use (PEoU)</td>
<td>5</td>
<td><strong>0.743</strong> (0.774 after PEoU03 and PEoU04 removed)</td>
</tr>
<tr>
<td>Perceived usefulness (PU)</td>
<td>5</td>
<td>0.793</td>
</tr>
<tr>
<td>Context-awareness feature (CA)</td>
<td>5</td>
<td>0.752 (0.807 after CA02's removal)</td>
</tr>
<tr>
<td>Story generation feature (SL)</td>
<td>4</td>
<td>0.832</td>
</tr>
<tr>
<td>Attitude toward using CAMRPG (ATT)</td>
<td>4</td>
<td>0.807</td>
</tr>
<tr>
<td>Intention of using CAMRPG (IT)</td>
<td>5</td>
<td>0.894</td>
</tr>
</tbody>
</table>

Note. Bold and underline = Cronbach's alpha value is lower than 0.75.

Validity analysis

Next, the internal commonality of items for each factor in the research model was examined using principal component analysis. Three items – PEoU03, PEoU04, and CA02 – were found to have factor loading less than 0.6 and therefore not good enough for presenting the construct. It was decided to remove these three items. The removal of PEoU03 and PEoU04 also improved the Cronbach's alpha value of "Perceived easy of use" construct by bringing it to 0.774, and the removal of CA02 improved the Cronbach's alpha value of "Context-awareness feature" construct to be 0.807. The remaining items could then be used to represent the factors respectively. Lower factor loading may have occurred due to unclear questions or misunderstanding. They need to be revised to fit the presented constructs in future studies and experiments. Table 3 lists results of all constructs in principle component analysis.

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1: Perceived ease of use (PEoU)</strong> α = 0.743</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I2: It is easy to learn how to play</td>
<td>.850</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I31: The system flow is clear and simple to me</td>
<td>.817</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I32: The terms and functions in the game are easy to understand</td>
<td>.799</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I35: User interface are easy to use</td>
<td>.759</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I36: I can get familiar with the learning objects quickly</td>
<td>.580</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Factor 2: Perceived usefulness (PU)</strong> α = 0.793</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I36: It provides me enough information for what I want to know</td>
<td>.762</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I33: I can get needed information quickly within the game</td>
<td>.758</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I34: The learning activities can save my time in learning</td>
<td>.741</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I3: This game makes me want to explore the game's world</td>
<td>.733</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I37: This game provides me enough information for learning</td>
<td>.703</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Factor 3: Context-awareness feature (CA)</strong> α = 0.752</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I15: The learning objects are associated to my chosen theme</td>
<td></td>
<td>.844</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I16: If a quest required multiple learning objects, all of them can be found in the authentic learning environments</td>
<td></td>
<td>.761</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
At the end, a valid and reliable technology acceptance model questionnaire for measuring participants’ attitude towards mobile educational game with six constructs and twenty eight items was determined and confirmed. Quantitative statistical method was then used to get the answers for the research questions.

### Data analysis and results

In order to answer the proposed research question, descriptive statistics was initially used to summarize the collected data and compare the constructs' mean and standard deviation values for different groups (e.g., gender smartphone use and player types). Independent t-test was then used to explore whether or not different groups of participants have different attitudes toward CAMRPG.

#### Descriptive statistics

The demographic questionnaire collected participant's gender information, experiences of playing games, time spent in playing games, and experiences with smartphones. Table 4 lists basic information for the final sample of 62 participants.

<table>
<thead>
<tr>
<th>Gender</th>
<th>N</th>
<th>Smartphone(s) using experience</th>
<th>Playing video games</th>
<th>Playing handheld video games</th>
<th>Playing computer games</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>34</td>
<td>10 (29.4%)</td>
<td>30 (88.2%)</td>
<td>29 (85.2%)</td>
<td>34 (100%)</td>
</tr>
<tr>
<td>Female</td>
<td>28</td>
<td>9 (32.1%)</td>
<td>21 (75%)</td>
<td>22 (78.5%)</td>
<td>27 (96.4%)</td>
</tr>
<tr>
<td>Total</td>
<td>62</td>
<td>19 (30.6%)</td>
<td>51 (82.2%)</td>
<td>51 (82.2%)</td>
<td>61 (98.3%)</td>
</tr>
</tbody>
</table>
Table 4 shows that most participants had rich experiences of playing games, especially computer games. Video and computer games are both found to be major entertainment activities for them. In addition, only 30.6% of participants had experiences of using smartphones.

Quantitative analysis

Independent t-test was used to explore whether there were significant differences in technology acceptance between different groups of participants (e.g., gender, time spent playing computer games, and experiences of using smartphones). The statistical data analysis in Table 5 shows two meaningful results: (1) female participants have more positive feedback than male participants for all constructs; and, (2) there is no obvious difference between male and female participants in their responses for six constructs. The results are in line with the findings of previous researchers (Gwee, Chee, & Tan, 2010; Law, 2010; Papastergiou, 2009).

<table>
<thead>
<tr>
<th>Construct</th>
<th>Gender</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived ease of use</td>
<td>Female</td>
<td>28</td>
<td>4.3429</td>
<td>.40682</td>
<td>1.579</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>4.0765</td>
<td>.81205</td>
<td></td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>Female</td>
<td>28</td>
<td>4.3000</td>
<td>.37515</td>
<td>1.987</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>3.9882</td>
<td>.75629</td>
<td></td>
</tr>
<tr>
<td>Context-awareness feature</td>
<td>Female</td>
<td>28</td>
<td>4.0000</td>
<td>.29313</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>3.9882</td>
<td>.56126</td>
<td></td>
</tr>
<tr>
<td>Story generation feature</td>
<td>Female</td>
<td>28</td>
<td>4.0982</td>
<td>.51523</td>
<td>1.065</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>3.9412</td>
<td>.62480</td>
<td></td>
</tr>
<tr>
<td>Attitude toward using CAMRPG</td>
<td>Female</td>
<td>28</td>
<td>4.2589</td>
<td>.36945</td>
<td>1.519</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>4.0368</td>
<td>.74907</td>
<td></td>
</tr>
<tr>
<td>Intention of using CAMRPG</td>
<td>Female</td>
<td>28</td>
<td>3.9857</td>
<td>.60106</td>
<td>1.422</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>3.7471</td>
<td>.70032</td>
<td></td>
</tr>
</tbody>
</table>

From Table 5, it can be seen that although the mean values of both groups are quite high (positive) for all constructs, male participants have relatively higher standard deviation. This circumstance shows that male participants may have extreme high or low responses for these constructs. It is notable that the statistical analysis for experience of using smartphones shows no obvious difference between smartphone users and traditional mobile phone users.

Multiple regression

To explore the cause-effect relationships in the research model, a simple linear regression (i.e., use of attitude towards using CAMRPG to determine the intention of using CAMRPG) and several multiple linear regressions (e.g., use of perceived ease of use, perceived usefulness, context-awareness feature, and story generation feature to determine attitude towards using CAMRPG; and, use of perceived ease of use, context-awareness feature, and story generation to determine perceived usefulness) have been used. Such multiple regression analysis is typically used to examine and predicate the linear relationship between one dependent construct and one or more independent construct(s).

First, the independent factors were analyzed before entering the regression model in order to know whether there is a collinear problem in the statistics. A collinear problem is a statistical situation in which two or more predictors (independent constructs) in a multiple regressions are highly correlated. This situation causes an abnormally high R-square (i.e., explanatory power) in the regression model because the variances, standard error, and parameter estimates of predictors are probably inflated. It may also cause insignificant or incorrect coefficients (e.g., positive to negative) between predictors and affected variables.

The existence of linear dependence in the independent constructs can be determined by observing the collinearity statistic fields in Table 6. A collinearity statistic indicates that the construct may have a serious overlap (i.e., a collinearity problem, which means there is high correlation between the independent constructs) if the variance
The inflation factor (VIF) is over 10 and tolerance tends to zero (Hair, Anderson, Tatham, & Black, 1998). The results show that there is no serious collinearity issue between the independent constructs in Table 6, in which tolerance > 0.1, VIF < 10, and no two variables' variances > 0.8 at the same line.

### Table 6. Coefficients of perceived ease of use, perceived usefulness, context-awareness and story generation feature

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Unstandardized coefficients</th>
<th>Standardized coefficient</th>
<th>t</th>
<th>Significance</th>
<th>Collinearity statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>ß</td>
<td></td>
<td>Tolerance</td>
</tr>
<tr>
<td>(constant)</td>
<td>.441</td>
<td>.423</td>
<td>1.044</td>
<td>.301</td>
<td></td>
</tr>
<tr>
<td>Perceived ease of use</td>
<td>.296</td>
<td>.137</td>
<td>.323</td>
<td>2.162</td>
<td>.035$^{*}$</td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>.338</td>
<td>.146</td>
<td>.346</td>
<td>2.311</td>
<td>.024$^{*}$</td>
</tr>
<tr>
<td>Context-awareness feature</td>
<td>.015</td>
<td>.156</td>
<td>.012</td>
<td>.099</td>
<td>.921</td>
</tr>
<tr>
<td>Story generation feature</td>
<td>.248</td>
<td>.111</td>
<td>.234</td>
<td>2.235</td>
<td>.029$^{*}$</td>
</tr>
</tbody>
</table>

*Note. Dependent variable: Attitude toward using CAMRPG*

### Table 7. Collinearity diagnostics

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Eigenvalue</th>
<th>Condition index</th>
<th>Variance proportions</th>
<th>(constant)</th>
<th>PEoU</th>
<th>PU</th>
<th>CA</th>
<th>SL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.968</td>
<td>1.000</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>2</td>
<td>.015</td>
<td>18.274</td>
<td>.41</td>
<td>.12</td>
<td>.07</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>3</td>
<td>.009</td>
<td>23.118</td>
<td>.21</td>
<td>.06</td>
<td>.00</td>
<td>.00</td>
<td>.84</td>
<td>.01</td>
</tr>
<tr>
<td>4</td>
<td>.005</td>
<td>32.640</td>
<td>.30</td>
<td>.03</td>
<td>.32</td>
<td>.71</td>
<td>.01</td>
<td>.13</td>
</tr>
<tr>
<td>5</td>
<td>.003</td>
<td>38.685</td>
<td>.09</td>
<td>.79</td>
<td>.61</td>
<td>.28</td>
<td>.28</td>
<td>.13</td>
</tr>
</tbody>
</table>

*Note. Dependent variable: Attitude toward using CAMRPG. PEoU = Perceived ease of use; PU = Perceived usefulness; CA = Context-awareness feature; SL = Story generation feature.*

Table 8 lists the coefficients of four independent constructs towards the dependent construct – attitude towards using CAMRPG. Three constructs (i.e., perceived ease of use, perceived usefulness, and story generation feature) present significant coefficient measures ($\beta = 0.323, 0.346, \text{and } 0.234, p < 0.05$).

### Path analysis for the multiple regressions model

Figure 5 shows the path diagram of the research model. The result of path analysis shows that the attitude towards using CAMRPG (ATT) has strong effects on the intention of using CAMRPG as $0.455 (p < 0.001)$ of path coefficient. The effects of perceived ease of use (PEoU), perceived usefulness (PU), context-awareness feature (CA) and story generation feature (SL) explain 74% of the attitude towards using CAMRPG, while perceived ease of use (PEoU), perceived usefulness, and story generation feature have significant effects ($\beta = 0.331, 0.437, \text{and } 0.234, p < 0.05$) on the attitude towards using CAMRPG, but context-awareness feature does not ($\beta = 0.012$). For the cause-and-effect relationship between the independent variables, the effects of perceived ease of use, context-awareness feature, and story generation feature explain 75% of perceived usefulness, while perceived ease of use ($\beta = 0.678, p < 0.001$) and story generation feature ($\beta = 0.206, p < 0.05$) have significant effects on perceived usefulness, but context-awareness feature does not ($\beta = 0.066$).
Findings and discussions

Data analysis revealed several findings that can help in understanding users' attitudes towards and acceptance of the proposed mobile educational game as well as exploring the answer of the research question: do the two game features—context-awareness and story generation—really influence students' attitudes towards using educational mobile role-playing games?

These findings are categorized into three categories: common findings (i.e., those that have been proven in other research), important findings (i.e., those that are supported by this research), and unexpected findings (i.e., those that did not support our assumptions in this research).

Common findings

Findings suggested that the original technology acceptance model presents good results in cause-and-effect relationship of all factors (e.g., PEoU, PU, ATT, and IT). In particular, for the path coefficients found between perceived ease of use and perceived usefulness, the results indicate that ease of use is an important factor in context-aware mobile educational game design as well as other technology acceptance issues. Users appreciate a simple and easy-to-use interface, and a user-friendly interface directly impacts perceived usefulness. In addition, attitude towards using CAMRPG and intention of using CAMRPG also present strong significant coefficients in our research model. These findings have been proven in many studies that have focused on the acceptance towards information systems.

Important findings

First of all, the descriptive statistical data (i.e., Table 6) shows that the responses from both males and females were positive in terms of appreciation of the proposed CAMRPG. In addition, responses of female participants to all factors were relatively higher than those of male participants in the pilot.

The result did not show any significant differences between participants who have experience using smartphones and those who only have experience using traditional mobile phones. The reason may be that the participants in this pilot were undergraduate students and they were all familiar with mobile phones and games. Therefore, experience of
using smartphones did not affect acceptance of innovative technology. On the other hand, from the perspective of national research program, this result suggested that there is no need to worry about whether or not a user has used a smartphone while deploying such context-aware mobile role-playing game for learning.

**Unexpected findings**

From the path analysis results, it was found that most of the proposed factors qualified to explain the dependent variables, except the context-awareness feature factor. One reason perhaps is that the context-awareness feature is transparent to its users. For instance, a participant will receive from the game only those learning activities that involve learning objects in a library if the game detects that the participant is in library at that moment. So the participant would not feel what exactly the feature does for him or her. Another reason could be that the experiment environment in the pilot might not have represented the concept of context-awareness well enough to make participants aware of this game feature. For instance, the pilot was conducted using a virtual science park that was built in a university building used by the participants regularly for attending classes, which made it difficult for participants to have immersive feelings that they were in San Francisco or Helsinki. Finally, this pilot did not cover different buildings and did not continue over a longer time period. In such case, the participants could not experience scenarios like signing on to get quests at different places. These shortcomings might have caused the context-awareness feature factor to present relatively lower measures and an insufficient cause-and-effect relationship on the path coefficient.

**Conclusions**

This paper presented the outcome of the first stage of the context-aware sub-project, under the auspices of the 5-year national research chair program, namely the context-aware mobile role playing game, in which its kernel – learning activity generation engine and story generation engine – can automatically generate a series of story-based learning activities. This game can help users in learning by role-playing in authentic learning environments. The story makes up the learning activity chain resulting in more interesting and immersive learning process. Integrating story into a mobile educational game increases the perceived effectiveness and satisfaction toward the game, especially for the male students. On the other hand, the story reduces the perceived efficiency of using the system.

The findings indicate that participants in the pilot found the context-awareness feature of the game to be less important for the game-play and this factor did not affect their attitude towards using the game. The findings also identified the importance of authentic environment in mobile learning. The pilot study designed in this research clearly demonstrated that the context-awareness ability of the system was not even noticed by students, since they were asked to imagine a floor of a teaching building in school campus as a country in the world, hence there was a lack of an authentic environment. Such mismatch between the virtual and the real world has potential to reduce the perceived usefulness of the context-awareness functionality.

To make users aware of the advantages of a context-aware mobile educational game, subject selection (e.g., learning environment, selected learning topic, and learning materials) would be an important issue. The current game seems to work well for outdoor teaching/learning as well as learning based on treasure hunting paradigm at particular sites (e.g., museums, botanical gardens, and historical sites). It is also suitable for replacing orientation/training courses for freshmen and new students of the graduate programs. However, such game might not be suitable in environments in which learning objects have no strong connections to either the learning topic or the environment (e.g., trying to learn a business intelligent system from a desktop computer in a laboratory).

The pilot study encompassed only a short-term intervention whose effect may not be carried for long run. The research results also provided a clear picture of what learning topics may be more appropriate for applying context-aware mobile role-playing games, what authentic environment and learning objects are the best for deploying context-aware mobile learning systems, and what features are important to students who use mobile role-playing games.

The next stage research will focus on continuing the architecture design and proof of concept of the services developed under the national research program. This includes the incremental improvement of various modules.
based on proof of concept evaluations in Canada as well as continuing to integrate the developments within the overall system. As the research results show that the context-awareness feature is difficult to be noticed when the learning environment is a mix of mismatched virtual and real worlds, augmented reality may help in enhancing the perceived usefulness of context-awareness feature of the game. Also, in order to increase the effectiveness of the game, ordinary learning activities, such as field trips and remedial learning can be integrated into the game. Future plans of the context-aware sub-project include: (1) study and application of augmented reality concept within the interactive mobile learning systems in order to provide students the benefits of context-awareness; (2) develop Android version of CAMRPG and deploy it in rural areas for K-12 education and field trips; and, (3) integrate the outdoor remedial instructions and worksheet idea together to provide students an even more personalized ubiquitous learning experience according to their academic performances and the context surrounding them.

References


Potentials of Mobile Technology for K-12 Education: An Investigation of iPod touch Use for English Language Learners in the United States

Min Liu*, Cesar C. Navarrete and Jennifer Wivagg

Department of Curriculum & Instruction, The University of Texas at Austin, 1912 Speedway Stop D5700, Austin TX 78712, US // mliu@austin.utexas.edu // ccnavarrete@utexas.edu // jwivagg@keystoneschool.org

*Corresponding author

ABSTRACT

This case study investigated a m-learning initiative by a large school district in the United States to provide iPod touch devices 24/7 to teachers and students of English Language Learners. We described the initiative and presented the research findings of its implementation for two years at elementary and middle school levels. The results revealed the iPod touch was used to support language and content learning, provide differentiated instructional support, and extend learning time from classroom to home. However, several challenges were identified such as significant time demand on the teachers, technical issues, the need for professional training and dedicated support staff. Implications for teachers, instructional technologists, school administrators, and researchers were discussed.

Keywords

iPod touch, Mobile device, M-learning, English language learners, K-12 education

Introduction

Emerging mobile technology is having a transformative impact on how people live, learn, work, and play. According to 2012 Horizon Report for K-12 Edition (Johnson, Adams, & Cummins, 2012), “mobile devices & apps and tablet computing as technologies expected to enter mainstream use in the first horizon of one year or less.” Incorporating mobile devices in K-12 education has experienced increased interests and schools in the United States (US) are looking for new opportunities and possibilities introduced by the mobile technology. We describe an initiative by a large school district in the southwest region of US to provide iPod touches 24/7 to teachers and students of English Language Learners (ELL). We present research findings of its implementation for two years from 2010 to 2012 at elementary and middle school levels. ELL students are those who speak diverse languages such as Spanish, French, Portuguese, Chinese, and Japanese. With different levels of English proficiency (some have been in US schools for several years, while others are new to English language instruction), they represent distinct academic challenges in language acquisition. Especially in the region, Spanish-speaking students are a rapidly growing population. The goal of investigating iPod touch use over two implementation cycles is to gain an in-depth understanding of the potential benefits of such devices as a teaching and learning tool for the ELL population. The findings of this research have implications for teachers, instructional technologists, school administrators, and researchers as they explore and consider mobile technologies for K-12 education.

Relevant research

Mobile learning has garnered considerable interests from the educational community (Koole, 2009; Traxler, 2011) and literature on mobile learning has identified such affordances as flexibility, accessibility, interactivity, and motivation and engagement. Implementing flash card learning, for example, Kiger, Herro, and Prunty (2012) found third-graders improved their multiplication skills using iPod devices. Studies have shown Internet-enabled mobile devices can support cognitive learning (Peng & Chou, 2007); mathematics learning (Kalloo & Mohan, 2011); language and literacy learning (e.g., Coe & Oakhill, 2011; Kemp & Bushnell, 2011); and game-based learning (e.g., Liao, Chen, Cheng, Chen, & Cha, 2011).

For ELL students, ready access to information technology is considered a critical factor (Cummins, 2000). Mobile Internet-enabled devices can provide ELL students with access to learning resources for “comprehensible input” in language acquisition necessary for academic success (Cummins, 2000, p. 541). Using the mobile devices for audio with trade books provided by the school teacher-librarian in collaboration with teachers, Patten and Craig (2007) found iPod shuffle devices to support reading and writing with ELL students. By adding audio support to their text
reading and journal writing, the use of the iPod devices were also found to significantly increase student engagement and allow the students greater connection to the “U. S. popular culture” (Craig & Paraiso, 2007, p. 1840).

Despite the perceived affordances of m-learning, there is a paucity of research on using mobile devices with ELL students, especially how ELL teachers and students use such devices in classrooms for teaching and learning. This study investigates ELL teachers’ and students’ use of iPod touches in elementary and middle school classrooms as well as their perceptions of using it as a teaching and learning tool. Our research question is: In what ways can iPod touch devices serve as a teaching and learning tool for English Language Learners?

Research project

Research background

The mobile initiative, for which this research was conducted, took place in a large school district in the southwest region of US. The district has a 2011 enrollment of about 18,000 students in grades K-12 and covering an area of about 600 square miles. Of that population, about 4.5% are students with limited English proficiency. Spanish is the primary language spoken at home for about 90% of ELL students in the district, while the remaining ELL students speak a variety of other languages at home. Because of different levels of English proficiency, teaching ELL students is a unique academic challenge. In an attempt to address a substantial gap between the test scores of ELL students versus other students in the district, in September 2009 the school district purchased the latest mobile technology with Internet at the time - iPod touch - for every ELL student and teacher at the middle school level. By allowing ELL students, 24/7 utilization of the devices for accessing additional educational resources, the district is hoping to boost the ELL students’ English language proficiency. In this district, ELL students in grades K-5 are placed in bilingual instruction classrooms that consist of ELL students only. In sixth grade, ELL students transition away from bilingual classroom into mainstream regular education classroom where ELL students are mixed with non-ELL students with support from English as Second Language (ESL) teachers. In 2011-2012, this mobile initiative expanded from middle school level (grade 6-8) to 4-5 grades at elementary level for ELL teachers and students only. In this paper, we report the research conducted in two cycles: Year One implementation from 2010-2011 and Year Two implementation from 2011-2012.

Participants

The participants in this research were two ELL middle school teachers and their students for the first cycle (during 2010-2011 school) and two ELL elementary school teachers and their students for the second cycle (during 2011-2012 school year). Two sixth-grade teachers were selected because they were among the first to implement the initiative. Two teachers, fourth and fifth grade, were selected because they were part of the initiative during the second cycle and were teaching only ELL students. Teachers’ demographics are provided Table 1. In addition to iPod touch, all these teachers had access to other technologies such as an interactive whiteboard, a desktop computer, a Mac laptop, and a document camera in their classrooms.

<table>
<thead>
<tr>
<th>Data collection period</th>
<th>Level</th>
<th>Teacher</th>
<th>Years of teaching</th>
<th>Grade levels taught</th>
<th>Total students taught/managed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle one: 2010-2011</td>
<td>Middle school</td>
<td>Virginia</td>
<td>26</td>
<td>6-8</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Claire</td>
<td>7</td>
<td>6-8</td>
<td>31</td>
</tr>
<tr>
<td>Cycle two: 2011-2012</td>
<td>Elementary school</td>
<td>Clara</td>
<td>14</td>
<td>4-5</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lydia</td>
<td>23</td>
<td>4-5</td>
<td>42</td>
</tr>
</tbody>
</table>

Table 1. Demographics of the teachers
Data sources and analysis

The primary data sources were (1) interviews with the teachers, (2) classroom observations, and (3) surveys with the students.

Interviews were conducted with each teacher at three different times: the beginning of their iPod touch implementation, the mid-school year and toward the end of the school year. The goal of these interviews was to understand the teachers’ and students’ experience and perceptions toward iPod touch use. Questions varied because of the different focus for each interview during the year. Sample questions are:

At Beginning-of-Year:
- How are the students using the iPods?
- How are the students using the iPods?

During Mid-year:
- How are you using the iPod touch?
- Please describe how does the iPod touch help student learning? How does it constrain student learning? Provide examples.
- What, if any, challenges are there in developing iPod touch activities? Examples?
- How is iPod touch used outside of the school day?

At End-of-Year:
- How did you use the iPod touch in your teaching? What has worked and what has not?
- In your opinion, does the iPod touch make a difference in the students’ learning? Provide examples?
- In what way does iPod touch support your students’ learning? How does that compare to student learning without the iPods?
- Do you see any challenges in using the iPod? Explain.

All interviews were conducted face-to-face or through conferencing software, audio-recorded, and transcribed. Researchers also observed each teacher’s use of iPod touch in their classrooms. Field-notes were taken and details of the activities were written afterwards. The guidelines of Strauss and Corbin (1990) were followed in the data analysis: First, each interview was independently coded by one researcher to generate a list of initial codes, then another researcher reviewed and verified the codes. Necessary coding modification, realignment, and refinement were made during the process. This data analysis went through several iterative cycles until codes were categorized and themes emerged. Any disagreement was discussed until 100% inter-rater reliability was reached. Interview quotes were presented below as unedited. Observation notes were analyzed in a similar way and were used to supplement the interview data.

Three surveys were given to the students: at the beginning, middle, and end of the school year. The goal of the surveys was to seek students’ self-reported usage (as there was no tracking software available for iPods) and find out their perception toward having iPod touch as a learning tool. Sample survey questions are “How often do you use the following applications/features at school?” “How often do you use the following applications/features at home?” “How helpful or fun do you think the following applications or features are?” and demographic questions. Survey responses were analyzed descriptively and frequency data were tallied. Since not all students responded to all survey items, the results provided the percentages of those who responded to the questions. Throughout the data collection process, the research team met regularly to discuss what they observed in different classrooms, shared insights, and performed peer debriefing.

Findings

Year-one implementation

Virginia taught three classes per day to ELL students and two of these were dedicated to beginner ELL students while Claire taught four periods of ELL classes with three of these dedicated to beginners. In their second year of
iPod use in their classrooms, Virginia and Claire enthusiastically embraced the opportunity the iPod touch initiative presented and were the first group of teachers who implemented it in their classrooms.

Interviews with Virginia and Claire and classroom observations revealed both teachers had very positive perceptions toward using iPod touch and found ways to incorporate the device in their daily teaching activities as well as home assignments. Data analysis highlighted a few affordances of mobile technology in their case: Using iPod touch to (1) support language and content learning with Internet-based multimedia resources, (2) provide differentiated instructional support to accommodate students’ individual needs and create collaborative learning opportunities, and (3) extend learning time from classroom to home use and establish a better home-school connection.

**Support language and content learning**

The student iPod touches were preloaded and synced by the teachers with native apps along with an Internet browser as well as a variety of learning games, videos, audio books, multimedia, and textbooks. Virginia and Claire had students access a wide range of multimedia resources to support English language learning in classes as well as assigned homework.

During the classroom instruction, the students were able to access resources such as translation dictionaries that provided audio pronunciations along with images to support the vocabulary acquisition and audio textbooks. Claire described an English learning activity, “They go find a picture and the name of the picture, we use [the app] ComicLife for that, so they have a picture and vocabulary word and at the beginning they add a Spanish word to it.” Students used the devices for just-in-time translations as well as multimedia materials to support content learning in math, science, and social studies. For example, Virginia described, “I used a lot of the science videos and I try to keep it on the topic that the science teachers are teaching.” The iPod touch, as a small and portable computer with Internet access, was available to students 24/7 and provided them with a plethora of multimodal resources for English language and content learning.

**Provide differentiated instructional support**

In addressing various English language learning levels of these students, the teachers used varied apps and resources to allow students for access to appropriate levels of language and academic instruction. As an example, Claire spoke of how the iPod “allows more development, more customized learning” and said that with the iPod, “they have their private tutor” with non-judgmental support in their activities. Virginia also spoke of how the iPod use helped the students, “It’s private, it’s between them and, so they don’t feel stupid.”

The teachers assigned the activities and games appropriate for students’ specific language levels and were able to scaffold students more easily with iPod touch from basics in phonetics and sight words to more fluency and comprehension in advanced topic and subjects. Virginia described her differentiated instruction: “I give instructions: 7th grade, you go to this video, 8th grade you go to this video.” Collaborative learning was also afforded in allowing the students to share information gathered individually with other students in classroom discussion. Claire illustrated, “They feel free to say whatever they think and they don’t feel that pressure [to participate].” Besides supporting the students in shared learning activities, Virginia had her students access current events and collaborated in identifying key information and answered questions on the topic.

**Extend learning time from classroom to home**

The students were assigned homework activities using iPod touch for extended learning at home. Virginia elaborated on her literacy learning focus with iPods for “helping them read better and being oral speakers” with audio recordings. The ability to take the iPods home allowed the student further academic learning away from the classroom. Emily, a seventh grade student, summarized her iPod use for learning, “I practice more in the iPod.” With the multifarious learning games and apps, the teachers believed their students were more engaged in further practices than traditional homework materials and activities offered.
Claire considered the importance of allowing student access to the iPod away from school and said, “The iPod brings more a family into the issues of learning.” Some students’ parents were able to access the students’ grades at school as well as being able to monitor their children’s actual work on the iPod. Additionally, the teachers were able to assign the student homework that included the siblings and parents. For example, Claire had the students conduct a video interview of their parents talking of their educational experiences as children. Virginia included homework that had the students read to their siblings, “... take one of my baby books and read it to your little bothers and sisters.” Parents were able to access the Internet via iPod touch for their own use as well. Virginia identified an incident of finding “job application” and discovered that the student’s parent was seeking employment information via the device.

Middle school students’ usage and perception

The 6th, 7th and 8th graders were asked to report how often they used the available features on their iPod touch on a weekly average both at school and at home. The features available to the students can be grouped into three categories: Resource tools (e.g., calculator, calendar, accessing Internet, maps, listening to music and podcasts, and checking weather), media creation tools (voice recorder, notes, still and video cameras) and other applications (often in the form of games). Table 2 showed students’ responses for school and home use. Comparing mid-of-school usage to end-of-school usage, the data showed that a considerable percentage of students used resource and media creation tools for two or more hours per week during the mid-school year; however, at the end-school year, almost all of the students reported using these applications 0-1 hour per week. The results indicated that the highest use of the iPod touch occurred during mid-school year when teachers and students were engaged in using the resource and media creation tools as part of weekly learning activities. The noticeable decrease of usage pattern toward the end of the school year could be explained in part by the fact that much of the school time during the end-school year was spent on preparing for various state mandated tests and end-of-year curriculum exams and testing itself. Comparing school and home usages, the data indicated approximately 63% of the students spent 2 or more hours using the Internet at home (with 25.9% spent more than five hours at home while 16.7% spent more than five hours at school). This is also true for listening to music with 40.7% spent 2-5 hours at home while 11.5% at school and 25.9% spent more than 5 hours at home while 15.4% at school. That is, students spent more time using the Internet (both school or non-school related) and listening to music (non-school related) at home than at school.

We asked the students to name three features they used most often and found most helpful. Accessing Internet, listening to music, and using the calculator were the three features the students used most often both during mid-school year and at the end-school year for resource tools. Voice recorder, video camera, and notes were used most often for media creation tools. Students found Internet and calculator most helpful for the resources tools and voice recorder and notes for media creation tools.

<table>
<thead>
<tr>
<th>Resource tools</th>
<th>Duration in hours</th>
<th>Mid-year %</th>
<th>End-year %</th>
<th>Mid-year %</th>
<th>End-year %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculator</td>
<td>0-1</td>
<td>48.1</td>
<td>94.2</td>
<td>66.7</td>
<td>82.3</td>
</tr>
<tr>
<td></td>
<td>(n = 27)</td>
<td>(n = 17)</td>
<td></td>
<td>(n = 27)</td>
<td>(n = 17)</td>
</tr>
<tr>
<td></td>
<td>2-5</td>
<td>40.7</td>
<td>5.9</td>
<td>33.3</td>
<td>17.7</td>
</tr>
<tr>
<td></td>
<td>(n = 27)</td>
<td>(n = 17)</td>
<td></td>
<td>(n = 27)</td>
<td>(n = 17)</td>
</tr>
<tr>
<td></td>
<td>More than 5</td>
<td>11.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Calendar</td>
<td>0-1</td>
<td>77.8</td>
<td>100</td>
<td>85.2</td>
<td>93.8</td>
</tr>
<tr>
<td></td>
<td>(n = 27)</td>
<td>(n = 17)</td>
<td></td>
<td>(n = 27)</td>
<td>(n = 16)</td>
</tr>
<tr>
<td></td>
<td>2-5</td>
<td>11.1</td>
<td>0</td>
<td>7.4</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>(n = 27)</td>
<td>(n = 17)</td>
<td></td>
<td>(n = 27)</td>
<td>(n = 16)</td>
</tr>
<tr>
<td></td>
<td>More than 5</td>
<td>11.1</td>
<td>0</td>
<td>7.4</td>
<td>0</td>
</tr>
<tr>
<td>Internet</td>
<td>0-1</td>
<td>45.8</td>
<td>94.1</td>
<td>37.0</td>
<td>93.8</td>
</tr>
<tr>
<td></td>
<td>(n = 24)</td>
<td>(n = 17)</td>
<td></td>
<td>(n = 27)</td>
<td>(n = 16)</td>
</tr>
<tr>
<td></td>
<td>2-5</td>
<td>37.5</td>
<td>5.9</td>
<td>37.0</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>(n = 24)</td>
<td>(n = 17)</td>
<td></td>
<td>(n = 27)</td>
<td>(n = 16)</td>
</tr>
<tr>
<td></td>
<td>More than 5</td>
<td>16.7</td>
<td>0</td>
<td>25.9</td>
<td>0</td>
</tr>
<tr>
<td>Maps</td>
<td>0-1</td>
<td>66.7</td>
<td>100</td>
<td>64.0</td>
<td>94.1</td>
</tr>
<tr>
<td></td>
<td>(n = 24)</td>
<td>(n = 16)</td>
<td></td>
<td>(n = 27)</td>
<td>(n = 17)</td>
</tr>
</tbody>
</table>
### Challenges revealed

Several challenges were observed during this first cycle: Significant amount of time required of the teacher to learn to use iPod touch, finding appropriate apps, developing lessons that would integrate the mobile devices, managing the devices by charging, synching, downloading, and updating, and dealing with technical problems. These tasks presented challenges for the teachers who already had a demanding workload with teaching ELL students. In addition, the teachers had to attend necessary technology training required by the school district.

Another challenge was encountering technology issues such as loss of Wi-Fi capacity for use of multiple devices at once at school as well as teachers helping individual students with login problems. The devices were also potentially prone to breakage or loss as the students took them home and other settings away from the classroom. For example, during the school year, several devices were broken or lost. A few thefts occurred throughout the year and one theft of 15 iPod touches towards the end-school year with only one recovered. Other examples of damage included accidental mishandling and washing with clothes.

### Table 3: iPod touch uses by teachers (26
classes, 17 teachers)

<table>
<thead>
<tr>
<th>Category</th>
<th>0-1 (n=26)</th>
<th>2-5 (n=17)</th>
<th>More than 5 (n=17)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Music</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=26)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>73.1</td>
<td>76.5</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>11.5</td>
<td>23.5</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>15.4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Weather</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=25)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>56.0</td>
<td>87.5</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>36.0</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>8.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Podcasts</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=25)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>64.0</td>
<td>94.1</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>24.0</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>12.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Voice recorder</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=24)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>58.3</td>
<td>94.1</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>29.2</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>12.5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Notes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=26)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>57.7</td>
<td>94.1</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>26.9</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>15.4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Still camera</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=21)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>76.2</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>19.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>4.8</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Video camera</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>50.0</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>35.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>15.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Other applications</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>63.1</td>
<td>81.2</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>21.1</td>
<td>18.8</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>15.8</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Not all students responded to each question. N indicated the number of the students who completed the question. This is also the case for Table 3. Other applications mostly included games the teachers downloaded for the students.
Similar data collection procedure was followed for both cycles. However, as researchers, we realized asking middle school students to self-report their weekly iPod usage can be a challenge as this age group often had a hard time remembering an activity within a week’s duration and had trouble understanding hours vs. minutes. Therefore, in surveying elementary students for the second cycle, we asked the students to report daily usage and made the response time duration from hours to minutes.

**Year-two implementation**

Clara and Lydia were partner teachers for 4th and 5th grade ELLs and both held a bilingual teaching certification. Clara taught math and science and Lydia taught English Language Arts and Social Studies. As a team, they collaborated on group activities, student discipline, and iPod integration. Clara also served as the manager of the devices, updating and syncing the devices for all 42 students. Compared to Virginia and Claire, Clara and Lydia, in their initial year of iPod integration, were less enthusiastic about this mobile initiative and they implemented the project because they were encouraged by the school district to participate in the initiative. The analysis highlighted a few affordances of mobile technology in their case: Using iPod touches to (1) support language and content learning with Internet-based multimedia resources and (2) provide differentiated instructional support to increase student engagement and collaboration in the classroom.

**Support language and content learning**

Clara and Lydia shared insights on their experience, explaining their efforts in incorporating iPod touches into their teaching. Similar to the middle school teachers, their experiences showed they used the tool for engaging students in language as well as content learning. They described how the iPod touch provided ELL students access to multiple resources that facilitated their learning. Clara detailed the process for multimedia use for vocabulary development: “They go online and find a picture that is associated with the word and download it or save it to their iPod and transfer it to their *StoryKit*. Then, they write the definition that goes with it. Then, they go into recording themselves reading their definition.” As another example, Lydia spoke of how she used the iPod for language and reading skills development, “We are using them in terms of recording their fluency or recording to listen to their reading.” Clara also stated, “All their science homework is uploaded into their iTouches” for easy sharing and grading, not possible with traditional instruction.

**Provide differentiated instructional support**

Clara described the differentiated learning opportunities made possible by being able to access the resources at students’ own levels as “none of my kids have ever been on the same level.” The students were engaged in using such resources as dictionaries and playing educational games at a level appropriate to their learning and therefore reinforcing skills such as reading and multiplication. Clara described her typical use in this way, “There is no more downtime in my classroom because any down time we had, even when I'm passing out papers, their iTouches are always on their desk and they know that: ‘Okay, get it out and start working on your times tables’ or ‘get it out and work on divisibility.’” Moreover, the students enjoyed using the iPod touch for learning. Lydia spoke of a writing activity in which the students went on a field trip and created a story using *StoryKit* with images. “They loved it. They had enjoyed the field trip anyway and were able to go in to create their own story pretty much just with peer assistance.” Additionally, Internet access allowed students to pursue topics of their interest as Lydia described a student’s use, “He will look up his favorite baseball players, what their stats are, and he will have a new line of interests depending on what he what he’s reading in his novels sometimes.” The students found the writing activities compelling when using images from the field experience. Clara added, “It’s fun for them. It’s not paper and pencil.”

**Elementary students’ usage and perception**

The 4th and 5th graders were asked to report how often they used the available features on their iPod touch on a daily average both at school and at home. The features available to the students were also grouped into three categories as
with the middle school students: Resource tools, media creation tools, and other applications. Table 3 shows elementary students’ responses for school and home use.

The overall usage patterns indicated an increase in use with the resource tools at school but a decrease in home use; an increase in use for the media creation applications at school, except for the camera, while the usage remained the same or decreased at home. In addition, home use of the video camera appeared to increase at the end-school year. For calculator, calendar, music and reminders, more than 50% of the students indicated not using them both at school and at home. These elementary students typically spent 30 minutes or less in using applications/features daily for approximately 2 to 2.5 hours each week. While relatively few apps/features were used for over an hour daily, of particular interest was that audio book use exceeded an hour.

For resources tools, when comparing the mid- and end-of-year usage, it showed a decrease in calculators use both at school and at home. Conversely, the use of Internet, and maps increased at school and at home. Internet use at school increased by 32% at the 31-60 minute duration while Internet use at home remained the same at the mid-year and end-school. For the media creation tools, the data indicated the voice recorder, notes, and video camera showed a modest increase in usage at the 1-30 minute time frame at school, while at home the voice recorder use dropped and video camera use increased. These usage patterns corroborated with teachers’ interviews indicating the types of school activities the students were required to do. For the other applications category, the trend indicated that both school and home use increased at the 31-60 minute duration at school and home.

When the students were asked to indicate how helpful or fun the applications/features are, more than half of the students found all of the applications/features helpful with the exception of music. Reporting the percentages of the features that exceed 50% of student responses, all resource tools, except for music, were found helpful and all media creation tools were found helpful. The features with the highest percentage of the students finding it helpful was audio books of the resource tools (helpful with 92.3% at mid-year and 92.1% at end-year), and the voice recorder of the media creation tools (82.0% at mid-year and 81.6% at end-year). The high usage of audio books and voice recorder was consistent with the teachers’ description of assigned learning tasks with the iPod touch.

The responses of “how fun” features that exceed 50% of students’ responses showed: calculator was identified as Not fun (55% at end-year), Internet was indicated as Very fun (55% at mid-year and 63.1% at year-end) and other application category (predominantly as games) was indicated Very Fun (68.7% at mid-year). Responses of the other features were spread out and offered mixed results.

*Table 3. How often students used iPod applications/features on average each day*

<table>
<thead>
<tr>
<th>Resource tools</th>
<th>Duration in minutes</th>
<th>At school</th>
<th></th>
<th>At home</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mid-year %</td>
<td>End-year %</td>
<td></td>
<td>Mid-year %</td>
<td>End-year %</td>
</tr>
<tr>
<td>Calculator</td>
<td>0 (n = 39)</td>
<td>66.7</td>
<td>78.9</td>
<td>51.3 (n = 39)</td>
<td>76.3 (n = 38)</td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>30.7</td>
<td>21.1</td>
<td>48.7</td>
<td>23.7</td>
</tr>
<tr>
<td></td>
<td>31-60</td>
<td>2.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>61-90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Calendar</td>
<td>0 (n = 39)</td>
<td>71.2</td>
<td>52.6</td>
<td>69.2 (n = 39)</td>
<td>57.9 (n = 38)</td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>25.6</td>
<td>47.4</td>
<td>30.7</td>
<td>42.1</td>
</tr>
<tr>
<td></td>
<td>31-60</td>
<td>2.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>61-90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Internet</td>
<td>0 (n = 40)</td>
<td>22.5</td>
<td>10.5</td>
<td>48.7 (n = 39)</td>
<td>44.7 (n = 38)</td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>67.5</td>
<td>44.7</td>
<td>33.3</td>
<td>36.8</td>
</tr>
<tr>
<td></td>
<td>31-60</td>
<td>7.5</td>
<td>39.5</td>
<td>15.4</td>
<td>18.4</td>
</tr>
<tr>
<td></td>
<td>61-90</td>
<td>2.5</td>
<td>5.3</td>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td>Maps</td>
<td>0 (n = 38)</td>
<td>52.6</td>
<td>36.8</td>
<td>55.3 (n = 38)</td>
<td>26.3 (n = 38)</td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>44.7</td>
<td>60.5</td>
<td>36.8</td>
<td>65.8</td>
</tr>
<tr>
<td></td>
<td>31-60</td>
<td>2.6</td>
<td>23.7</td>
<td>5.2</td>
<td>7.9</td>
</tr>
<tr>
<td>Time</td>
<td>Music</td>
<td>Weather</td>
<td>Reminders</td>
<td>Clock</td>
<td>Audio book</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td>0-10</td>
<td>92.3</td>
<td>56.4</td>
<td>0</td>
<td>15.0</td>
<td>30.8</td>
</tr>
<tr>
<td>11-20</td>
<td>5.1</td>
<td>41.0</td>
<td>13.2</td>
<td>50</td>
<td>65.0</td>
</tr>
<tr>
<td>21-30</td>
<td>2.6</td>
<td>2.6</td>
<td>2.6</td>
<td>2.5</td>
<td>17.5</td>
</tr>
<tr>
<td>31-40</td>
<td>56.4</td>
<td>47.3</td>
<td>2.6</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td>41-50</td>
<td>0</td>
<td>42.1</td>
<td>0</td>
<td>42.1</td>
<td>0</td>
</tr>
<tr>
<td>51-60</td>
<td>0</td>
<td>42.1</td>
<td>0</td>
<td>42.1</td>
<td>0</td>
</tr>
<tr>
<td>61-90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>92.1</td>
<td>35.9</td>
<td>71.1</td>
<td>20.5</td>
<td>27.5</td>
</tr>
<tr>
<td></td>
<td>(n = 38)</td>
<td>(n = 38)</td>
<td>(n = 38)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
</tr>
<tr>
<td></td>
<td>5.3</td>
<td>60.0</td>
<td>21.0</td>
<td>5.3</td>
<td>44.7</td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td>5.1</td>
<td>2.6</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>24.3</td>
<td>37.5</td>
<td>27.5</td>
<td>42.1</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td>(n = 33)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
</tr>
<tr>
<td></td>
<td>52.6</td>
<td>64.8</td>
<td>54.0</td>
<td>47.3</td>
<td>50.0</td>
</tr>
<tr>
<td></td>
<td>23.7</td>
<td>12.1</td>
<td>53.8</td>
<td>5.2</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td>12.1</td>
<td>5.3</td>
<td>2.6</td>
<td>2.6</td>
</tr>
</tbody>
</table>
**Challenges revealed**

During the second year of implementation, the two elementary teachers communicated similar challenges as those by the two middle school teachers. For these two technology novices, they especially indicated significant amount of additional time required of them to carry out the project. In addition, these teachers pointed out they had to attend additional technology training during the holidays as well as in the classroom. Although the teachers indicated the benefits of the training, they also pointed out this was in addition to their regular workload. Apart from the technical issues, appropriate use of the devices was identified as an issue in the elementary setting. The teachers described individual students who misused the camera function as well as cases of attempts to bypass teacher-managed settings. Monitoring appropriate student use while supporting mobile device use in K-12 setting poses a specific challenge. Moreover, the teachers indicated a need for change in teaching approaches. While the purpose of this mobile initiative was to improve ELL students’ success on state assessments, the various affordances provided by mobile devices were found to encourage and support teachers to adopt more independent, learner-centered pedagogical approaches.

**Discussion and implications**

**Discussion**

The goal of this research is to investigate how iPod touch devices can serve as a teaching and learning tool for English Language Learners. Encompassing a two-year period with two implementation cycles, the findings revealed that iPod touch use can offer ELL students critical support for language and content learning (Banister, 2010; Craig et al., 2007; Kiger et al., 2012; Patten & Craig, 2007), as well as engaging students in enjoyable learning (Ullman, 2010). Three affordances of mobile technology are highlighted with the results: Using iPod touch to support language and content learning, provide differentiated instructional support, and extend learning time from classroom to home (Chan et al., 2006; Looi, Seow et al., 2010; Looi, Zhang et al., 2010). For these ELL students, several features are especially important: audio books, Internet access, and media creation tools. The iPod touch provides them with another learning platform in a just-in-time manner to facilitate their language acquisition and content learning. Teachers also have easier access to resources and tools to adapt and adjust their instruction to meet students’ needs (Lacina, 2008; Ullman, 2010). The noticeable changes in the teacher’ curriculum practices is to allow students to have more control over their learning. Because of the instant access to useful resources (e.g., translator, dictionary, voice recorder, and other applications) at anytime and anywhere, the students can use tools they previously did not have access to learn content materials. This is especially important for ELL students as they can practice language skills in different ways using multiple modes at school and at home. The teachers in this study noted a need to shift their pedagogy toward a more student-centered approach with affordances offered by mobile devices. Kukulski-Hulme and Traxler (2005) described the “discursive/didactic dichotomy” within m-learning as a complexity in teaching approaches in traditional environments (p. 27). Additionally, consistent with previous research suggesting that engaging games can provide enjoyable content learning (Huizenga et al., 2009), the results showed that the students often accessed the game apps both at school and home and considered such apps as fun. Educators should take advantage of the new opportunities and possibilities provided by mobile devices to “engage, motivate, support, and interest students” (Liao et al., 2011, p. 86) to positively influence learning.

Although there are clear benefits for using iPod touch in teaching and learning as this research indicated, the results also showed teachers will need significant added support for learning how to use and manage the devices. This is especially so for teachers who themselves are not enthusiastic technology users or technology novices. Using mobile technology requires teachers to allocate significant time and effort to learn to use the device as well as learn to integrate effectively in their teaching (Kukulski-Hulme & Traxler, 2005; Traxler, 2011). The management of the devices for classroom wide use, in terms of charging, synching and dealing with multiple technical demands can also add a substantial load on the teachers (Franklin & Peng, 2008; Vogel, Kennedy, & Kwok, 2009).

**Implications**

Given the preliminary encouraging evidences of m-learning for ELL students and teachers, what we have learned from this research project for the past two years not only has implications for this district, but also should offer insights for teaching and learning practices for other schools that are considering similar initiatives.
**For teachers**

Teachers who plan to incorporate mobile devices should be prepared to invest considerable amount of time in learning the functionalities of the devices as well as in designing, planning, and developing effective learning activities. Preparing to use the device for instruction will require prior preparation during the summer, before the school year commences. Moreover, teachers should be flexible in their teaching practice, reconsidering curriculum as well as teaching and learning approaches in order to effectively leverage the full range of affordances that mobile devices can offer. ELL teachers should take advantage of the new possibilities to offer differentiated instruction in meeting the specific language and academic needs of ELL students. To accommodate students’ learning needs and encourage student-directed learning, teachers can consider management plans to monitor student learning and create activities that utilize a wide range of apps/features (e.g., audio recordings, video productions, digital stories) available on mobile devices.

**For instructional technologies**

School district’s instructional technology department played a key role in this implementation. Prior to and during the implementation, the instructional technologists should provide training for the teachers, focusing on the basic capabilities and operations as well as identifying appropriate apps and suggesting ideas on how to integrate the iPod touch into the overall curriculum. Additionally, instructional technologists must anticipate, identify and quickly resolve technical challenges as well as help teachers in making sure that devices are synced with the appropriate apps, media, software, and security settings. They need to make sure students and teachers know how to log in to their devices and access wireless Internet, while coordinating with the network infrastructure staff in ensuring for sufficient and reliable network resources availability. Technology staff also needs to remain available for troubleshooting technical problems so that these types of issues do not lead to frustration and disuse.

**For school administrators and policy makers**

The results of this research suggest that realizing the benefits of the iPod touch program requires initial and ongoing support from the school administration. Support should include dedicated instructional technology staff to provide training and on-going support as well as provide funding to purchase and maintain the iPod touches. It is suggested that policies and processes for fixing or replacing broken, lost, and stolen devices should be in place. Administrators should also consider how to get buy-in from teachers when rolling out a technology initiative such as this one and set and enforce guidelines for safe and appropriate use of the devices and their apps. Equipping entire classes with mobile devices can create a significant demand for wireless bandwidth and administrators need to make sure the infrastructure can meet the demand. Without the support to ensure a fully functional device most of the time, student and teacher frustration may lead to declining use and thus undermine or negate the affordances of these devices and their potential to enhance learning.

**For researchers**

Conducting classroom-based, mobile learning research is a challenge, especially in ELL classrooms. Given the nature of ELL instruction in US public schools, there are many classroom transitions in middle schools on any given day. ELL students can be in different content classes and capturing classroom iPod interactions consistently poses a challenge. Because there was no tracking software on iPods, usage has to rely on students’ self-reporting. Students’ responses often are short and less specific. Follow-up interviews could be considered to further examine students’ perception.

**Conclusion**

The experience of the two-year implementation of this mobile initiative and research findings suggest that mobile devices such as iPod touch can provide ELL students significant support for language and content learning and extend learning time from classroom to home. Audio books, Internet access, and media creation tools are found to be
especially important for these ELL students. While several affordances are revealed both at elementary and middle school ELL settings, substantial support in the form of professional training and administration support and encouragement is also highlighted.

References


The Impact of a Principle-based Pedagogical Design on Inquiry-based Learning in a Seamless Learning Environment in Hong Kong

Siu Cheung Kong and Yanjie Song

Department of Mathematics and Information Technology, The Hong Kong Institute of Education, Hong Kong // sckong@ied.edu.hk // ysong@ied.edu.hk

ABSTRACT

An inquiry-based learning pedagogy coupled with a seamless learning environment is a potential way to realise the educational goal of learner-centred learning in digital classrooms in the 21st century. An overarching research framework is proposed for preparing teachers to effectively develop pedagogical designs that are premised on theoretical principles and facilitate inquiry-based learning in a seamless learning environment. We carried out an initial study using the overarching framework. Three questions are addressed: how a principle-based pedagogical design was developed and implemented, the effect that the principle-based pedagogical design had on students’ domain knowledge gains and inquiry skills and how students advanced their domain knowledge and developed their inquiry skills. One teacher and 27 students from a local primary school were involved in the study. Both qualitative and quantitative data were collected and analysed over two weeks. Six inquiry-based learning lessons focusing on a scientific ‘rustproofing’ learning unit were conducted in a seamless learning environment, initiated in a digital classroom and extended to online discussions on a social network platform. The results reveal innovative ways of developing and implementing the pedagogical design in the rustproofing learning unit and demonstrate the pedagogical design’s positive effect on students’ domain knowledge gains and inquiry skills. In addition, how the students advanced their domain knowledge and inquiry skills were also explored and discussed.

Keywords

Inquiry-based learning, Primary school education, Principle-based pedagogical design, Seamless learning environment

Introduction

Educational reform calls for a paradigm shift to learner-centred domain knowledge learning. It is well recognised that the inquiry-based learning approach is a useful pedagogy for realising learner-centred learning (Marshall, Smart, & Horton, 2010). The inquiry-based learning process helps learners to develop inquiry skills, which are an important type of 21st century skill. The development of inquiry skills takes root during a child’s senior primary school years (Lakkala, Lallimo, & Hakkarainen, 2005). Digital classrooms are on the rise; students are connected and learn in a ‘one learner to one computer’ setting, and teachers are expected to be prepared to lead students to learner-centred learning in such classrooms as early as the primary school stage. The use of online learning platforms inside and outside digital classrooms supports resource access and peer interaction to develop students’ domain knowledge and inquiry skills (Kong & So, 2008; Lakkala et al., 2005). Incorporating the inquiry-based learning pedagogy into a seamless learning environment may thus be a potential method for realising learner-centred educational goals and driving teachers to apply and reflect on pedagogical designs. This study presents a design-based research framework for principle-based pedagogical design for inquiry-based learning in seamless learning environments. It details and reports the results of an initial study conducted using this framework in a Hong Kong primary school.

Research framework

With the goal of finding a meaningful and sustainable method of developing the teacher competence necessary to facilitate inquiry-based learning in a seamless learning environment, this research study seeks to address two issues. First, the method for developing teacher competence should be in line with the method for developing learner competence in inquiry-based learning. Second, the method for developing teacher competence should present evidence of the development of teacher competence and student learning improvement. This study adopts a principle-based approach to developing and implementing pedagogical designs for inquiry-based learning. This approach, which differs from the conventional approach that emphasises “best practices” with prescribed procedures, provides more flexible scaffolding under guiding principles. It attempts to build up teachers’ capacity to promote inquiry-based learning in a manner aligned with learners’ inquiry-based learning practice. This study also adopts a
design-based research approach to developing and refining pedagogical designs for inquiry-based learning guided by instructional principles. It exposes teachers to the process of progressive refinement in pedagogical designs driven by principles and supported by empirical evidence. In light of these issues, we propose an overarching research framework on the use of principle-based pedagogical designs for inquiry-based learning in a seamless learning environment, as shown in Figure 1.

A design-based research method is adopted to prepare teachers to effectively develop pedagogical designs that are premised on instructional principles. Learners learn in a seamless learning environment to develop the necessary skills to practice inquiry-based learning.

Inquiry-based learning includes three approaches: structured, guided and open inquiry, listed in ascending order of the learner’s autonomy over setting investigation problems and planning problem-solving procedures (Colburn, 2000). The literature suggests that the guided inquiry approach is especially suitable for young learners, as teachers only select core issues that are worthy of a learner’s inquiry (Hakkarainen, 2003; Marshall et al., 2010; Song & Looi, 2012). According to Wong and Looi (2011), seamless learning environments provide learners with opportunities to make use of diverse resources and tools in digital formats for learning and communication, which is initiated in digital classrooms and extended to online interactions. The technological support of a seamless learning environment allows learners to conveniently share and store multimedia resources, and to easily exchange and track discussion ideas with peers during the inquiry process. During class time, learners in digital classrooms are connected, and use digital technologies in a ‘one learner to one computer’ setting (Chan, 2010; Kong, 2011). Beyond the limited class time, learners typically use learning platforms to communicate with peers online, mostly to extend discussions or to engage in deeper discussions after class.

Pedagogical design refers to the organisation plan for learning activities and the actual implementation of the plan in a learning unit (Lakkala et al., 2005). Researchers have reported that principle-based pedagogical designs are more adaptable and conducive to transforming inquiry-based learning practices (Schwarz, 2009; Song & Looi, 2012; Zhang, Hong, Morley, Scardamalia, & Teo, 2011). The principle-based approach to pedagogical design defines the core principles of learning and teaching. According to Schwarz (2009), Zhang (2010) and Zhang et al. (2011), principle-based pedagogical designs focus on guiding principles and customizable practices. Teachers are afforded the flexibility to reflectively judge and adapt classroom decisions to accommodate different learning and teaching possibilities. Based on knowledge-building and social-constructivism theories (e.g., Scardamalia, 2002; Vygotsky, 1978), a set of theoretical principles premised on 12 knowledge-building principles and progressive inquiry principles (Lakkala, Muukkonen, Paavola, & Hakkarainen, 2008; Scardamalia, 2002; Song & Looi, 2012; Zhang et al., 2011) is considered suitable for pedagogical designs for inquiry-based learning.
Teachers need support from evidence-based research to make continuous pedagogical reflections. As such, the design-based research approach is suitable for gaining new insights. Design-based research attempts to combine theory-driven design with empirical analyses of practices in real settings. It creates a path to connect interventions to outcomes through an iterative mechanism of design, evaluation and refinement (Bell, 2004; Hoadley, 2004). Teachers are provided with iterative opportunities (as shown in Figure 1) to use principle-based pedagogical designs to enhance their competence in leading inquiry-based learning in a seamless learning environment. This study details and reports the results of an initial study on design-based research that explored the effect of using the principle-based approach to pedagogical designs for science inquiry in the seamless learning environment of a Hong Kong primary school.

The initial study was conducted using the overarching framework in a learning unit on ‘rustproofing’ conducted in the school’s Primary 4 class.

This study

Research context

The study took place in the initial cycle of design-based research on principle-based pedagogical designs for inquiry-based learning that aimed to develop teacher competence and learners’ science domain knowledge and inquiry skills in a seamless learning environment at the primary level in Hong Kong. According to recent territory-wide surveys on the development of technology-enhanced education in Hong Kong (Li & Kong, 2011), local primary school teachers are typically capable of integrating technology into their daily teaching methods. Further, local primary school learners are ready to use technology for inquiry-based learning, as they demonstrate a basic information literacy competency that is important in the inquiry process. These surveys reveal that primary schools in Hong Kong have built a foundation for introducing educational innovations that integrate pedagogical designs for inquiry-based learning into technology-supported learning environments.

The study purposefully sampled a primary school in Hong Kong as its partner school. One experienced science teacher and one Primary 4 class with 27 students (15 female and 12 male) were invited from the partner school to participate. The science inquiry focused on a learning unit on rustproofing conducted in six lessons over 2 weeks for senior primary school learners. The following research questions were addressed.

1. How did the teacher develop and enact the principle-based pedagogical design?
2. What effect did the principle-based pedagogical design have in helping students to gain domain knowledge and inquiry skills in the seamless learning environment?
3. How did the students advance their domain knowledge and develop their inquiry skills?

To address these questions, our pedagogical design involved the adoption of the 5E inquiry-based learning model to guide the students’ science inquiry, and five instructional principles for pedagogical practice in a seamless learning environment supported by a social network (i.e., Edmodo). These principles are elaborated in the remainder of this section.

5E inquiry-based learning model

According to EDB (2008), the focus of science education is to promote students’ scientific thinking through inquiry-based learning approaches. Although open inquiry provides optimal opportunities for students’ cognitive development and scientific reasoning, teacher-guided inquiry may provide better opportunities for students to focus on the development of particular science concepts (Song & Looi, 2012). To balance the two inquiry approaches, we developed a 5E inquiry-based pedagogical model as follows: (a) “engage” in inquiry topics and questions, (b) ‘explore’ the inquiry methods and processes, (c) “explain” the inquiry analyses and outcomes, (d) “evaluate” the inquiry processes and outcomes and (e) “extend” the inquiry topics and questions. The process is cyclic and progressive but not linear, and may not involve all of the components in each learning cycle.
Five instructional principles

To explicate the processes and dynamics of science inquiry for knowledge advancement using the inquiry-based pedagogical approach, we adapted five core instructional principles from a set of progressive inquiry principles (Song & Looi, 2012) and other related research (Scardamalia, 2002; Yeo & Tan, 2010). These principles are premised on social constructivist principles, and include (a) working on real problems, (b) encouraging diverse ideas, (c) providing collaborative opportunities, (d) using authoritative sources constructively and (e) performing a formative assessment (see Table 1).

<table>
<thead>
<tr>
<th>Principles</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working on real problems</td>
<td>Setting up real-life problems rather than abstract concepts (Scardamalia, 2002; Song &amp; Looi, 2012).</td>
</tr>
<tr>
<td>Encouraging diverse ideas</td>
<td>Encouraging students to express their ideas voluntarily. There is no right or wrong answer. Every idea is valued and unique (Song &amp; Looi, 2012).</td>
</tr>
<tr>
<td>Providing collaborative opportunities</td>
<td>Emphasis on the importance of collective effort and responsibility in the learning process (Scardamalia, 2002; Song &amp; Looi, 2012).</td>
</tr>
<tr>
<td>Using authoritative sources constructively</td>
<td>The meaningful use of authoritative sources for continual science meaning-making (Yeo &amp; Tan, 2010).</td>
</tr>
<tr>
<td>Performing a formative assessment</td>
<td>Provision of peer assessment and teacher feedback concurrently in the collaborative process (Scardamalia, 2002; Song &amp; Looi, 2012).</td>
</tr>
</tbody>
</table>

Teacher principle-based understanding

Before the beginning of the study, the teacher received two 1.5-hour training sessions from two researchers. In the first session, the researchers prompted and discussed the inquiry-based learning model and five instructional principles with the teacher using PowerPoint slides. The researchers then asked the teacher to reflect and pose questions. In the second session, the teacher chose one of his lessons to elaborate how he would conduct it using instructional principles, and discussed his pedagogical design with the researchers.

Seamless learning environment supported by a social network—Edmodo

This study took place in a seamless learning environment. The teacher provided pedagogical support in implementing the 5E model and five instructional principles for inquiry-based learning. The seamless learning environment comprised digital technologies that allowed students to access learning resources and interact with peers in inquiry-based learning. Inside the digital classroom, each student was given a mobile computing device comprising a tablet PC with Internet connectivity and an embedded camera. The social network Edmodo (see Figure 2), was used as a learning communication platform to support the students’ learning at the individual, group and whole-class levels both inside and outside the classroom. Edmodo is a secure microblogging medium conducive to collaborative knowledge construction (Ma, Ko, Chu, & Song, 2012). It can be used across formal and informal learning settings, allowing students to collaborate, communicate, submit assignments and upload and download files, and teachers to share lecture notes with students, connect to useful websites, upload and download learning references for students, create online quizzes and release news and events. The platform can run on different operating systems (e.g., iOS or Android).

Figure 2. Interface of Edmodo for the Primary 4 science class
The learning activities organisation plan

The principle-based pedagogical design included an organisation plan of the learning activities (see Table 2) and the actual enactment of the learning unit plan on rustproofing, which comprised six lessons and activities that Primary 4 students carried out at home or between lessons over two weeks. According to the plan, the students formed six groups of four or five members each and were expected to collaboratively lead their own experimental inquiries into an “expert rustproofing design” project. Two prompts were provided on Edmodo during the experimentation process. First, the ‘Forms for Experimental Rustproofing Designs’ prompt asked students to record their inquiry plans. The students were required to fill out three design methods with hypotheses (Appendix I), and each student was required to take on a responsibility in the design. Second, the ‘Observational Forms for Rustproofing Experiments’ prompt helped students to monitor their experimental process and scaffold their reflections. The students were required to document the rustproofing process over a week and to include the observers’ names (Appendix II). Both of the Edmodo forms were linked to GoogleDocs, which allowed the students to fill them in directly.

Table 2. Learning activities organisation plan for the rustproofing learning unit

<table>
<thead>
<tr>
<th>No.</th>
<th>Aim</th>
<th>Activity</th>
<th>Means of interaction</th>
<th>Teaching and learning resources</th>
</tr>
</thead>
</table>
| Lesson 1 | To engage students on the topic of rustproofing                      | - Storytelling was used to make students understand why they needed to rustproof and to arouse their curiosity on how to do so.  
- Individual students were required to discover rustproofing methods and prepare to share their discoveries with group members in the next lesson. | F2F + online          | LCD projector, Tablet PCs, Internet, Social network: Edmodo |
| At home | To discover rustproofing methods individually                       | - Individual students continued researching rustproofing methods and uploaded their findings to Edmodo to share with their peers.  
- The students posted questions to Edmodo and commented on or responded to other students’ posts. | Online learning       | Desktop, laptop, iPad, iPhone, etc. Edmodo |
| Lessons 2 and 3 | To determine the three best experimental rustproofing designs in groups | - Each group member shared and explained his or her findings on the experimental rustproofing designs.  
- The three best experimental rustproofing designs were discussed and worked out by combining every member’s ideas.  
- The three best experimental designs were filled in on the ‘Forms for Experimental Rustproofing Designs’, prepared by the teacher on Edmodo to share with peers. | F2F + online          | LCD projector, Tablet PC, Edmodo |
| At home | To plan and prepare the material for the rustproofing experiment in the next session | - Students could post their questions to Edmodo. They could also comment on or respond to other students’ posts. | Online               | Desktop, laptop, iPad, iPhone, etc. Edmodo |
| Lessons 4 and 5 | To conduct the experiment based on the proposed experimental designs in groups | - Each group conducted three rustproofing experiments by placing three iron clips into three plastic cups full of water and certain other materials.  
- The cups containing the clips and materials were placed on windowsills. | F2F + online          | Three iron clips provided by the teacher for the students to conduct the experiments. Students brought the rustproofing materials from home, including the plastic cups. |
Data collection and analysis

To understand the effects of the principle-based pedagogical design on students’ domain knowledge of rustproofing and inquiry-based learning skills, and on how the students advanced their domain knowledge and inquiry strategies, we collected the following data.

- Data on the development and implementation of the pedagogical design, including posts on Edmodo, the learning activities organisation plan, lesson videos, group experimental design forms, group experimental observational forms, the assignment and teacher interviews and reflections.

- Data on the effects of principle-based pedagogical design on domain knowledge gains, including pre- and post-domain tests and assignments. The pre- and post-domain tests were identical and consisted of 20 multiple-choice questions and five open-ended questions on rustproofing. The assignments took the form of worksheets on rustproofing knowledge (Appendix III). The students submitted their assignment directly to Edmodo, and the teacher commented on their work directly to provide immediate feedback. The data on the effects of the principle-based pedagogical design on student inquiry skills included pre- and post-questionnaire surveys on perception changes towards inquiry learning skills before and after the inquiry-based learning approach. The questionnaire focused on students’ perceptions of inquiry learning skills, and comprised 12 items rated on a five-point Likert scale (with 5 indicating strong agreement and 1 indicating strong disagreement). The 12 items were designed to address the five inquiry skills under the 5E inquiry-based learning model, with items 1-3 addressing the “questioning” skill; items 4-5 addressing the “exploring” skill; items 8-11 addressing the “explaining” skill; items 6-7 addressing the “evaluating” skill and item 12 addressing the “extending” skill. The Cronbach’s alpha reliability scores were 0.849 and 0.902 for the pre-test and post-test, respectively, implying that the questionnaire was reliable.

- Data on how the students advanced their domain knowledge, including group experimental designs (experimental design forms on Edmodo), group experimental results/products (observational forms on Edmodo) and group artefacts (photos documenting the experimental process), and data on how the students advanced their inquiry skills, including posts on Edmodo, lesson videos, teacher and student interviews and field notes.

Table 3 shows the data sources used to investigate the three research questions.

<table>
<thead>
<tr>
<th>Data</th>
<th>Research questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre- and post-domain tests</td>
<td>*Q1</td>
</tr>
<tr>
<td>Pre- and post-questionnaire surveys</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Data sources for analysis
Posts on the social network (Edmodo) | x | x
Learning activities organisation plan | x |
Lesson videos | x | x
Group artefacts |
Group experimental design forms | x | x
Group experimental observational forms | x | x
Group experimental field notes |
Assignments (worksheets) | x | x
Teacher interviews and reflections | x | x
Student interviews |

Note. Q1 development and implementation of the pedagogical design; Q2(a) effect on domain knowledge; Q2(b) effect on inquiry skills; Q3(a) methods of gaining domain knowledge; Q4(b) methods of developing inquiry skills.

To investigate how the teacher developed and implemented the pedagogical design, we adopted three iterative and complementary streams of content analysis: (a) a “preliminary exploratory analysis” to obtain an understanding of the data (Creswell, 2008, p. 250); (b) categorising strategies to code the inquiry skills using the five elements of the inquiry learning model (i.e., engage, explore, explain, evaluate and extend) and instructional strategies as a coding scheme and (c) contextualising strategies to identify the inquiry strategies and instructional principles implemented in the learning unit (Maxwell, 2005).

The effect of the principle-based pedagogical design on the students’ gains in domain knowledge was examined by analysing the pre- and post-domain tests using quantitative methods and SPSS software to determine the changes in students’ domain knowledge before and after the adoption of principle-based pedagogical design. The students’ assignment results were then analysed to identify content knowledge gains and problems. To investigate the effect on the students’ inquiry skills, pre- and post-questionnaire surveys were conducted to understand changes in the students’ perceptions before and after the adoption of the principle-based pedagogical design.

To scrutinise how the students’ science domain knowledge was advanced and to understand how the groups worked together to make sense of a problem inquiry situation (Stahl, 2002), we traced the students’ development of certain artefacts (Hakkarainen & Paavola, 2009), including concrete objects (e.g., experiment materials) and conceptual artefacts (e.g., text, pictures and drawings). The data analysis methods used to examine how the students advanced their inquiry skills were similar to the three iterative and complementary content analysis streams adopted to analyse the development and implementation of the pedagogical design. We also counted the numbers of students’ posts on Edmodo using a coding scheme, as a means of “counting” is necessary in qualitative data analysis in some circumstances (Miles & Humberman, 1994, p. 253). Whenever necessary, field notes were used in the data analysis process for clarification and confirmation.

**Results**

**Development and implementation of the principle-based pedagogical design**

The data analysis results on the development of the principle-based design show that the learning activities organisation plan included the five elements of the inquiry-based learning model in a seamless environment. The inquiry learning activities went from engaging students on the topic of rustproofing, to exploring rustproofing methods and making hypotheses for the experimental designs, to evaluating the experimental rustproofing designs through active experimentation and explaining and sharing the designs and finally to consolidating and extending the rustproofing knowledge to help the students to become rustproofing ‘experts’. The entire inquiry process was carried out seamlessly between classes and the students’ homes with the support of the social network platform (Edmodo).

The data analysis results on the implementation of the principle-based pedagogical design indicate that the teacher premised the rustproofing learning activities organisation plan on the five instructional principles (see Table 1). The demonstration of the five inquiry-based elements and five instructional principles in a seamless learning environment during the enactment of the rustproofing learning unit is illustrated in Table 4.
Table 4. Principle-based pedagogical implementation of inquiry-based learning in a seamless learning environment

<table>
<thead>
<tr>
<th>Implementation of organisation plan</th>
<th>Inquiry elements</th>
<th>Instructional principles</th>
<th>Seamless learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesson 1</td>
<td>Engage: Engaging in the topic of rustproofing</td>
<td>P1</td>
<td>Class</td>
</tr>
<tr>
<td>At home</td>
<td>Explore: Exploring rustproofing methods</td>
<td>P4</td>
<td>Home</td>
</tr>
<tr>
<td>Lessons 2 and 3</td>
<td>Explain: Sharing and planning rustproofing experiments and making hypotheses in groups</td>
<td>P2 and P3</td>
<td>Class</td>
</tr>
<tr>
<td>At home</td>
<td>Engage: Preparing and coordinating the rustproofing experiments</td>
<td>P1 and P3</td>
<td>Home</td>
</tr>
<tr>
<td>Lessons 4 and 5</td>
<td>Evaluate: Evaluating the hypotheses through conducting experiments</td>
<td>P1, P3 and P5</td>
<td>Class</td>
</tr>
<tr>
<td>Breaks between lessons</td>
<td>Evaluate: Observing and documenting the rustproofing process</td>
<td>P1, P3 and P5</td>
<td>Breaks</td>
</tr>
<tr>
<td>Lesson 6</td>
<td>Explain: Explaining and sharing group work</td>
<td>P2, P3 and P5</td>
<td>Class</td>
</tr>
<tr>
<td>At home</td>
<td>Extend: Consolidating and extending knowledge related to rustproofing</td>
<td>P5</td>
<td>Home</td>
</tr>
</tbody>
</table>

Note: *P1 = working on real problems; P2 = encouraging diverse ideas; P3 = providing collaborative opportunities; P4 = using authoritative sources constructively; P5 = performing a formative assessment.

Table 4 shows that the teacher flexibly adopted different principles at different stages of the science inquiry. We also asked the teacher to reflect on his pedagogical plan and enactment process and outcomes based on guided questions on the inquiry-based learning approach and instructional principles, and arranged a time to interview him to hear his reflections. Some of the questions (Q) and excerpts from his reflections (R) are presented as follows:

Q1: When teaching the rustproofing learning unit, do you think it is important for the students to conduct hands-on experiments in an authentic environment? Why?

R1: If I told the students the reasons for rusting and how to prevent rusting directly, it would take 3 minutes. However, by providing opportunities for the students to lead their own science inquiry, they not only tested their own hypotheses in the experiments, but also underwent a process of discovery and collaboration: to identify problems in their everyday lives, raise questions to explore resources and conduct experiments to solve the problems.

Q2: Did you encourage the students to express their diverse ideas in their inquiry? How?

R2: I valued each student’s questions. How? Digital technology and the Internet extend our learning spaces. I seldom asked students to ask questions face to face in class. They could post their questions anytime, anywhere to the Edmodo social network platform, both inside and outside the classroom. They can get quick feedback from their peers or from me. If I found that many students were concerned about a problem, I would discuss the problem in class. Using Edmodo, all students’ questions are treated equally and their learning is extended beyond the classroom. This can be called seamless learning.

It is worth noting that the teacher’s good understanding of inquiry-based pedagogies and the principles of working on real problems and encouraging diverse ideas allowed him to apply the instructional principles in his pedagogical practices in multiple contexts, with the support of the social network platform.

Effect on students’ domain knowledge gains and inquiry skills development

We investigated the effect of principle-based pedagogical design on students’ domain knowledge gains through pre- and post-domain tests on rustproofing. Table 5 shows the results of the tests. Significant differences were found between the pre- and post-domain test results (pre-average score = 11.64; post-average score = 22.50, \( p < 0.05 \)). We can thus conclude that the students made significant advancements in their rustproofing knowledge after their inquiry.

Table 5. Pre- and post-domain test results on the rustproofing learning unit

<table>
<thead>
<tr>
<th>Class</th>
<th>No. of students</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4C</td>
<td>27</td>
<td>M=11.64</td>
<td>S.D.=2.652</td>
<td>M=22.50</td>
</tr>
</tbody>
</table>

Note: Total test score = 40

* \( p < .05 \)
In terms of the rustproofing learning unit assignments, the average scores for the 27 students were 79% (7.11 out of 9 questions in total). Although some of the students’ scores were not high, their worksheets revealed that in many cases marks were deducted due to incorrect rendering of Chinese characters rather than their content knowledge. Figures 3(a) and 3(b) show screen captures of two students’ worksheets marked by the teacher. Figure 3(a) shows that the student used the incorrect Chinese word “份 (part)” rather than “分 (component)” – the two words are the same in Pinyin, and their characters are also similar. Figure 3(b) shows that one student did not know how to write the Chinese word “熄 (extinguish).”

![Figure 3. Screenshots of a student’s worksheet (a – left; b – right)](image)

The effect of the principle-based pedagogical design on students’ inquiry skills was examined through pre- and post-questionnaire surveys. The results are shown in Table 6. They reveal that only the pre- and post-questionnaire results for item 1 (I know how to start thinking about how to solve a scientific problem) (mean = 4.35) and item 10 (I know how to explain my ideas to my peers when learning science) (mean = 4.16) showed significant differences. Items 1 and 2 relate to the “questioning” and “explaining” inquiry skills, respectively, and indicated an improvement in the students’ skills in raising questions and explaining ideas and concepts to peers.

<table>
<thead>
<tr>
<th>Question</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I know how to start thinking about how to solve a scientific problem</td>
<td>3.77</td>
<td>4.35</td>
<td>4.573*</td>
</tr>
<tr>
<td>2. I know how to solve a scientific problem step by step</td>
<td>3.88</td>
<td>4.31</td>
<td>2.026</td>
</tr>
<tr>
<td>3. I know how to find scientific problems that I am interested in solving</td>
<td>4.08</td>
<td>4.15</td>
<td>0.440</td>
</tr>
<tr>
<td>4. I know where to find the information to solve a scientific problem</td>
<td>4.00</td>
<td>4.27</td>
<td>1.272</td>
</tr>
<tr>
<td>5. I know how to explore information/resources on my own when solving a scientific problem</td>
<td>3.96</td>
<td>4.08</td>
<td>0.486</td>
</tr>
<tr>
<td>6. I know how to improve the ways to solve a scientific problem</td>
<td>3.88</td>
<td>4.27</td>
<td>1.917</td>
</tr>
<tr>
<td>7. I know how to try different ways of solving a scientific problem</td>
<td>4.13</td>
<td>4.13</td>
<td>0.000</td>
</tr>
<tr>
<td>8. I know when to ask help from my peers when learning science</td>
<td>4.04</td>
<td>4.08</td>
<td>0.132</td>
</tr>
<tr>
<td>9. I know when to ask help from teachers when learning science</td>
<td>3.88</td>
<td>3.88</td>
<td>0.000</td>
</tr>
<tr>
<td>10. I know how to explain my ideas to my peers when learning science</td>
<td>3.72</td>
<td>4.16</td>
<td>2.290*</td>
</tr>
<tr>
<td>11. I know how to explain my ideas to my teacher when learning science</td>
<td>4.24</td>
<td>4.20</td>
<td>-0.146</td>
</tr>
<tr>
<td>12. I know how to work together with my peers to solve a scientific problem</td>
<td>3.88</td>
<td>4.29</td>
<td>2.005</td>
</tr>
<tr>
<td>Total</td>
<td>3.87</td>
<td>4.15</td>
<td>1.973</td>
</tr>
</tbody>
</table>

*p < .05

Ways that the students advanced their domain knowledge and developed their inquiry skills

We traced the artefact development of each student group in the rustproofing learning unit. Each group designed three methods. Among the 18 experimental methods proposed by the groups, only one method (i.e., using paint to coat the clip) was the same between two groups. All of the other methods were different from each other (see Table 7). In addition, after tracking the experiment outcomes, the degree of rusting in each method was evaluated and scored from 0 (no rusting) to 10 (most severe rusting), which is also shown in Table 7.
Table 7. Experimental designs and rustproofing outcomes by group

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of design</th>
<th>Hypothesis of rustproofing theories using the methods. (All of the following methods were hypothesised to prevent air from entering the cup or to prevent water from having direct contact with the clip.)</th>
<th>Degree of rusting</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>(1)</td>
<td>Use Vaseline and paint to coat the clip and put oil into the cup along with the water.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use Vaseline to coat the clip.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Use paint to coat the clip.</td>
<td>7</td>
</tr>
<tr>
<td>G2</td>
<td>(1)</td>
<td>Use Vaseline to coat the clip and then wrap the clip with plastic wrap. Put oil and W-4 into the cup.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use Vaseline to coat the clip and then wrap the clip with plastic wrap. Put oil into the cup.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Use Vaseline to coat the clip and then wrap the clip with plastic wrap.</td>
<td>3</td>
</tr>
<tr>
<td>G3</td>
<td>(1)</td>
<td>Put the clip into a storage bag.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use Vaseline to coat the clip, and then use the dryer to dry the clip before putting it into the water.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Use a candle to dry the clip after removing it from the water.</td>
<td>3</td>
</tr>
<tr>
<td>G4</td>
<td>(1)</td>
<td>Wrap the clip with glue paper.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Wrap the clip with plastic wrap.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Put oil into the cup.</td>
<td>2</td>
</tr>
<tr>
<td>G5</td>
<td>(1)</td>
<td>Use Vaseline to coat the clip and then put oil and soya oil into the cup.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use Vaseline to coat the clip and then put oil into the cup.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Wrap the clip with glue paper and then put the oil into the cup.</td>
<td>7</td>
</tr>
<tr>
<td>G6</td>
<td>(1)</td>
<td>Use paint to coat the clip, seal the clip inside a bottle and then put the bottle into the water.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use paint to coat the clip, use a dryer to dry the paint, seal the clip inside a bottle and then put the bottle into the water.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Use paint to coat the clip.</td>
<td>7</td>
</tr>
</tbody>
</table>

We also examined the students’ captured photos daily to keep an observational record of the three experiments in each group. We obtained some of the artefacts created by Group 2 as an example. Figure 4(a) shows a screenshot of the three experimental results on the first day using the three methods (see Table 7). Figure 4(b) shows a screenshot of the experimental results on the last day (1 week after). Figure 4(c) shows the degrees of rusting (0-10) in the group’s three methods, evaluated by the group itself, peers in other groups and the teacher (3, 0 and 0, respectively). These artefacts documented the students’ deepened understanding of rustproofing from experimental design through to observation, presentation and evaluation.

Figure 4. Photo of the first day of the experiment; photo of the second day of the experiment; photo of the experimental results evaluation (a – left; b – middle; c – right)

Group 2 was chosen as having the best experimental rustproofing design, and was awarded a badge declaring its members rustproofing “experts,” which encouraged the students to make further science inquiry. It is worth noting...
that although the students captured photos of the experimental results each day, they rarely entered observational records on the degree of rusting into the “Observational Forms for Rustproofing Experiments.”

To examine how the students advanced their inquiry skills, we investigated students’ inquiry processes supported by Edmodo, where seamless learning across physical spaces (in the classroom and between lessons using tablet PCs, and at home using various devices) and individual and social spaces (online learning and class interactions) was documented. Figure 5 indicates the limited numbers of students’ posts on Edmodo relating to their science inquiry skills (questioning = 1.7%, exploring = 2.2%, explaining = 5.1%, evaluating = 3.9% and extending = 0%). However, other posting categories accounted for the majority of the contributions: coordinating the experimental designs and observations (47.9%), social interactions (23%), greetings (14%) and news sharing (8%). This indicates that the social network platform played an important role in students’ project work orchestration and establishment of intimate relationships.

<table>
<thead>
<tr>
<th>Inquiry skills</th>
<th>No. (%) of postings on Edmodo</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questioning</td>
<td>7 (1.7%)</td>
<td>What shall I use in order not to make the oil not spill over?</td>
</tr>
<tr>
<td>Exploring</td>
<td>9 (2.2%)</td>
<td>How many ways are there to prevent rusting (with Hyperlink)?</td>
</tr>
<tr>
<td>Explaining</td>
<td>21 (5.1%)</td>
<td>This is the method of using oil to prevent rusting (Hyperlink: iron rusting - [Wiki – Wikipedia - the free encyclopedia])</td>
</tr>
<tr>
<td>Evaluating</td>
<td>16 (3.9%)</td>
<td>Good!</td>
</tr>
<tr>
<td>Extending</td>
<td>1 (0%)</td>
<td>Glided iron or metal can prevent rusting</td>
</tr>
<tr>
<td>Others</td>
<td>355 (86.7%)</td>
<td>Chu L, don’t forget to bring the hair dryer.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>You get my phone, Ho Y?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good morning, everyone!</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The second group got the honor of “Rustproof Experts”.</td>
</tr>
</tbody>
</table>

*Figure 5. Categories of students’ posts on Edmodo related to rustproofing and other inquiry skills*

We also interviewed a group of students using the questions guided by the five elements in the inquiry-based learning model to understand their rustproofing inquiry processes. The students expressed great passion for leading their own rustproofing research in real life (supported by Edmodo), where their inquiry experience bridged seamlessly across different spaces. One student made the following statement.

In the past, in science class, we only learned from textbooks. It was boring. Now, we can explore and discuss the best ways for our experimental designs on rustproofing. It is interesting that we can do hands-on experiments ourselves, observe the process of the rustproofing process and take pictures [of the clip rusting results] day by day... If I learned [rustproofing] from textbooks, it would be easy for me to forget what I have learned. But doing hands-on experiments makes it difficult to forget the methods and principles [of rustproofing]. I understand the rustproofing theories better.

The students’ interview results indicate that they acquired solid domain knowledge and developed inquiry skills from the experience of hands-on experimental design and practice.
Discussion

**Principle-based pedagogical design—gains**

This study attempted to develop and implement a principle-based pedagogical design for inquiry-based learning in the seamless learning environment of a Hong Kong primary school. The results show that a teacher with training experience and a good understanding of the principles and inquiry-based learning model was better able to plan and enact student-centred inquiry activities in which students had control over their own learning. This echoes the findings in previous studies that suggest that enhancing teachers’ inquiry-based pedagogical competence in a technology-supported learning environment requires teachers to understand the basic theoretical principles behind inquiry-based pedagogies and how to apply the principles to pedagogical practices (Scardamalia, 2002; Song & Looi, 2012, Zhang et al., 2011). In our study, the students showed improvements in their domain knowledge learning and inquiry skills, especially in terms of “questioning” and “explanation,” which are considered essential elements of inquiry-based learning (Hakkarainen, 2003).

However, students cannot develop a meta-cognitive awareness of inquiry strategies without adequate scaffolding (Lakkala et al., 2005). Our study adopted a guided inquiry-based learning model by providing prompts (“Forms for Experimental Rustproofing Designs” and “Observational Forms for Rustproofing Experiments”) for the students’ experimental designs and observations. Although each group of students generated different experimental designs and hypotheses, they were all on the right track in the inquiry. According to Lin and Lehman (1999), metacognitive skill development is typically fostered by providing students with opportunities to reflect on and monitor their learning performance and revise their investigative strategies. In this regulative process, students are reflective inquirers looking to accomplish projects and gain a deeper understanding of domain knowledge and inquiry skills (Loh et al., 2001). Further, in the “Forms for Experimental Rustproofing Designs,” the students were required to take on different responsibilities in completing the experiment, which increased their awareness of taking collective responsibility for advancing the group’s knowledge (Scardamalia, 2002; Zhang et al., 2011).

In this study, the Edmodo social network platform provided a seamless learning environment for the students to coordinate the inquiry projects and establish a rapport in groups and with peers, which played an important role in advancing their rustproofing knowledge and developing their inquiry strategies. In addition, the students could share their groups’ products with their peers at any time and anywhere on the platform, which allowed them to evaluate other groups’ work and construct knowledge collaboratively. The embedded peer group assessment in the pedagogical design meant that the assessment responsibility was turned over to the students, helping them to develop increased agency when evaluating their own learning progress (Zhang et al., 2011). The students could also submit assignments to the platform and obtain the teacher’s feedback in a timely manner, which encouraged them to learn. This study contributes to the literature on the use of principle-based pedagogical design for guided inquiry-based learning in science in a seamless learning environment at primary level. Nevertheless, it also has some limitations.

**Principle-based pedagogical design—losses**

We identified several issues and limitations to be addressed in future work. First, in terms of the five instructional principles adopted in the research, the teacher was not able to grasp the gist of the principle of using authoritative information constructively in his reflections. Some students copied and pasted information from the Internet directly without acknowledging the source. However, the teacher believed that as long as he asked the students to explore learning resources on the Internet, he was adhering to the principle. He did not further scaffold the students on making constructive use of sources. According to Yeo and Tan (2010), the constructive use of authoritative sources involves the interpretation of meaning in context and plays an important role in deepening and expanding students’ science domain knowledge. Hence, in our next cycle of research, we must elaborate this principle to the teacher. Second, the teacher designed the ‘Observational Forms for Rustproofing Experiments’ (see Appendix II) for the students to document the daily degree of rust over five consecutive weekdays. The findings show that none of the six groups completed this task. The ‘Degree of rusting of the iron clip: 0/10’ item might have been too abstract for Primary 4 students to estimate and record. Nevertheless, the students took some pictures to document the daily rusting process during their break time. In providing scaffolds such as prompts, we suggest that the teacher must
consider the students’ level and cater to their needs. Finally, the results of the research cannot be generalised due to the short time span. Further interactive studies are required to investigate whether the five instructional principles suffice for developing teachers’ principle-based understanding of inquiry-based learning, whether the inquiry-based learning model must be refined and how to make better use of social network technology to support seamless learning environments.

Conclusions and future work

This study explores the use of a principle-based pedagogical design for inquiry-based learning in a seamless learning environment, with resource access and peer interactions initiated in classrooms and extended to online interactions. The inquiry-based learning approach was integrated into the domain knowledge learning process to promote students’ development of inquiry skills. The results demonstrate the effective development and implementation of the pedagogical design in a rustproofing learning unit and the positive effect of the design on students’ domain knowledge gains and inquiry skills. The results reveal the need for further research to accumulate experience and scale-up pedagogical interventions within and across schools. The future scale-up of research efforts should be planned under the “design-based implementation” and “designing for diffusion” approaches (Dearing & Kreuter, 2010; Penuel, Fishman, & Cheng, 2011) to build capacity among teachers in the within- and cross-school settings. Further research will instil in target teachers a positive perception of the research efforts that address persistent practice problems from multiple perspectives during the capacity scale-up across the Hong Kong primary school sector.

References


Li, K. M., & Kong, S. C. (2011). *A report on “review survey(s) on the third strategy on information technology (IT) in Education.”* Hong Kong, China: Education Bureau.


Appendix I
Forms for experimental rustproofing designs

<table>
<thead>
<tr>
<th>Group:</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Group members:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

*Experimental design 1 (Experimental designs 2 and 3 used the same forms as experimental design 1.)*

<table>
<thead>
<tr>
<th>Material</th>
<th>Student in charge</th>
<th>Hypotheses of rustproofing theories</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Design diagrams (in order)

Appendix II
Observational forms for rustproofing experiments

<table>
<thead>
<tr>
<th>Observation date</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation time</td>
<td></td>
</tr>
<tr>
<td>Observer(s)</td>
<td></td>
</tr>
</tbody>
</table>

Experimental design 1 (Also experimental designs 2 and 3)

Degree of rusting of the iron clip: 0/10
Mindtool-Assisted In-Field Learning (MAIL): An Advanced Ubiquitous Learning Project in Taiwan

Gwo-Jen Hwang*, Pi-Hsia Hung, Nian-Shing Chen and Gi-Zen Liu

Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan // Graduate Institute of Measurement and Statistics, National University of Tainan, Tainan, Taiwan // Department of Information Management, National Sun Yat-Sen University, Kaohsiung, Taiwan // Foreign Languages & Literature Department, National Cheng Kung University, Tainan, Taiwan // gihwang.academic@gmail.com // hungps@mail.nutn.edu.tw // nianshing@gmail.com // gizen@mail.ncku.edu.tw
*Corresponding author

ABSTRACT

Scholars have identified that learning in an authentic environment with quality contextual and procedural supports can engage students in thorough observations and knowledge construction. Moreover, the target is that students are able to experience and make sense of all of the learning activities in the real-world environment with meaningful supports, such that their learning motivation can be promoted, knowledge can be sensibly constructed, and skills can be fully developed. To develop potential tutoring strategies and learning activity models using mobile, wireless, and sensing information and communication technologies (ICT) in a real-world learning environment, a four-year national e-learning research project entitled “Mindtool-Assisted In-field Learning (MAIL)” has been funded by the National Science Council of Taiwan since 2008 in an effort to lead the development and innovation of Learning Technology. The integrated project aimed to develop Mindtool-assisted knowledge construction models, assessment models, guidance models, and reflection strategies for cutting-edge context-aware ubiquitous learning. Moreover, a series of learning activities has been conducted to examine the effectiveness of the proposed learning strategies and models. Each year, more than 1,500 students have participated in the in-field learning activities with the designed approaches. Based on the results of a series of experiments, it was found that the students’ learning performance as well as their in-field inquiry ability was significantly improved, showing the effectiveness of the Mindtool-assisted ubiquitous learning approach and the success of the MAIL project. In this paper, the background, objectives, theoretical foundations, systems, research issues, applications, and findings of the MAIL project are presented. Finally, the scaling-up plan for applying these research-proven learning models to all levels of educational settings in Taiwan is also addressed.

Keywords
Mobile learning, Ubiquitous learning, Mindtools, Concept maps, Repertory grid

Background and objectives

Learning Technology (LT), as a trans-disciplinary, professional field of leading human developments and innovations in various situations, disciplines, settings, and industries with advanced technological uses, is always in need of creative applications of hard and soft technologies in order to bring about positive changes (Jonassen, 2004; Liu, 2008). Many educators have identified the importance of situating students in real-world contexts for developing and acquiring knowledge and skills (Brown, Collins, & Duguid, 1989; Lave, 1991; Wong & Looi, 2011). In the meantime, researchers have also pointed out the importance of providing personalized learning supports or knowledge sharing facilities during in-field activities (Sharples, Milrad, Arnedillo-Sánchez, & Vavoula, 2009; So, Seow, & Looi, 2009). The popularity of mobile and wireless information and communication technologies (ICT) has provided good opportunities which match this emerging trend in LT; moreover, the advancement of sensing technology has further enabled learning systems to detect real-world information with various types of information-generating e-readers and e-tags. With the help of these technologies, students are able to learn anytime, anywhere. That is, they are encouraged to learn in various real-world environments with supports from and access to the digitalized world (Hwang, Tsai, Chu, Kinshuk, & Chen, 2012; Looi et al., 2009; Wong, 2012); moreover, dynamic learning systems are developed for the user to engage in more active interactions with other learners as well as the learning system itself (Ogata, Li, Hou, Uosaki, El-Bishouty, & Yano, 2011; Okamoto & Tseng, 2008). Generally speaking, this kind of learning strategy has been called “context-aware ubiquitous learning,” and is a state-of-the-art, particular form of ubiquitous learning (u-learning) as defined by Hwang, Tsai and Yang (2008).

Recently, context-aware u-learning has become a popular issue and research topic in the area of e-learning (Ogata & Yano, 2004; Sollervall, Otter, Milrad, Vogel, & Johansson, 2012; Syyänen, Beale, Sharples, Ahonen, & Lonsdale, 2005). Researchers have attempted to conduct context-aware u-learning activities for various courses; however, it
has been found that, without effective learning strategies or tools, students’ learning performance could be disappointing (Chu, Hwang, & Tsai, 2010; Chen & Li, 2009; Liu, Peng, Wu, & Lin, 2009). Several studies have pointed out that u-learning scenarios could be too complex for most students without some proper guidance or supports, because the students need to make use of both real-world and digitalized world learning resources at the same time (Shih, Hwang, Chu, & Chuang, 2011). Therefore, it has become an important and challenging issue to provide effective learning supports in mobile or ubiquitous learning activities.

Among the various learning strategies and tools, Mindtools have been recognized as an effective way of assisting students to learn in complicated learning contexts with all kinds of ICT. Educators have indicated that “technologies should not support learning by attempting to instruct the learners, but rather should be used as knowledge construction tools that students learn with, not from” (Jonassen, Carr, & Yueh, 1998, p. 1). Mindtools are cognition tools that are able to assist students to think and learn in a meaningful and constructive way through stimulating them to expand their cognitive ability in interpreting, analyzing, synthesizing and organizing their knowledge. Jonassen (1999) defined Mindtools as “a way of using a computer application program to engage learners in constructive, higher-order critical thinking about the subjects they are studying” (p. 9). With the assistance of Mindtools, students’ knowledge can be constructed to reflect what they have learned and realized, instead of merely memorizing or recalling content taught by their teachers.

Mindtools and their logical learning design have been widely developed and used with various computer-based application programs, which include database systems, spreadsheets, expert systems, semantic nets (e.g., concept maps), video conference systems, multimedia and hypermedia editing tools, programming tools, and Microworld environments (Jonassen, 1999). In the past two decades, scholars all over the world have paid much attention to using Mindtools in various practical applications related to in-class learning, blended learning, and totally online learning; nevertheless, using Mindtools in u-learning activities remains an important but challenging issue (Lee, Lee, & Leu, 2009).

To develop Mindtool-supported u-learning approaches and to investigate their effectiveness, a four-year national e-learning project was initiated in Taiwan in 2008. The aim of the project was to develop Mindtool-assisted u-learning environments and strategies. Numerous experiments were conducted to evaluate the effectiveness of applying the Mindtool-assisted u-learning approaches to various in-field activities in terms of students’ learning achievement, motivation, attitudes, cognitive load and technology acceptance. Moreover, the students’ in-field observation and question-raising abilities were measured as well.

**Mindtool-assisted ubiquitous learning approaches**

To facilitate in-field learning within context-aware u-learning environments, two kinds of Mindtools were developed to support u-learning activities in the integrated, collaborative project; that is, the grid-based approach which originated from a knowledge elicitation method for developing expert systems and the concept mapping approach that has been widely adopted in in-class learning, blended learning and totally online learning environments.

**Grid-based Mindtools for ubiquitous learning**

An expert system is a computer system that simulates expert-level reasoning based on the knowledge elicited from domain experts. The process and know-how of acquiring and organizing knowledge from domain experts for building knowledge bases of expert systems is called knowledge engineering (Feigenbaum, 1977). Jonassen (1999) indicated that such a process of collecting and organizing domain knowledge for constructing knowledge bases could engage students in critical thinking; that is, an effective way of employing expert systems as Mindtools is surely to engage students in collecting and organizing knowledge related to the course/learning content they aim to learn following a knowledge acquisition approach.

Among various knowledge acquisition approaches, the repertory grid method originating from the Personal Construct Theory proposed by Kelly (1955) has been widely adopted and discussed (Aranda-Mena & Gameson, 2012; Canning & Holmes, 2006; Boose & Gaines, 1989). A repertory grid can be viewed as a matrix whose columns are element labels and whose rows are construct labels. Elements could be decisions to be made, objects to be
identified, or concepts to be learned, while constructs are traits for featuring the similarities or differences between the elements. A construct consists of a trait (e.g., “Long”) and the opposite of that trait (e.g., “Short”). Meanwhile, a five-scale rating mechanism is usually used to represent the relationships between the elements and the constructs, where “1” represents that the element is inclined to have the trait and “5” represents that the element is inclined to have the extreme opposite characteristic of that trait.

Referring to the “Expert systems as Mindtools” conception proposed by Jonassen (1999) and the repertory grid method, a repertory grid-oriented Mindtool was developed in the MAIL project for supporting in-field ubiquitous learning in several ways. In the earlier stage of this project, the repertory grid-based ubiquitous learning system was used as a guiding system for helping students observe learning targets in the field, collect data based on their observations, and develop repertory grids for organizing the collected data. Since 2008, a series of learning activities has been conducted with the tool to help students identify and classify a set of learning targets (e.g., plants on a school campus, butterflies in ecology gardens, or rocks in laboratories) via guiding them to observe the learning targets and organize what they have found in a repertory grid using mobile devices (Chu, Hwang, & Tsai, 2010; Wu, Hwang, Su, & Huang, 2012).

Before the learning activities, teachers were asked to develop an objective repertory grid (i.e., a repertory grid with correct ratings for each <element, construct> pair) to guide the students to make observations in the field and develop their own repertory grids. In such a learning-guiding approach, the elements and constructs were provided by the teachers; therefore, the students only needed to fill in the rating for each <element, construct> entry based on their observations in the field. Users were encouraged to consider the objective repertory grid in Table 1, in which the elements were "Lalang Grass," "Arigated-leaf croton," "Cuphea," "Indian almond," "Money tree," "Crown of thorns" and "Pink ixora," and the constructs were "leaf shape," "leaf point," "leaf edge," and "number of leaf vein branches." For example, the value of the <Lalang Grass, Leaf-shape> entry is 1, indicating that the leaf shape of Lalang Grass is "Long and thin." On the contrary, the value of the <Indian almond, Leaf-shape> entry is 4, implying that the leaf shape of Indian Almond tends to be "flat and round."

### Table 1. Example of a repertory grid for guiding students to observe plants on a school campus

<table>
<thead>
<tr>
<th>Trait</th>
<th>Lalang Grass</th>
<th>Arigated-leaf croton</th>
<th>Cuphea</th>
<th>Indian almond</th>
<th>Money tree</th>
<th>Opposite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf-shape long and thin</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>Leaf-shape flat and round</td>
</tr>
<tr>
<td>Perfectly smooth leaf edge</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>The leaf edge has deep indents</td>
</tr>
<tr>
<td>The leaf vein has few branches</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>The leaf vein has many branches</td>
</tr>
</tbody>
</table>

During the learning activities, the students were asked to complete their own repertory grids by filling in the rating for each <element, construct> pair based on their observations in the field. If the students failed to give the correct ratings in comparison with those given by the teachers in the objective repertory grid, the learning system would guide them to observe a comparative learning target with the “incorrect main-feature” and would ask them to compare it with the original learning target. For example, if a student observed the plant (learning target) “Lalang Grass” and described its “Leaf point” as “round with a blunt tip” (by giving rating "4"), by comparing the student’s answer with the rating given by the teacher (i.e., “tapering to a long point” with rating "2"), the learning system detected that the student’s answer was incorrect. Accordingly, the learning system tried to find a comparative plant with “round with a blunt tip” leaf points from the objective repertory grid. In this case, “Indian almond” met the condition; therefore, the student was guided to observe the “Indian almond,” and compared its “Leaf point” with that of “Lalang Grass,” as shown in Figure 1. Having made the comparison, the student was then asked to answer the question again. If the student still failed to correctly answer the question, the learning system then provided supplementary materials and the teacher’s rating to the student.

In addition to the guidance from the learning system, the students were able to browse their own repertory grids via the mobile device during the field trips, which was helpful to them in identifying and distinguishing the learning targets via comparing the corresponding ratings of the features of the targets in the grid.
The student is asked to observe the leaf shape of “Indian almond” and compare it with the leaf point of “Lalang Grass”. The answer “Round with a blunt tip” to the “leaf point” of “Lalang Grass” is incorrect.

Figure 1. An example of guiding the student to the plant with the “incorrect feature”

To engage students in higher-order thinking during the in-field learning activities, the repertory grid-based Mindtool has further been used to support collaborative knowledge construction in context-aware ubiquitous learning activities in the MAIL project. For example, one of the applications conducted in 2010 aimed to engage students collaboratively in a more challenging learning task; that is, the teacher only showed them the learning targets/objects (i.e., plants on the school campus) without providing any other guidance during the in-field learning process, meaning that the students needed to determine the constructs for identifying and differentiating the plants themselves as well as observing the plants and collecting data. To enable the students to share their repertory grids and make reflections after referring to the repertory grids developed by their peers, a knowledge-sharing system was developed. Through this knowledge-sharing system, the students could upload their repertory grids to the system, browse others' repertory grids, receive feedback from teachers, and discuss with their peers. The experimental results showed that the quality (i.e., completeness and correctness) of individual students' knowledge structure (i.e., the constructs they used and the ratings they gave to represent the relationships between elements) was significantly improved after the ubiquitous learning activity.

Concept mapping as Mindtools for ubiquitous learning

Concept mapping is a well-known learning tool for helping students organize and visualize knowledge and learning experiences (Chiou, 2008; Fischer, Bruhn, Grasel, & Mandl, 2002; Hwang, Wu, & Kuo, 2013; Novak & Cañas, 2006). It is also an effective assessment tool for helping teachers evaluate students' cognitive levels and knowledge structures (Ingeç, 2009; Liu, Don, & Tsai, 2005; Peng, Su, Chou, & Tsai, 2009; Trent, Pernell, Mungai, & Chimedzu, 1998). In the past decades, many studies have shown the effectiveness of concept mapping in engaging students in meaningful learning, and hence their learning achievements could be improved (Amadieu, Tricot, & Mariné, 2010; Anderson-Inman & Ditson, 1999; Horton et al., 1993; Markham, Mintzes, & Jones, 2006).

In the MAIL project, concept mapping has been employed in context-aware ubiquitous learning activities in several ways. For example, one of the learning activities was conducted for butterfly ecology observations. Before the learning activity, the students were asked to develop a concept map about butterfly ecology based on what they had learned from the textbook. During the in-field observations, the students browsed their concept map when observing the butterflies in the ecology garden. They could modify the concept maps if something different or interesting was observed. Alternatively, they could take notes and modify the concept maps when they went back to the classroom. Figure 2 shows the scenario of the in-field learning activity. The students were situated in the butterfly ecology garden with wireless communication networks. The garden consisted of 25 ecology areas in which particular kinds of butterflies and the related host plants of the butterflies are raised. In this earlier study, an RFID (Radio Frequency Identification) tag was placed in each ecology area and each student held a PDA (Digital Personal Assistant) with an RFID reader. When the students walked into an area, the learning system could detect the information within the tag in that area via the RFID reader and confirm individual students’ locations, such that corresponding learning guidance or support could be provided.
RFID reader

Please find the idea leuconoe clara, as shown in this figure. Use the PDA to sense the orange tag if you find it.

Figure 2. A learning scenario of butterfly ecology observations

In recent studies of the integrated MAIL project, smartphones and the QR-code sensing technology have been used to support extensive self-directed learning in the field. A series of large-scale and long-term activities has been conducted in Chiku Ecology Park in southern Taiwan, where various species of mangroves grow. Figure 3 shows the plan for one of the learning activities. In this activity, the students were equipped with a smartphone to interact with the learning system as well as a telescope for long-distance observations. The learning system provided online instant feedback (e.g., hints to remind the students that part of their answers to the questions raised by the learning system were incorrect) and learning guidance (e.g., clues to find the correct answers to the questions) to the students via wireless communications. Moreover, an e-library was developed to provide supplementary materials for the field-based activities.

A series of concept mapping tasks was developed to scaffold the students' knowledge construction during the field trips in a progressive manner. In the first stage, the learning tasks included multiple-choice questions, short-answer questions, and a structured two-level concept map in order to help the students clarify their knowledge about the basic features of individual mangrove species. In the second stage, the learning tasks were designed to encourage the students to describe the advanced features for classifying different species via the use of developing multiple-level concept maps. In the third stage, the learning tasks aimed to guide the students to compare the species and find the relationships among the species based on what they had learned and observed via developing the multiple-level and cross-relationship concept maps.

Figure 3. The learning objectives of the Chiku Ecology Park observations
Figure 4 shows an illustrative example of a student’s emerging concept map while working in the field. The title of the concept map is “the life of Idea leuconoe clara,” which is a species of butterfly found in Taiwan. Before the field trip, the student developed an initial concept map consisting of four stages (i.e., egg, pupa, larva and imago) to describe the life of Idea leuconoe clara. Later, he went to the butterfly garden to complete the learning tasks. When observing the butterfly ecology in the field, the student browsed the initial concept map and found several facts to be added: (1) In stage 1 of the Idea leuconoe clara, he had only described the egg as being “white or lemon yellow;” however, in the field, he noted that the egg was also “translucent.” Therefore, this new feature was added to the concept map, as shown in block A of Figure 4. (2) In stage 2 of the Idea leuconoe clara, he had not described the features of the larva. When learning in the field, he noted two of its features, that is, “red spots on the sides of the body” and “black and white;” therefore, these two features were added to the concept map, as shown in block B of Figure 4. (3) In stage 4 of the Idea leuconoe clara, he had not given examples of food plants. When observing in the field, he found that the Idea leuconoe claras were acquiring honey from magnolias; therefore, the proposition “magnolia is an example of food plants” was added to the concept map, as shown in block C of Figure 4.

Applications and research items

From August 1, 2008 to September 30, 2012, the research team conducted 73 u-learning activity-based studies to try out the learning system, Mindtools, learning models and strategies, evaluation scales, and the design of the learning content and activities. The in-field learning environments included the Chiku Mangrove Conservation Area, the Chiku Black-faced Spoonbill Conservation Center, the butterfly ecology garden in Cheng-Kung Elementary School, the science parks and museums in several cities across Taiwan, and the campuses of several educational settings in Taiwan. The learning content of the in-field activities not only focused on natural science, but has also been extended to various academic disciplines. So far, there are 38 natural science studies, 19 social science studies, 5 computer science studies, 2 nursing training studies, 7 language learning studies, 1 mathematics study and 1 Art learning study conducted as part of the MALL project.

The total number of participants has increased each year, from about 500 participants in 2008 to nearly 3,000 in 2012, as shown in Figure 5. To date, the total number of participants in this integrated LT project has reached more than 6,000 students. Most of these experiments have been conducted by comparing the learning performance of experimental groups and control groups; moreover, the students’ pre-and post-test scores as well as their perceptions collected based on several measures have been analyzed. It is evident that based on these experiment results, both the
quantity and quality of this integrated, collaborative research project are satisfactory in terms of academic rigor and knowledge innovation.

![Number of participants](image)

Figure 5. The number of students participating in the MAIL ubiquitous learning studies in 2008-2012

**Achievements and implications**

In the early experiments, we aimed to investigate the students’ perceptions of learning with the ubiquitous learning approach in comparison with their past experiences of learning with the traditional one-to-many in-field instruction. In the meantime, the teachers’ perceptions of conducting u-learning activities were also investigated. For example, one of the experiments was conducted to collect the feedback from 30 elementary school students and 9 teachers after they experienced a u-learning activity in the butterfly ecology garden (Peng et al., 2009). From the questionnaire survey, it was found that, in comparison with the traditional instruction, the students’ learning motivation and interest were significantly promoted with the help of the personalized guidance and feedback provided by the u-learning system in the field. The average rating given by the students was 4.53 in a five-point Likert rating scheme. Moreover, from the interviews, it was found that the teachers highly accepted the u-learning approach owing to several reasons: (1) it provided the students with better access to online resources during the field trip; (2) it enabled the students to make observations and collections with learning guidance without being constrained by time or location; (3) the u-learning activities engaged the students in learner-centered activities seamlessly across locations and contexts; and (4) the u-learning approach was able to provide step-by-step expert advice and record the students’ learning portfolios.

To improve the students’ learning achievements, in the second stage of the MAIL project, we aimed to compare the effectiveness of the Mindtool-assisted in-field learning approaches with that of the conventional tour-based u-learning approach, which guides individual students in the field, providing them with supplementary materials and giving feedback to them based on their observations and input. It was found that, with the assistance of the Mindtools, the students’ learning achievements, as well as their learning attitudes, were significantly improved. Moreover, it was also found that via the sharing of the constructed knowledge (e.g., repertory grids or concept maps), the knowledge structures as well as learning achievements of the students were further improved. For example, in an experiment for conducting a “plant identification” activity at an elementary school campus, the repertory grid method was implemented in the u-learning system to serve as a Mindtool to help the students summarize the features of the plants observed in the field (Chu, Hwang, & Tsai, 2010). From the experimental results, it was found that the Mindtool-integrated approach not only enhanced the learning interest (the average rating changed from 4.85 to 5.31 in a six-point Likert rating scheme), but also improved the learning achievements in comparison with the conventional u-learning approach via ANCOVA analysis ($F = 9.573$, $p = 0.011$ and $d = 1.39$) for the two groups of students. Another experiment was conducted in the butterfly ecology garden with embedded concept mapping in the u-
learning system to help students organize what they had observed in the field and compare the acquired knowledge
with their prior knowledge learned from the textbooks (Hwang, Shi, & Chu, 2011). The experimental results showed
that, after the learning activity, the students who learned with the Mindtool-based u-learning approach showed a
significantly positive change in their attitudes toward learning science (from an average rating of 3.97 to 4.38 in a
five-point Likert rating scheme); moreover, their learning achievements were significantly improved in comparison
with the achievements of those who learned with the traditional concept maps (with paper and pencil) in the field and
the conventional u-learning approach based on the ANCOVA result ($F = 4.257, p < 0.05$).

In the third stage of the MAIL project, Mindtool-integrated u-learning was included in the formal science
curriculums of several selected schools in Taiwan. Accordingly, several long-term activities were conducted in field
trips to observe the growth of students’ inquiry competences with the u-learning approach, such as problem-posing
and problem-solving abilities. For example, in one of the activities conducted in the Chiku Ecology Park, the
participating students were forty-nine elementary school students aged 11.5 years old on average. The students
experienced the field trips within four months to complete a series of learning tasks. Twenty-five of them who were
assigned to the experimental group learned with the u-learning approach. Another twenty-four students who were the
control group learned with the traditional in-field instruction; that is, they were guided and instructed by the teacher
on the field trip. The students’ inquiry performances were evaluated by the teachers based on several criteria, including the quantity and accuracy of the descriptions of the learning targets for completing the learning tasks, the number and quality of the questions raised and the responses to the peers’ questions during the field trip, and the relevance and correctness of the features and relationships used to describe their findings in the learning diaries. It was found that through the assistance of Mindtools, the students’ inquiry behaviors, such as the quantity and quality of the questions they raised and the depth of their descriptions of their observations in the field, were significantly increased in comparison with traditional in-field learning based on the ANCOVA result ($F = 4.72$ and $p < 0.05$); in the meantime, the students’ learning performances were significantly improved.

Another three-month experiment was conducted to compare the learning performance of 18 gifted students and 30
average students who were 11.5 years old on average. The participants were scheduled to learn with the concept
map-based u-learning approach in the Chiku Ecology Park. Within the three months, the two groups of students
showed remarkable progress in ecology observations based on the Computerized Ecology Observation Competence
Assessment (CEOCA) developed by Hung, Hwang, Lin, Hung and Wu (2010). The CEOCA consisted of three facets,
that is, knowledge, observations and conceptual relationships. The test items were presented with real pictures, films
or concept maps. In the pre-test, the average performance of all of the participants was close to the norm (0.04 vs.
0.00) of the students of the same age in Taiwan. After the learning activity, the post-test scores showed that the
average growth slope of all of the participants was significant ($\mu = 0.27, p < .01$) in comparison with their pre-test
scores with effect 0.53; however, there was no significant difference between the two groups. By conducting a
follow-up test one month later, a significant difference was found in the CEOCA scores between the two groups. The
gifted students revealed positive performance growth, while the performance of the average students decayed after
the learning activity, showing the need to provide continuous supports to average students after field trips (Hung,
Hwang, Lin, & Su, 2012).

Furthermore, some experimental results also showed that the Mindtool-assisted u-learning approach can help
students improve not only their learning achievements, but also their higher-order critical thinking competences.
For example, in one of the u-learning activities conducted in the butterfly ecology garden, the students were asked to
develop repertory grids based on what they observed on the field trip (Hwang, Chu, Lin, & Tsai, 2011). The experimental results showed that the students who learned with the Mindtool-based u-learning approach showed better learning achievements than those who learned with the conventional u-learning approach. By comparing the students’ answers to the learning sheets before and after participating in the repertory grid-based u-learning activity using a t-test, it was found that the students’ ability of determining the characteristics for differentiating the butterflies and their competence for identifying and differentiating the butterflies had significantly improved with $t = 7.13 (p < 0.001)$ and $t = 9.23 (p < 0.001)$, respectively. This implies that their higher order thinking (i.e., analysis and evaluation) performance was improved.

In addition to the lead-in of various Mindtool-based u-learning strategies, it should be noted that some of the
participating schools of the MAIL project have already included such Mindtool-assisted u-learning approaches as
part of their regular curricula. For example, a nursing school in southern Taiwan not only prepared their own u-
learning equipment (i.e., mobile devices, wireless networks and sensing devices) after participating in one of the
Another issue raised in the MAIL project was the cognitive load of the students who participated in the u-learning activities (Hwang, Wu, Zhuang, & Huang, 2013). As the students needed to interact with the real-world learning environment as well as the e-learning system simultaneously, there was a concern that their cognitive load might be too great in some cases; therefore, several experiments of the MAIL project measured the students’ cognitive load using the measures developed by Paas (1992) and Sweller, van Merriënboer, & Paas (1998). It was found that, with a proper learning design, the Mindtool-assisted u-learning approach could significantly decrease students’ cognitive load; on the contrary, students were likely to meaningfully expand their cognitive capability after the practice of integrating in-field observation and technology-driven knowledge construction into situated learning. This decrease could be due to the fact that the Mindtools were able to assist the students in organizing the collected data from the field by linking the chunks of information in a well-structured form, which eased their load in interpreting the data (Verhoeven, Schnotz, & Paas, 2009). For example, the repertory grid-based Mindtools can help students organize the observed features of the learning targets in a unified form (i.e., ratings ranging from 1 to 5), which is very helpful to them for comparing the learning targets and identifying the significant features that can be used to distinguish the targets. Consequently, students’ cognitive load could be decreased; in the meantime, their learning achievements could be improved owing to learning in a more efficient and effective way.

From the series of related studies conducted in the MAIL project, it is found that technologies are not the key or solution to cope with in-field learning problems. Without proper learning supports in the field, students might feel helpless, frustrated and aimless, and hence their learning attitudes or motivations could be affected. Moreover, their cognitive load can be high owing to the strategies or tools used to link what they have learned and observed together, and hence their learning achievements could be disappointing. On the other hand, from the experimental results, it is also suggested that grid-based tools are effective in helping students identify and differentiate a set of learning targets, while concept mapping tools are helpful to students in linking and organizing what they have observed in the field and have learned from the textbooks. That is, grid-based tools help students observe the learning targets with a "micro view," while concept mapping enables them to see things with a "global view." For example, if the aim of a context-aware activity of a language course is to help students learn to use vocabulary, phrases and sentence patterns related to the contexts, concept mapping could be useful; similarly, if the aim of a social studies course is to let students have a whole picture of a cultural asset, concept mapping is also a good choice. Nevertheless, if the aim is to foster students’ ability of identifying or differentiating the artifacts from different historical periods, grid-based Mindtools are good candidates.

Therefore, when designing Mindtool-based u-learning activities for different subjects, the following procedure is suggested:

1. Review the nature of the learning content to see if the aim of the subject unit is relevant to identifying and differentiating a set of learning targets based on their features, or organizing the relevant concepts by finding the relationships between them. Accordingly, the Mindtools to be employed in the learning activity can be determined.

2. Design the learning tasks based on the aims of the activity. For the learning activities with grid-based Mindtools, both problem-based and inquiry-based learning tasks are recommended, depending on the level of learner control. For novice or younger learners, problem-based learning with instant feedback would be preferable; for experienced or older ones, inquiry-based learning with supplemental materials in e-libraries or on the web would be better. On the other hand, for the activities which incorporate concept mapping strategies, inquiry-based learning is recommended.

3. Determine the technologies used in the learning activities. It is suggested that at least mobile and wireless communication technologies are required, while sensing technologies are optional. One of the reasons for adopting sensing technologies is to provide students with learning tasks, learning supports or supplementary materials at the right place and at the right time, which not only reduces the load of students in searching for the information, but also makes the learning process more efficient.

4. Determine the way to measure the learning performance of students and provide feedback to them. For the activities using grid-based Mindtools with the problem-based learning approach, automatic scoring and instant
Conclusions and future work

In this paper, an advanced u-learning project entitled “MAIL” with various dimensions of research design, issues and contributions, has been presented. The integrated project has aimed to develop Mindtool-assisted u-learning environments to improve the in-field learning performance of students. A total of 73 u-learning activity-based studies have been conducted in the past five years to investigate the effectiveness of the Mindtool-assisted u-learning approach in terms of improving students’ learning achievements, learning motivation, learning attitudes, and technology acceptance degrees, among other aspects. The experimental results show that the research-proven approach with multiple practices in various settings is both promising and appealing.

In terms of LT innovation, the findings of the MAIL project provide several new contributions to the field of mobile and ubiquitous learning. It has been demonstrated that simply adopting new technologies for students to learn in a real-world learning environment is not good enough. What is more important is for us to design appropriate pedagogical strategies as Mindtools for providing better support to students in an authentic in-field ubiquitous learning environment with procedural and contextual components. It is expected that the accomplishments of the MAIL project can provide research-proven LT know-how of Mindtool-assisted ubiquitous learning as well as references for those researchers and practitioners who are interested in conducting in-field activities with instant supports from technologies.

The various designs and experiments described in this paper can serve as a good reference model for practitioners and researchers who are interested in this emerging field of LT. This paper also reveals several essential future research topics, which are summarized as follows:

- Track students’ learning activity logs as a way to support learning analytics studies in mobile and ubiquitous learning environments. For example, it would be interesting to analyze the students’ learning patterns and investigate the relationships between the patterns and their learning performance. Moreover, it is important to further examine the effects of the Mindtool-based u-learning on students’ higher order thinking based on the learning logs of students’ in-field learning behaviors.

- Provide instant and personalized learning supports based on the learning logs and profiles of individual students. Although Mindtools are theoretically helpful to students in constructing and organizing knowledge, students might find it difficult to effectively use Mindtools during the in-field learning activities. For example, some students might have difficulty in developing concept maps without appropriate assistance. That is, while learning with Mindtools in the field, students might require instant and personalized supports. Therefore, it is important to provide instant learning supports by analyzing the learning logs and profiles to identify their problems and needs in the field.

- Develop seamless learning environments by integrating front-end in-field learning experiences with the backend support of Learning Management Systems (LMS) by using the cloud technology so as to apply versatile Mindtools in more courses. In addition to the concept map and grid-based Mindtools developed in this paper, other Mindtools reported by Jonassen (2004) could be included for helping students learn in more effective and constructive ways. For example, spreadsheets could be an effective Mindtool for helping students infer the relationships between variables in Mathematics and Physics courses; database management systems could be Mindtools that engage students in analytical tasks; simulation software could be helpful to students in associating abstract theories with real-world scenarios. Therefore, it is worth investigating the possibility and effectiveness of applying those Mindtools to different u-learning activities.

These cutting-edge research topics and issues are worth our efforts to shed more light on this ever-changing, promising field of context-aware ubiquitous learning in Learning Technology. Recently, the Ministry of Education in Taiwan has initiated a large-scale program for applying mobile and ubiquitous strategies and tools in all levels of schools. In 2012, one hundred schools were selected as demonstration sites, and the number of schools participating in the program will be increased each year. In those schools, each student in the selected classes is equipped with a
mobile device. In each city or county, a cloud-based educational service system has been established to support the anywhere and anytime learning. Moreover, a series of training programs has been proposed to train teachers in how to design in-class and in-field activities with the strategies and tools developed based on the experiences and findings of MAIL and some other studies. It is expected that mobile and ubiquitous learning will become a regular form of learning in the coming five years in Taiwan.

Acknowledgments

This study is supported in part by the National Science Council of the Republic of China under contract numbers NSC 100-2631-S-011-002 and NSC 100-2631-S-011-003.

References


Mobile Phones for Spain’s University Entrance Examination Language Test

Jesús García Laborda1*, Teresa Magal Royo2, Mary Frances Litzler1 and José Luis Giménez López2

1Departamento de Filología Moderna, Universidad de Alcalá, calle Trinidad, 3, 28801 Alcalá de Henares (Madrid), Spain // 2Departamento de Ingeniería Gráfica, Universidad Politécnica de Valencia, Camino de Vera, s/n, 46022 Valencia, Spain // jesus.garcialaborda@uah.es // tmagal@degi.upv.es // mf.english.uah@gmail.com // jojilo@degi.upv.es

*Corresponding author

ABSTRACT

Few tests were delivered using mobile phones a few years ago, but the flexibility and capability of these devices make them valuable tools even for high stakes testing. This paper addresses research done through the PAULEX (2007-2010) and OPENPAU (2012-2014) research projects at the Universidad Politécnica de Valencia and Universidad de Alcalá (Spain) to provide a powerful but low cost delivery system for the foreign language paper of the Spanish College Entrance Examination (henceforth PAU). The first project, PAULEX, intended to create a robust mobile platform for language testing while the second, OPENPAU, examined the specific applications of ubiquitous devices to create more dynamic forms of assessment. This paper focuses on the projects’ design, testing theory, and technical evolution including visual ergonomics. The current results demonstrate the technical and didactic feasibility of mobile-based formal assessment that aligns student needs with the kind of inferences that the mobile based language test should provide academic authorities.

Keywords

Mobile learning, High-stakes testing, College entrance examination, Foreign language, Higher education

Introduction

Mobile phones have been playing an increasingly significant role in education in the last years, and although until recently very few tests were delivered through them, their flexibility and capability to do so have suggested their potential even for high stakes-testing. High-stakes testing can be defined as those tests with important consequences for the test taker such as acceptance to university, a scholarship, or a license to practice a profession, all of which may have a great influence of the testee’s life. Using mobiles beyond their traditional uses such as podcasts, mp3 applications, and even learning apps seems to be a real challenge at this point, yet they have already been used for language testing as in PhonePass, previously called SET-10 (http://www.7act.net/7ACT_files/set10.pdf) test. The validity of the test has repeatedly been supported (Downey, Farhady, Present-Thomas, Suzuki & Van Moere, 2008) but little evidence has been provided of its operativeness in real educational contexts. As a consequence, the potential opportunities for mobiles for language testing are still open (Valk, Rashid, & Elder, 2010) but sound projects need to be implemented. This is the kind of research that the Ministry of Education started to support in 2007. By that year, the Spanish and regional educational authorities responsible for the high stakes University Entrance Examination (“Prueba de Acceso a la Universidad”, henceforth PAU) had determined the need to design a new test with greater validity than the current paper-based test, which only included the traditional tasks of reading and writing along with grammar questions. The new test had to include listening and speaking activities. However, budget cuts reduced the validity than the current paper-based test, which only included the traditional tasks of reading and writing along with grammar questions. The new test had to include listening and speaking activities. However, budget cuts reduced the possibilities of implementing a new exam that could include speaking and listening activities unless a low-cost delivery system was found. With a view toward designing a modern test in the hopes of saving the possibility of including the two skills, the Universidad Politécnica de Valencia obtained funds for the development of an online testing system (PAULEX, “PAU en Lenguas Extranjeras”) project from 2007-2010 and the OPENPAU (“PAU abierta”) later on between 2012-2014 whose results were described by García Laborda (2012). After testing over 150 students online it was concluded that implementation of the computer test would save human resources and be economically feasible in a period of two to three years. As a subproject, the PAULEX project addressed the use of mobile phones (García Laborda & Giménez Lópe, 2010), which is the main focus of this paper.

Literature review and theoretical approach

When addressing this sub-project, the research team felt that the use of mobile phones, like any other delivery system, could not challenge three main testing features: validity, reliability, and practicality (Bachman & Palmer, 1996, among others). Validity here means that if a student gets a score of X on the test, it means that he should be
able to study using that foreign language at university; reliability provides information on the “the precision of the test measurement” (Salmani-Nodoushan, 2009, p. 1); and practicality implies that the test can be implemented in real life. In addition, the test construct in which mobiles are to be applied has taken into account current theories in language testing and Communicative Competence (Canale & Swain, 1980; Canale, 1984; Bachman & Palmer, 1996). Thus, the questions considered were (1) why use mobile phones for language testing? (2) How can the basic testing features be assured? (3) What learning theories are implicated in their use?

For most part, the collection of evidence in both projects was based on Weir’s validation framework and the Evidence Centered Design (Mislevy, Steinberg, & Almond, 2002; Mislevy & Haertel, 2006). Weir (2005) feels that the reliability of a test depends mostly on its conditions for validation. For him, it is necessary to have warrants that there will be two main types of validity in implementing a test: context validity and theory-based validity. Both are interrelated and need to be considered interdependently. Context validity is divided into three parts: task, setting and administration, and task demand (similar to test construct), while theory-based validity includes executive processes and executive resources. Given this framework, the theoretical application for the mobile application of the PAU took into account the aspects included in figure 1.

![Figure 1. Delivery framework (based on Weir, 2005)](image-url)
As indicated in figure 1, the theoretical aspects that needed to be considered in the implementation of the mobile-based test were grouped into three main phases: design, delivery, and consequential. The limitations and scope of this paper only allow for a discussion of the most important aspects of the design phase. Within this framework, this paper will mostly focus on contextual validity because this is where Weir places the delivery system factor. However, there is no question that the implications of using certain delivery systems—whether pen-and-paper or mobile—are present in all the aspects presented in figure 1. This approach is based on performance-oriented tasks, which intend to resemble communicative acts of the language. In designing the test process in the test, two main options were included: a cognitivist and a social constructivist intervention. While the cognitivist approach supports the notion that students bring knowledge at the time of testing and this is represented by observable behaviors, the social constructivist approaches in language testing are very much related to the development of the Zone of Proximal Development (ZPD) (for further discussion see Poehner, 2008) through the examiner’s intervention and moderation. The Zone of Proximal Development is defined as “...the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978, p. 86). Based on the development of the ZPD, Poehner (2008) also suggests that Dynamic Assessment includes both approaches. Both PAULEX and OPENPAU projects follow the principles of Dynamic Assessment (DA). This approach considers that students have a current level of language knowledge (the one they show without test moderation) and a potential level (the one that they show when their output is computer or human-based moderated and improved through this interaction). According to DA, both can be included in a two-part assessment if the first is moderated (socio-cognitive approach) and the second just serves to obtain current language evidence (cognitivist approach) without any tester’s intervention. This can be seen the process shown on figure 2.

According to these principles and given their experimental nature, the PAULEX and OPENPAU projects placed more emphasis on achieving a sound design based on experimental evidence than on potential achievement scores through the use of the test. In practice, evidence was collected and recorded through the use of mobiles. There were five main benefits that justified the decision: (1) the lower cost of mobile based hardware; (2) immediacy of rating and results; (3) ease of recording during oral interviews (hence, data available for further revision of the test and research); (4) the candidates’ familiarity with the delivery means; (5) possibilities for students to rehearse; and (6) ease of rating and administration. Additionally, accessibility for schools and/or official testing centers would enable the optimization of space.

Evidence obtained from the tests, which was moderated, was processed through the Evidence Centered Design (figure 1) (Mislevy, Steinberg, & Almond, 2000; Mislevy & Haertel, 2006) after the first interview, and implied the design of adequate tasks that considered all the linguistic requirements (as seen in figure 1) and could be delivered through mobile phones. According to the cognitivist approach, used in the second testing session, tasks had to be automatically delivered and recorded without moderation to provide current real data. Then the responses were rated (and the scores validated) and with a view toward having an impact on decisions for teaching and high stakes decisions.

Overall, the research team felt that the use of mobile phones was strongly founded but they recognized that the advantages and disadvantages needed to be weighed. Table 1 presents the pros and cons of their use:
Table 1. Use of mobile phones according to test characteristics

<table>
<thead>
<tr>
<th>Test characteristics</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place and time circumstances</td>
<td>Convenience of location and time because they require no external human presence</td>
<td>Security and technical assurance of full functionality required; otherwise test is at stake</td>
</tr>
<tr>
<td>Test rubric and process for responses</td>
<td>Tasks are consistent and delivered exactly in the same way to all testees</td>
<td>Testees may have different levels of understanding of the instructions according to proficiency levels</td>
</tr>
<tr>
<td>Test delivery</td>
<td>Current developments in mobile phones increasingly enable the inclusion of audio and image, thus enhancing contextualization and richness of responses</td>
<td>Standardization of mobile phones used for the test is needed; otherwise significant differences in responses can be found even by the same speaker.</td>
</tr>
<tr>
<td>Construct, rating, and scoring</td>
<td>Automated rating validates the equanimity in multiple choice items; separation of the rater from the testee enables rating protocols to be followed without the influence of contact with the testee. Thus assessments are more objective.</td>
<td>Human protocols do not assure complete equanimity (Baldwin, Fowles, &amp; Livingston, 2008).</td>
</tr>
</tbody>
</table>

Mobile phones in high stakes testing: The PAULEX project (2007-2010)

Since the implementation of the originally planned computer-based language testing platform was costly - albeit assumable in the long term-, one of the suggestions for the researchers was the use of mobile phones for the Speaking test only until the online platform could be used. However, while mobiles were originally thought to support student training, almost from the beginning the project management felt that they could also have a very positive effect on learning and they could encourage after-test washback effects. The main reasons to implement mobile phones were that the hardware was less costly than for computers, their use could be more accessible as they can easily be delivered and collected to and from each school, and their use could facilitate rapid assessments by testing units (which would resemble calling centers in their functioning and organization). These testing units could potentially organize and deliver a large number of tests in a limited time. The tests could be delivered automatically; the students’ responses could be recorded and assessed later by human raters.

From the beginning of this three-year project, it was clear that a well-trimmed double design project was needed for the delivery, ergonomics, and content inclusion. Figure 1 describes the organization of the PAULEX project.

![Figure 2. Organization diagram of PAULEX project](image-url)
As can be seen in figure 2, two branches were organized: one devoted to the linguistic and validation aspects, and the other focused on the technological design of the online and mobile platforms. From the beginning it was clear that most of the significant difficulties were associated with the test design since the technology group had already been involved in similar projects before. Because the validation process was central to the project, the mobile application was designed and tested considering a variety of students and also bearing in mind that the PAU project served to obtain inferences of whether students would be able to use English for university work. Furthermore, the mobile technology branch considered that not all students have the same ability in using mobile technology so the technological specifications were relevant and accessible to students with special needs.

The development of the mobile phone subproject within the larger PLEVALEX project was intended to provide information on three aspects: (1) student adaptability to the new environment, (2) content and test validity for the listening/speaking tasks, and (3) delivery reliability. As mentioned above, mobile phones have been thought to foster learning more than to be used to assess students. Learning would take place by providing them with test samples that could be used anywhere and at any time. In this way mobile phones would bring to the fore the required testing skills in combination with similar listening and speaking tasks along with affective considerations but in a more interactive and usable manner. This process would also provide opportunities for authentic learning and the elimination of test fear would probably favor motivation. In this sense, the mobile phone sub-project sought to engage students in terms of motivation, high stakes test practice, and language learning. The results for this project were obtained through triangulating linguistic achievements, field notes, and a usability analysis carried out through a 20-item questionnaire computer delivered to all the students who took part in the research (García Laborda, Giménez López, & Magal Royo, 2011).

Validation method

We used five types of validation analysis for the PAULEX project. First, we did a Delphi analysis (Custer, Scarcella, & Stewart, 1999) to foresee potential issues in the mobile phone test with experts and then a reduced number of regular users (3). Second, we observed the intended scores of the mobile phone users and compared them with those obtained with the online platform (Sariola, 2003). Next, we analyzed the video recordings from the pilot studies. After that, inspection techniques were followed to do a usability analysis (Nielsen & Mack, 1994). Finally, the students’ attitudes toward mobile phone use for the test were analyzed (as seen above).

Usability analysis

As discussed above, the first test was conducted as a pilot test with a small sample of students, to detect faults in the design of the application and to debug and test the viability. After making some corrections such as adapting some aspects of the content and navigability, a second review of the application was made using potential users. In this second test the number of the sample was expanded to 144 individuals in the last year of high school (aged from 17-18 years), all of whom lived and studied in the area of La Oliva-Gandia (Valencia, Spain) (see table 2).

<table>
<thead>
<tr>
<th>Table 2. Students’ school of origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>High school</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>1 IES Tirant lo Blanc</td>
</tr>
<tr>
<td>2 IES Monduver</td>
</tr>
<tr>
<td>3 IES Veles e Vents</td>
</tr>
<tr>
<td>4 IES Maria Enriquez</td>
</tr>
<tr>
<td>5 IES Ausias March</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

To evaluate the usability of the interface, a likert scale questionnaire ranging from 1 to 4 (to avoid indecisions) was used.
Results of the second test

Once collected, the data were processed using the SPSS statistical program. The first part of the test, which related to knowledge of the environment, focused on aspects to justify routine use and availability of phones, adaptation to the environment of the test items, and utility-satisfaction.

<table>
<thead>
<tr>
<th>Responses</th>
<th>Frequency</th>
<th>Percentage</th>
<th>Valid Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totally disagree</td>
<td>9</td>
<td>6,3</td>
<td>6,3</td>
</tr>
<tr>
<td>Disagree</td>
<td>35</td>
<td>24,3</td>
<td>24,3</td>
</tr>
<tr>
<td>Agree</td>
<td>89</td>
<td>61,8</td>
<td>61,8</td>
</tr>
<tr>
<td>Totally Agree</td>
<td>11</td>
<td>7,6</td>
<td>7,6</td>
</tr>
<tr>
<td>Total</td>
<td>144</td>
<td>100,0</td>
<td>100,0</td>
</tr>
</tbody>
</table>

Table 4. Students’ attitudes toward the mobile-based tool usefulness

<table>
<thead>
<tr>
<th>Responses</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totally disagree</td>
<td>9</td>
<td>6,3</td>
</tr>
<tr>
<td>Disagree</td>
<td>16</td>
<td>11,1</td>
</tr>
<tr>
<td>Agree</td>
<td>100</td>
<td>69,4</td>
</tr>
<tr>
<td>Totally Agree</td>
<td>19</td>
<td>13,2</td>
</tr>
<tr>
<td>Total</td>
<td>144</td>
<td>100,0</td>
</tr>
</tbody>
</table>

Table 5. Students’ attitudes toward the mobile-based tool time facilitator

<table>
<thead>
<tr>
<th>Responses</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totally disagree</td>
<td>10</td>
<td>6,9</td>
</tr>
<tr>
<td>Disagree</td>
<td>21</td>
<td>14,6</td>
</tr>
<tr>
<td>Agree</td>
<td>69</td>
<td>47,9</td>
</tr>
<tr>
<td>Totally Agree</td>
<td>44</td>
<td>30,6</td>
</tr>
<tr>
<td>Total</td>
<td>144</td>
<td>100,0</td>
</tr>
</tbody>
</table>

Once the descriptive data had been surveyed and the group statistics had been examined, it was determined that the results were satisfactory as a whole. The results are above 1.5 on average (on a 0-3 scale).

Table 6. Students’ attitudes toward other factors in relation to mobile phone use for language testing

<table>
<thead>
<tr>
<th>Responses</th>
<th>Valid</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Mode</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A mobile-based task design helps me to perform better</td>
<td>144</td>
<td>1,71</td>
<td>0,058</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>A mobile-delivered test is useful</td>
<td>144</td>
<td>1,90</td>
<td>0,058</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Mobiles help to save time in taking this test</td>
<td>144</td>
<td>2,02</td>
<td>0,071</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Mobiles are adequate to cope with my needs for this test</td>
<td>144</td>
<td>1,74</td>
<td>0,060</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>I learned to use the application quickly</td>
<td>144</td>
<td>2,44</td>
<td>0,057</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>It is easy to remember how to use the application</td>
<td>144</td>
<td>2,45</td>
<td>0,053</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>I became familiar with the application easily</td>
<td>144</td>
<td>2,32</td>
<td>0,051</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>I think this is a good application</td>
<td>144</td>
<td>1,94</td>
<td>0,062</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>It is user friendly</td>
<td>144</td>
<td>1,94</td>
<td>0,060</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>The application works as I expected</td>
<td>144</td>
<td>1,92</td>
<td>0,059</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>I would recommend its use to other students</td>
<td>144</td>
<td>1,90</td>
<td>0,075</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
As observed, in general the students valued the use of mobile phones for language testing very positively. The results of these tests have led us to continue with the research, which is still currently being developed.

Advancing toward solutions for the PAU: The OPENPAU project (2012-2014)

In the years between the PAULEX project and the beginning of the OPENPAU project, Spain started experiencing one of the worst financial crises in its history. In that context, the research team of the PAULEX project observed that mobile phones would be a valuable asset in testing oral skills (speaking and listening) and reading efficiently at a low cost. However, the team also considered that mobiles would be inappropriate for writing due to the intrinsic difficulty of keyboard use (García Laborda, Giménez López, & Magal Royo, 2011; Park, 2011). The design principles for developing a mobile phone application were the following (also Keskin & Metcalf, 2011):

- Use of video communication with the examiner or a video delivery system if videos are used (possibly the most likely situation),
- Creation of a podcast library for student test preparation,
- Adequate real or deployed time access and adequate connectivity,
- Augmented reality possibilities,
- Mobile Blackboard or a similar platform for test preparation.

The first trials on a large scale are expected to begin by September 2013. The technology is designed to incorporate these conditions. The following sections address this concern. Thus far only the Delphi analysis and a very small sample of research have been complete.

Significant results of the PAULEX and OPENPAU project for m-testing technology

The results hereby presented are mostly related to the observations and research undertaken between the end of the PAULEX project and the beginning of the OPENPAU project. However, the ideas are based on the results from PAULEX and the triangulation of the Delphi method and focus groups. From their reactions and opinions, we concluded that the different kinds of interfaces for mobile devices favor tasks such as speaking and listening, even multiple choice tasks but are rather limited for reading and especially for writing. In the case of online exams with different kinds of tasks, adjustments must be made to navigation and content so that users can feel more comfortable when viewing and inputting information or data. The adaptation of user interfaces for certain tasks in a limited period of time requires prior understanding of certain determining factors such as the physical, functional, and formal accessibility of the application. For example, an interface with a hierarchical menu on a mobile phone is useful for beginning users because the appropriate options can be selected through the presentation of a series of menus. Hierarchical menus require relatively higher numbers of key clicks but this is acceptable for novel users who need help using unfamiliar navigation systems and thus leads to diminishing differences due to technology knowledge and serves to validate the use of mobiles as delivery system (Weir, 2005).

Interface design

The most recent interfaces developed have been designed with specific criteria for users taking the university entrance exam. The fundamental criteria studied in this period were accessibility, ergonomics, and the functionality or usability of the application.

Accessibility

Both projects followed the criteria for technical accessibility for interface design proposed by the World Wide Web Consortium, W3C, so that they reached the largest possible number possible of students as end users including those having visual or auditory impairments. This was achieved mainly by following the Web Content Accessibility Guidelines, WCAG 2.0 (http://www.w3.org/TR/WCAG20). In terms of accessibility of the contents of the university entrance exam, we considered the type of programming language used for navigation and also established guidelines for information access (Nelly et al., 2009). The first applications developed, PAULEX and OPENPAU, were created
in environments accessible on Internet with access to the contents delivered online by way of contextual menus for accessing the task management area and student management area. As for the exams created for the students, special attention was paid to visual and functional accessibility of navigation during the final exam tasks. The tests permitted us to determine the potential for mobile phones and the testees’ acceptance mobil phones (Magal-Royo, Fajarnes, Tortajada Montañana, & Defez Garcia, 2007). At the same time, they revealed the importance of two determining factors in the development of future applications: the present rate of technological progress of ubiquitous devices, and the adaptation of contents to the restrictive conditions imposed by them.

Ergonomics

The ergonomic aspects examined for the different applications created for mobile devices focused on the visual ergonomics that enabled students to focus easily and effectively on completing the different tasks. Various studies conducted after the experts’ research revealed the need to establish formal visual guidelines for the content of language learning tasks to enable navigation that is directed and transparent (Weining, Heng, & Guoping, 2007; Garcia Laborda, Magal-Royo, de Siqueira Rocha, & Álvarez, 2010). Ubiquitous devices (mobiles, PDAs, smart phones, netbooks, etc.) can be small in size. That is, they have small screens that limit the space for user interface and the information available on them thus writing and reading tasks may have an additional difficulty due to the fact that a global vision of the read or written text is always desirable. In fact, the information shown must be carefully selected and presented so that it facilitates user interaction not only with the device but mostly with the task content. The major problem with the large variety of screens on ubiquitous devices is the direct impact on information access and visualization because no normalized standards have been established so far (Chae & Kim, 2004; Piolat, Roussey, & Thunin, 1997). This problem was of the most significant ones in test validity. In theory, if the test is implemented, the Ministry should provide all the testees with the same mobile phone to avoid biased or unfair testing conditions.

Limited data input mechanisms

These devices are also limited in terms of data entry procedures because of their reduced size. The methods used most often for mobile devices nowadays are the keypad, which has more than one function associated to each key, and the touch screen. Both methods require a high degree of attention on the part of users and can lead to errors, a situation that limits how they can be used.

Thanks to improvements to user interaction mechanisms now found on ubiquitous devices, different channels can be used for data input, which can be simultaneous, synchronized, or combined for certain tasks. In the case of the mobile phones, which have different kinds of exercises (oral, comprehension, writing, etc.), data entry mechanisms can slow down or directly affect completion of the exam, for example, on the reading test, which can involve reading a long text or typing using a virtual physical (Giménez López, Magal Royo, Garcia Laborda, Garde Calvo, & Prefasi Gomar, 2009) (see figure 3). The approach to design this interfaces was taken from the socio-constructivist theory of language that uses images to trigger the testee’s output and the visuals to support and enrich the production. This can be also the case when the videoclips are interrelated in semi-interactive conversation through short questions or even in connection to other user to make dialogues between two testees in which the potential knowledge is visible after reconstructing one’s production (Poehner, 2008). At the same time, a robust recording system and clear rubrics support the cognitivist approach that can be best seen in the long responses for descriptions or the multiple choice responses in which the students need to show evidence of knowledge without external mediation or support (what has been called current knowledge).

Usability

In terms of the usability and functioning of the applications adapted to mobile phones created for the PAULEX project, the results show that the students considered it to be useful because it enabled them to save time while taking an exam of this kind (see figure 3). It was also determined that the students learned to use the application for mobile phones faster and independently due to their familiarity through daily use of mobiles, which enabled them to adjust quickly to it and to its guided interactivity. Analysis of the data related to level of satisfaction with the use of the
application was very high when there was a sensation of predictability that leads to a fast understanding of the method and learning how to do specific tasks on the mobile phone.

![Figure 3. PAULEX application on mobiles](image)

The overall conclusion of these first trials was that the students felt comfortable with the format (bearing in mind the limitations of the devices. The oral tasks with video presentations were evaluated with the same degree of confidence and reliability as the analogous activities on the web platform for personal computers.

![Figure 4. Usability test of the PAULEX application](image)

**Proposals for the design of an m-testing platform**

The proposals in this section were also applied to the OPENPAU project and any future project of a language testing m-platform and are strongly based on the findings from the PAULEX project. The OPENPAU project has incorporated the application to the HTC Desire model mobile device whose base technology allows multimodal use of different forms of data input and output (see figure 3). To do so, a study was carried out in advance to determine the initial conditions needed for completion of the tasks on an English language skills exam. These included the
functional structure and verification of the sections in order to ensure efficient, real navigation. Afterwards, the following premises were considered vital for the m-application:

- Exam access should be possible using the keypad and/or the touch function for navigation on the mobile device, enabling access to each of the parts of the exam and including the possibility of repeating tasks as long as the user has not finalized the section.
- There should be a final screen for confirming that tasks that have been completed but it does not have to indicate whether they have been finalized by the user.
- The user should be able to select in advance the method of interaction or main form of navigation to change screens and tasks and to finalize them.
- The user should be able to see and hear any digital content annexed to the tasks created using the device's generic speaker.
- The user should be able to see and hear any digital content annexed to the tasks created using the external earphones and microphones.

The general format of the application contained the following visual sections:

- The program header area. The application name and official logo of the program participants appear in this section of the screen.
- The user data area. This area is fundamental for the final coding of the exam and student for initial correction and any future corrections, as well as any official reviews required by law at the national level.
- Area for viewing progression through the exam. This area has numbers indicating the different tasks that must be completed on the exam. This section will enable the students to know their progression throughout the exam from the point at which they enter their application access code until they send the completed exam. It starts with the reading of the student's data before the actual completion of the exam and provides information throughout completion of the exam including selection of the interaction mode, and completion of the different tasks on the exam, etc.
- Test area. This area shows the questions or exercises to be completed on each of the tasks. The content will vary depending on the functional and/or content characteristics of the exercises.
- Help area. This section will show general as well as specific information about how to complete the exam including the maximum score assigned to each section.

Project results and conclusions

As observed in the PAULEX project, in situations of high stakes tests with a large number of students, mobiles have some advantages that may put them ahead of other testing systems in terms of budget, accessibility, familiarity, and sound quality. Additionally, although the results in the PAULEX project were limited, the validation methods provided information about the ergonomics, usability, integration, and motivation of the application. According to the data obtained, it was observed that prospective research should include the following aspects:
- Task adaptation to new types of mobile phones;
- Multiplatform systems;
- User satisfaction;
- External validity as compared to other delivery systems and other tests including similar pen and paper versions;
- Technical advances in software design;
- Pedagogical benefits;
- Delivery reliability;
- Functionality.

The students were eager to use mobile phones for language teaching and learning, but they mostly wanted to use them for speaking and listening. Still, the multiple choice items for grammar were also well regarded. However, the students predictably indicated that reading and writing were too difficult to be implemented, with reading rated in a better position than writing (García Laborda, Giménez López, & Magal-Royo, 2011). The PAULEX project also showed that mobiles were excellent for test preparation and an even more encouraging finding is that they offer great opportunities for the real test itself because the students would accept using them for real testing tasks. All three teachers indirectly involved in piloting their use supported mobiles and liked the sequencing and delivery procedure for questions, but they claimed that they had no software up to that time to implement the teaching at a large scale. They also found that, although the testing system could, in fact, be valuable to assess oral skills in the PAU in the long term, phones with bigger screens were desirable. At the same time, they doubted that the Ministry of Education would spend large sums of money on the terminals. However, they believed that the listening and speaking sections could be done online while the rest of the test could be done with pen and paper in order to lower the cost. Additionally, they mentioned that one set of mobile phones could serve more than one high school and maybe more than one year given adequate hygienic measures. Finally, they mentioned the convenience for raters since they could work from a distance either on synchronous or asynchronous testing.

Our experience also determined that technologies for developing user interfaces should focus on the requirement to offer simple interaction modes that are highly natural and adapted to future terminals and communication networks (Oviatt & Cohen, 2000). It is in this area in which technologies face their biggest challenge: attempting to integrate different modes of communication (visual, oral, auditory, gestural, etc.) in order to offer new more powerful methods of interaction with the user, grouped under the name of natural or multimodal interaction, thus overcoming the limitations of interfaces available today (Oviatt, 1999). The ultimate objective of natural interaction is to enable users to be able to use all the communication resources available to them, combining multiple modes of interaction and, therefore, creating a multimodal environment for information access (voice, audio, graphics, video, keypad, electronic pencil, pointer, mouse, etc.) (Oviatt & Larson, 2003). In this sense, the OPENPAU project is currently being driven by practical concerns. The current research is now exploring the potential for implementation and the pedagogical implications while extending the domains of the project to make it a multiplatform one. The study has shown the feasibility of using mobiles for the intended purposes and that the cost could probably be lower than the traditional face-to-face interviews while also permitting a better distribution of space for delivery and adequate rationalization of testing times. Most of the students might also engage in this testing means more easily than in a face-to-face interaction with the examiner. With the development of the OPENPAU application for ubiquitous devices, it has been found that technology has now progressed sufficiently to propose the offering of exams using multimodal access. The incorporation of new modes of interaction such as voice recognition for navigation, the use of touch screens, or synchronized use of the keypad will enable users more comfortable access in accordance with their needs and, thus, solve problems related to accessibility to the media (Magal-Royo, Giménez-López, Pairy, García Laborda, & Gonzalez-Del Río, 2011; Magal-Royo & Giménez López, 2012).

Progress in the use and research of mobile phones for language learning is receiving increased attention and their use in Mobile Assisted Language Learning (MALL) is an area of steady growth. Despite the advantages this area offers users in terms of the flexibility and ubiquitous nature of the device and environment, as well as advances in mobile applications and Internet access, it must still deal with the need to seek efficient adequate interfaces for user needs for information access and transfer. In the specific case of task completion or specific processes, it is important to evaluate the impact of functional environments that enable users to find comfort and accessibility in the information provided in order to favor this mode of learning.
Future lines of work

The potential of technologies adapted for multimodal interaction in language testing offers huge possibilities for development of innovative applications. In that sense, devices will enable users to select between using one mode or another exclusively (for example, using an online dictionary or making a voice call), to the possibilities of changing between modes of interaction in the same session (sequential multimodality, as in consulting the dictionary on occasions on a mobile during a test), to true freedom in combining and changing modes (simultaneous multimodality: talking, keying, dialing, viewing, etc.) on terminals or ubiquitous devices that enable simultaneous access to voice and data channels, and thus offers opportunities for new items that resemble more what speakers do with the language and how they use it.

In reference to the project impact in the Spanish educational system, it is believed that an inexpensive system to assess speaking skills may have two potential benefits: first, it will enable testing of this skill at a low cost; and second, the impact on the classroom of implementing speaking skills may lead to a great educational improvement in foreign languages. Thus, as a whole, the expected effect of the project if used in the near future is immense and certainly very significant for the educational system.

To conclude, while the use of mobile phones for high stakes testing may be feasible, it is necessary to obtain a commitment from all the stakeholders including the students and the administration authorities. Since the oral test is a social, professional, and educational demand, delivering the oral section of the PAU through mobile phones would require adequate facilities from all the high schools, a better understanding of technology from teachers and new ways to plan and prepare for the test on the part of students. Researchers should also seek ways to overcome the difficulties associated with hearing impairment or other restrictions. While mobiles could be a great asset in education, it is necessary to recognize that not all teachers may be equally prepared to face such as a technological change or eager to change their ways of teaching to cater to the students’ needs by facilitating them with the necessary strategies for taking the test. Thus, practitioners should also receive the necessary instructions and courses to facilitate their adaptation to the new context. Nevertheless, it is believed that this change would not be any more traumatic then others that they have seen in recent years. The ongoing work in the PAULEX project is expected to continue to address these issues. The information obtained so far, while initial, provides enough evidence for the potential of this innovation in both the national and international contexts in areas such as educational planning, course design, test delivery, specifications, and information and communication technologies development. It also takes the use of mobile phones far beyond their traditional perspective of mere supportive elements of courses or learning to enhance their role as high stakes testing facilitators.

Acknowledgements

The authors would like to thank the Ministry of Economy and Competitiveness for funding the research project (with co-financing by ERDF) within the framework of the National R + D + I (2011-2014) "Guidance, proposals and teaching for English section in the entrance examination to the University" (Reference FFI2011-22442). The researchers would also thank the Spanish Ministry of Education, Culture, and Sports because without the grant for the Senior Researchers Mobility this paper would probably have not been possible.

References


Effects of Mobile Instant Messaging on Collaborative Learning Processes and Outcomes: The Case of South Korea

Hyewon Kim1*, MiYoung Lee2* and Minjeong Kim3

1Center for Teaching and Learning, Dankook University, South Korea // 2School of Education, Virginia Commonwealth University, USA // 3Department of Teaching Education, Dankook University, South Korea

*Corresponding authors

ABSTRACT

The purpose of this paper was to investigate the effects of mobile instant messaging on collaborative learning processes and outcomes. The collaborative processes were measured in terms of different types of interactions. We measured the outcomes of the collaborations through both the students’ taskwork and their teamwork. The collaborative learning processes and outcomes in the Mobile Instant Messaging group (Mobile IM) were also compared with the Personal Computer-based Instant Messaging group (PC IM) and the Bulletin Board System group (BBS). A total of 48 students participated in this study, and the main results show that more cognitive and metacognitive interactions were found in the BBS group while social and affective interactions were the major types of interactions in the Mobile IM group and the PC IM group. As a result of the collaborative learning outcomes, the Mobile IM group shows better teamwork than the other two groups. However, better taskwork was found in the BBS group and the PC IM group rather than the Mobile IM group. Finally, the researchers discuss the implications of this study from the perspective of the educational potential of mobile learning.

Keywords

Mobile-based collaborative learning, Mobile instant messaging, Collaborative learning processes, Collaborative learning outcomes

Introduction

Many researchers have claimed that mobile learning will greatly influence the future of teaching and learning in collaborative learning contexts (El-Hussein & Cronje, 2010; Huang, Yang, Huang, & Hsiao, 2010; Ryu & Parsons, 2012). The main reason behind many researchers’ enthusiasm about mobile based collaborative learning stems from its spontaneous, portable, personalized, ubiquitous and situated characteristics (Motiwalla, 2007; Patten, Arnedillo Sanchez, & Tangney, 2006; Rau, Gao, & Wu, 2008; Ryu & Parsons, 2012). Moreover, mobile learning has gradually become stable and mature (Huang, Yang, Huang, & Hsiao, 2010) and has attracted an increased number of learners in recent years.

Educators in South Korea are particularly fascinated by the concept of mobile learning due to its potential to overcome the limitations of traditional education and web-based learning. According to Korea Internet & Security Agency (2011), the infrastructure for mobile learning (e.g., WiFi networks, high-speed internet connection) is well established in South Korea. The Organization for Economic Cooperation and Development (OECD) also recently reported South Korea has the most mobile wireless broadband subscriptions of 34 OECD counties (OECD, 2012). South Korea has 104.2 subscriptions per 100 inhabitants. Additionally, several South Korean universities have distributed free iPhones or smart phones and encouraged students to utilize them to participate in lectures, to access library sources, and to access educational administration system (Lee, 2010). Students’ adoption of mobile technology is not surprising, given recent statistics on Internet usage. The Korea Internet & Security Agency (2012) finds that the internet usage rate for university students is almost 100% (99.9%) and among instant message users, 49.4% use mobile instant messaging services. The rapid diffusion and use of mobile devices suggests students may be receptive to educators’ incorporation of these tools for learning or ubiquitous learning in South Korea (Park, Nam, & Cha, 2012).

However, the true extent of the impact of mobile learning on education is still contested, both theoretically and empirically (Motiwalla, 2007, Ryu & Parsons, 2012). Moreover, previous research is limited to two specific themes – the effectiveness of mobile learning and the design of mobile learning systems (Wu, Wu, Chen, Kao, Lin, & Huang, 2012). Researchers have typically measured the effectiveness of mobile learning using learning outcomes rather than learning processes (Chen, Chang, & Wang, 2008; Hwang & Tsai, 2011). These outcomes comprise motivations, perceptions, attitudes, academic achievement, and satisfaction of students.
In this respect, various research topics that can uncover the potential of mobile learning are warranted to present more practical guidelines in this area. To address this gap in the literature, the present study explores how mobile learning affects collaborative learning processes and outcomes. Specifically, we examine the extent to which students’ cognitive, metacognitive, and social/affective interactions vary in mobile-based collaborative learning environments. We also examine the quality of cognitive messages and the level of team effectiveness in order to measure taskwork and teamwork, respectively.

Theoretical background

Mobile-based collaborative learning in social and situated learning frameworks

It is important to emphasize that the use of technology in educational settings must be in accordance with educational theories and specific pedagogical considerations (Patten et al., 2006). According to Ryu and Rarsons (2012), social and situated learning can be experienced through mobile-based collaborative learning since mobile learning facilitates seamless social interaction in learners by providing them advanced functions such as mobility and instant connectivity. Social learning theory emphasizes that learning occurs within a social context, which means people learn through observing and modeling other learners’ behaviors (Bandura, 1977; Hung, Looi, & Koh, 2004). Mobile-based collaborative learning can maximize the quality and quantity of interactions and observations through its rich communication channels. On the other hand, situated learning theory emphasizes authentic contexts and real learning activities (Lave & Wenger, 1991). Situated learning occurs in educational settings, which provide authentic contexts and activities to promote social interaction and collaboration (Herrington & Oliver, 1995; Lave & Wenger, 1991). Unlike traditional classrooms that decontextualize learners from authentic and practical situations, mobile learning provides a borderless context where learners can reach their goals and needs through real-time interactions. Thus, learners will experience enhanced social and situated learning through mobile learning. Also, mobile learning grounded in social and situated learning will provide learners with more updated learning environments.

Mobile instant messaging for collaborative learning

Collaborative learning is defined as ‘a situation in which two or more people learn or attempt to learn something together’ (Dillenbourg, 1999, p. 2). Collaborative learning can be mediated through many different tools, such as discussion boards, blogs, and instant messenger. Like computer-based collaborative learning, mobile-based collaborative learning is mainly text-based, which can enable students to express their opinions and to ask questions without the pressure or feeling of threat that can accompany traditional classrooms (Kitsantas & Chow, 2005; Rau et al., 2008; Ting, 2012). However, Chen & Huang (2010) note that computer-based collaborative learning has a limitation with respect to meeting learners’ educational needs, especially for students who want a more informal and flexible learning environment. In this respect, mobile-based collaborative learning can be more in accordance with their needs by providing ubiquitous and situated learning environments (El-Hussein & Cronje, 2010).

Instant messaging is one of the most widely-used mobile applications for education (Rau et al., 2008). Rau et al. (2008) found that mobile instant messaging supported social bonding between students and instructors. Additionally, Yengin, Karahoca, Karahoca, & Uzunboyulu, (2011) investigate the potential of using mobile instant messaging for education, and they found the successful examples such as a quiz tool, an assessment tool and discussion tools in several previous studies (e.g., Attewell, 2005; Stone, Briggs, & Smith, 2002; Markett, Sánchez, Weber, & Tangney, 2006; Bollen, Eimler, & Hoppe, 2004; Holley & Dobson, 2008). Other studies suggested that when used as a discussion tool, mobile instant messaging can promote interactivity and led to more active collaboration (Markett et al., 2006; Bollen et al., 2004; Holley & Dobson, 2008). Despite positive findings from several studies, Ryu & Parsons (2012) and El-Hussein & Cronje (2010) point out that there is still a need to conduct additional research on how mobile instant messaging could facilitate collaborative learning beyond the ‘novelty effect’ of new mobile technology.

Collaborative learning processes: Cognitive, metacognitive and social/affective interactions

Mobile-based collaborative learning supports interactions among students as well as instructor-student interactions (Ting, 2012). Students can also enjoy the increased frequency of social interaction through mobile technology in group-based projects (Seppala & Alamaki, 2003). A number of researchers emphasize the quality of cognitive
interaction in learning environments, which is crucial for the success of collaborative learning. However, many researchers note that students’ metacognitive and social/affective interactions also play a fundamental role in collaborative learning (Efklides, 2008; Salonen, Vauras, & Efklides, 2005). Metacognition is defined as knowledge about knowledge or the regulation of cognition (Brown, 1987). Metacognitive interaction is regarded as the interactive activities that monitor, evaluate and revise other team member’s cognitive processes when they work as a team. They involve the sharing of metacognitive justification, evaluation and feeling (Efklides, 2006). Social/affective interactions are an inevitable part of human communication and play an essential role in collaborative learning (Shen, Wang, & Shen, 2009). Learners express a variety of emotional states (e.g., interest, curiosity and confusion) (Kort, Reilly, & Picard, 2001) as well as social expressions (e.g., greeting, complimenting, and expressing appreciation) (Rourke & Anderson, 2002) when they work together. Furthermore, Panitz (1999) argues that it is important to create an emotional environment that enables students to take initiative in expressing their opinions about any given topic while constructing a shared learning experience.

Interestingly, Ting (2012) suggests that mobile technologies can strengthen learners’ interactions and ultimately help learners achieve better collaborative learning outcomes. In addition, Rogers & Price (2006) indicate that mobile technology can change learners’ collaborative learning processes, particularly their cognitive, metacognitive and social/affective interactions. However, it is hard to find studies that focus on these specific types of interactions, even though much research has been done on the topic of computer-based collaborative learning (Guan, Tsai & Hwang, 2004; Hara, Bonk, & Angeli, 2000). In addition, Wu et al.’s (2012) meta-analysis on mobile learning using 164 published papers from 2003 to 2010 shows that evaluating the outcomes of mobile learning rather than processes was the most researched topic in the field of mobile learning. Thus, our study, which addresses how these interactions occur in mobile-based collaborative learning environments compared to collaborative learning via desktop computer or BBS, will be valuable to practitioners as well as researchers who are interested in facilitating students’ informal or seamless learning by applying mobile technologies to education.

**Collaborative learning outcomes: Taskwork and teamwork**

Unlike individual learning, collaborative learning not only needs task-related skills but it also needs team-related skills that enable team members to work together smoothly and effectively (Eccles & Tenenbaum, 2004). Moreover, a high performance team is characterized as a group of people that is effective in creating a balance between taskwork and teamwork (Johnston, Smith-Jentsch, & Cannon-Bowers, 1997). Mathieu, Heffner, Goodwin, and Salas & Cannon-Bowers (2000) describe taskwork as the skill necessary to accomplish a given task. Taskwork is identified by a learner’s cognitive activity. On the other hand, teamwork is described as the skills needed for effective team functioning such as proper role assignment/responsibility, using efficient communication channels and accurate decision making.

Although many researchers argue that teams develop both taskwork and teamwork through performing their team projects, the evaluation of collaborative learning tends to only focus on their task achievement in terms of how effectively and efficiently they accomplish their given tasks (Mathieu et al, 2000). However, Stott and Walter (1995) indicated that taskwork and teamwork are conceptually independent, but the nature of their functioning is intertwined and affects team performance. Therefore, it is more reasonable to measure both taskwork and teamwork as outcomes of collaborative learning instead of measuring taskwork by itself.

**Research questions**

To examine the extent to which learners’ cognitive, metacognitive and social/affective interactions vary in mobile based collaborative learning as well as the effects of mobile learning on collaborative learning outcomes in terms of taskwork and teamwork, the specific research questions are as follows.

First, are there any differences in collaborative processes in terms of learners’ three types of interactions when they use Mobile Instant Messaging in comparison to Personal Computer-based Instant Messaging and Bulletin Board Systems?

Second, are there any significant differences in collaborative outcomes in terms of learners’ taskwork and teamwork when they use Mobile Instant Messaging in comparison to Personal Computer-based Instant Messaging and Bulletin Board Systems?
Third, are there any differences in learners’ perceptions when they use Mobile Instant Messaging in comparison to Personal Computer-based Instant Messaging and Bulletin Board Systems?

**Method**

**Participants**

A total of 48 students in three classes from a large private Korean university participated in the study. All participants were enrolled in an introductory educational technology course which was a required course. Their average age was 21.57 (SD = 13). They participated in the study as part of their regular class activity. The three classes were randomly assigned to one of the following three groups: a mobile instant messaging group (Mobile IM; n = 22), a personal computer-based instant messaging group (PC IM; n = 12), and a bulletin board system group (BBS; n = 14).

**Three communication media for discussion**

Mobile Instant Messaging (Mobile IM): The Mobile IM group used the KakaoTalk application to conduct their discussion task. It is one of the most popular free mobile messenger applications in South Korea. It provides free text messaging and free calls. The students in the Mobile IM group can share various content and information such as photos, videos, and URL links. Group discussion is possible without the constraints of time and space.

Personal Computer based Instant Messaging (PC IM): The PC IM group used MSN Messenger in their desktop computers. The MSN Messenger is a form of communication over the internet on a PC that offers a quick transmission of text-based messages from sender to receiver. Computer instant messaging basically offers real-time online chat but students need to set a time and to log into the messenger for their group discussion.

Bulletin Board System (BBS): The BBS group used a discussion board system like Blackboard provided by the Learning Management System in a University. Through the BBS, students can do an asynchronous discussion while students can do synchronous discussion through Mobile IM or PC IM. Students in the BBS group are able to revisit their discussion board and post their message whenever they want.

To fairly compare the differences in the three communication media groups, both mobile and computer instant messaging groups are allowed to use only text-based messaging even though they can use voice chatting through their devices. Also, the BBS groups are only allowed to use the discussion board through their personal computer even though they can access it through mobile technology.

**Task and procedure**

The team task was an ill-structured problem describing a novice teacher who took on a very low achievement class with many troublemakers, and a school principal who directed her to increase student academic achievement within a year. Before students could solve the ill-structured problem, lessons on learning paradigms such as behaviorism, cognitivism, and constructivism were provided to the participants in a regular class. Then, they were randomly assigned to one of three communication media groups. Each group consisted of three or four students and they were asked to discuss a best solution to solve the given problem based on three learning paradigms within a week. All participants were required to discuss the topic using only an assigned communication medium. After the discussion week, participants answered an open-ended perception question which asked them what were the most and the least favorite aspects of the medium that they used for their discussion.

**Measures**

To examine the effects of Mobile Instant Messaging on collaborative learning, students’ interactions were measured as the learning processes of their collaborations, and taskwork and teamwork were measured as collaborative learning outcomes. The specific methods were described as follows.
Three types of interactions

The content analysis method was used to analyze the types of interactions. As Henri (1992) suggested, an individual theme or idea (thematic unit) was used as the unit of the analysis rather than a word, sentence or paragraph in order to maintain consistency in analyzing students’ discussion messages that occurred in the three different media. For example, the Mobile IM group expressed their opinions in a short phrase or word instead of using a full sentence (e.g., “when?”, “in this case”) while the BBS group usually posted at least one paragraph to state their idea. Therefore, the individual theme or idea was used as the unit of analysis in this study, so the unit of analysis can be any size text from a single word to a paragraph as long as it expresses a theme or idea.

The types of interactions are composed of three categories: cognitive or metacognitive interaction, social or affective interaction, and other interaction. Cognitive or metacognitive interaction is a task-related meaning unit. Social or affective interaction is a non-task-related meaning unit such as personal talks or the expression of feelings. Other interactions are interactions about managing the discussion such as scheduling for the task and setting discussion rules. Two researchers then developed a coding scheme and classified each thematic unit into one or more of the aforementioned categories. The coding scheme is described in Table 1 with samples of thematic units. Inter-rater reliability for the classification of categorical variables was determined by Cohen’s Kappa, which measures the agreement between two raters who each classify thematic items into mutually exclusive categories. Cohen’s Kappa for the inter-rater reliability was 0.96 for the agreement of thematic unit and 0.94 for the classification of interaction. The two raters discussed until they reached a consensus, and a total of 1,850 messages were analyzed in this study.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive/metacognitive</td>
<td>Talking about key concepts of learning theories</td>
</tr>
<tr>
<td>Interaction</td>
<td>Talking about learning theories’ principles</td>
</tr>
<tr>
<td></td>
<td>Talking about implication of learning theories</td>
</tr>
<tr>
<td></td>
<td>Sharing learner’s opinion on learning theories and application</td>
</tr>
<tr>
<td></td>
<td>Speculating some issues on learning theories</td>
</tr>
<tr>
<td></td>
<td>Questioning something about learning theories</td>
</tr>
<tr>
<td></td>
<td>Summarizing what they discussed on learning theories</td>
</tr>
<tr>
<td></td>
<td>Reflecting what they discussed on learning theories</td>
</tr>
<tr>
<td>Social/affective interaction</td>
<td>Praising the other student’s utterances</td>
</tr>
<tr>
<td></td>
<td>Chatting about student’s private lives</td>
</tr>
<tr>
<td></td>
<td>Chatting about non-task-related topics</td>
</tr>
<tr>
<td>Other interaction</td>
<td>Talking about scheduling for the task</td>
</tr>
<tr>
<td></td>
<td>Talking about taking turns</td>
</tr>
<tr>
<td></td>
<td>Talking about setting discussion rules</td>
</tr>
</tbody>
</table>

Teamwork

A survey was used in order to measure teamwork. The survey consisted of five questions about team effectiveness: Efficiency of team management, Observance of team schedule, Conviction of team output quality, Adequacy of team output quantity, and Satisfaction with team output (e.g., “Our team management was efficient,” “Our team members kept our team schedule,” “we think the quality of team output was excellent,” “we think the quantity of team output was appropriate,” and “we are satisfied with our team output.”). Students responded on a five-point Likert scale ranging from “Strongly Agree” to “Strongly Disagree,” depending on how well they thought that the statement described their team effectiveness. The responses were coded in the following manner: strongly agree = 5, agree = 4, not sure = 3, disagree = 2, strongly disagree = 1. The reliability of the survey was .80 for the pilot test and for this study it was .78.

Taskwork

To examine how well learners discussed a given topic, we evaluated the quality of their group discussion. Specifically, we measured their cognitive messages based on four criteria: novelty, importance, relevance, and ambiguity. Among the 10 criteria in Newman, Webb, & Cochrane’s (1996) study, four criteria which measure the
quality of cognitive messages were selected for this study. Two researchers who specialized in educational technology scored each cognitive message as 1 or 0 based on the four criteria described in Table 2. The inter-rater reliability through Cronbach alpha analysis was 0.92.

Table 2. Four criteria for rating taskwork

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Descriptions</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novelty</td>
<td>New information, ideas, solutions</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Repeating what has been said</td>
<td>0</td>
</tr>
<tr>
<td>Importance</td>
<td>Important points/issues</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Unimportant, trivial points/issues</td>
<td>0</td>
</tr>
<tr>
<td>Relevance</td>
<td>Relevant statements</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Irrelevant statements, diversions</td>
<td>0</td>
</tr>
<tr>
<td>Ambiguities</td>
<td>Clear, unambiguous statements</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Confused statements</td>
<td>0</td>
</tr>
</tbody>
</table>

Perception on communication media

For a more in-depth understanding of the characteristics of each communication medium, students’ perceptions on medium were measured by a survey which contained one open-ended question. The question asked students what was the most and the least favorite aspects of the medium in their discussion. One piece of paper was given to each student, and they described their thoughts about the given communication medium for 30 minutes.

Data analyses

Content analysis was conducted to examine how the types of interactions were different across the three groups. For the analysis of the comparisons of the three groups in terms of taskwork and teamwork, one-way ANOVAs were conducted. The perception survey data was analyzed qualitatively based on the main themes that students addressed as characteristics of the communication medium they used.

Results

Collaborative process: Types of interactions

This study was designed to discover if there are differences in the types of interactions such as: cognitive or metacognitive interactions, social or affective interactions and other interactions (not included in the two categories) among the three communication media groups. The interactions were analyzed by a content analysis and the percentage of each interaction compared to the total number of messages from each group was discerned. The results are shown in Table 3.

Table 3. Frequencies of Interaction Types by Communication Media Groups

<table>
<thead>
<tr>
<th></th>
<th>Cognitive/metacognitive interaction</th>
<th>Social/affective interaction</th>
<th>Other interaction</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile IM</td>
<td>614(50.00%)</td>
<td>449(36.56%)</td>
<td>165(13.44%)</td>
<td>1,228(100.0%)</td>
</tr>
<tr>
<td>PC IM</td>
<td>205(45.15%)</td>
<td>166(36.56%)</td>
<td>83(18.28%)</td>
<td>454(100.0%)</td>
</tr>
<tr>
<td>BBS</td>
<td>123(73.21%)</td>
<td>32(19.05%)</td>
<td>13(7.74%)</td>
<td>168(100.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>942(50.92%)</td>
<td>647(34.97%)</td>
<td>261(14.11%)</td>
<td>1,850(100.0%)</td>
</tr>
</tbody>
</table>

From the results, it was found that cognitive/metacognitive interaction accounted for approximately 50%, social/affective interaction 37%, and other interaction 13% in the group utilizing Mobile IM in their discussion. A similar tendency was found in the group utilizing PC IM with results that showed that cognitive/metacognitive interaction accounted for approximately 45% of the total number of messages, social/affective interaction 37%, and other interaction 18%. On the other hand, cognitive/metacognitive interaction accounted for more than 73%,
social/affective interaction 19%, and other interaction 8% in the group utilizing BBS for interaction.

In terms of the three types of interactions, the Mobile IM and PC IM groups showed similar results. However, the result reveals that the BBS group had more cognitive/metacognitive interactions and fewer social/affective interactions compared to the Mobile IM and PC IM groups. In addition, other interaction was also lower in the BBS group compared to the other two groups. The results are arranged into a pie chart as follows.

![Pie chart showing interaction types by communication media]

**Figure 1. Distribution of interaction types by communication media**

**Collaborative outcomes**

**Teamwork**

To identify if there are differences in teamwork scores across the three different communication media groups, the mean of the teamwork score was calculated as shown in Table 4. A one way ANOVA analysis was conducted to see if there were any statistically significant differences in the teamwork scores among the groups. The results are presented in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile IM</td>
<td>22</td>
<td>4.12</td>
<td>.65</td>
</tr>
<tr>
<td>PC IM</td>
<td>12</td>
<td>3.47</td>
<td>.41</td>
</tr>
<tr>
<td>BBS</td>
<td>14</td>
<td>3.65</td>
<td>.58</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>3.82</td>
<td>.63</td>
</tr>
</tbody>
</table>

From the results above, the Mobile IM group recorded the highest average points with 4.12 in the teamwork score, followed by the BBS group with 3.65, and the PC IM group with 3.47. As shown in Table 5, there were significant differences in teamwork scores between groups according to the type of media. From the results of the Scheffe verification, there was a significant difference between Mobile IM and PC IM at the p < .01 level, however, significant differences were not found between the BBS group and the other two groups (Mobile IM and PC IM). That is, the result reveals that the Mobile IM group showed higher teamwork at a statistically significant level compared to the PC IM group.

<table>
<thead>
<tr>
<th></th>
<th>Sum of Square</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3.929</td>
<td>2</td>
<td>1.965</td>
<td>5.803</td>
<td>.006</td>
</tr>
<tr>
<td>Within-groups</td>
<td>15.235</td>
<td>45</td>
<td>.339</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19.164</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Taskwork

To identify if there was any difference in the taskwork score depending on the type of communication media, the total taskwork score was divided by the total number of cognitive/metacognitive messages in each group of communication media. The results revealed that the mean scores of taskwork were 1.89, 2.59 and 2.62 in the Mobile IM, PC IM and BBS groups respectively. The highest mean of taskwork score was found in the BBS group, and the lowest mean of taskwork score was found in the Mobile IM group.

<table>
<thead>
<tr>
<th>N</th>
<th>Total # of cognitive/metacognitive messages</th>
<th>Total taskwork score</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile IM</td>
<td>22</td>
<td>614</td>
<td>1,160</td>
<td>1.89</td>
</tr>
<tr>
<td>PC IM</td>
<td>12</td>
<td>205</td>
<td>530</td>
<td>2.59</td>
</tr>
<tr>
<td>BBS</td>
<td>14</td>
<td>123</td>
<td>322</td>
<td>2.62</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>942</td>
<td>2,006</td>
<td>2.13</td>
</tr>
</tbody>
</table>

The ANOVA analysis was conducted to determine if the taskwork score was statistically different across the three communication media groups. The results are shown in the following Table 7.

<table>
<thead>
<tr>
<th></th>
<th>Sum of Square</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>109.640</td>
<td>2</td>
<td>54.820</td>
<td>55.370</td>
<td>.000</td>
</tr>
<tr>
<td>Within-groups</td>
<td>938.589</td>
<td>948</td>
<td>.990</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1048.229</td>
<td>950</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As shown in Table 7, there were statistically significant differences in taskwork scores between groups. To identify which groups showed the difference, a Scheffe verification was conducted, and its results revealed that there was no difference between the PC IM group and the BBS group, but significant difference was found in the Mobile IM compared to the PC IM and BBS groups at the p < .001 level. That is, the taskwork score of the Mobile IM group was significantly lower than the other two groups.

Perception on communication media

Based on the results of the student perception survey, the following seven major themes were perceived to be the most or the least favorable aspects of each communication medium for their discussion.

**Theme 1 - Time Constraint:** Mobile IM enabled students to contact their team members whenever they needed them. Most of the students using mobile devices said they liked that it did not have constraining time features of Mobile IM.

“I think our team discussed the topic all day long because we talked whenever we are available. Even though it’s short time... So, I love it. One day, I had a lot of classes but I could read the members’ opinion between the classes and respond them. So, I could keep up with my team members’ discussion.”

However, the students in the PC IM group pointed out that it was difficult to find an available time for all team members to log into the PC IM, even though they did not need to be in the same place.

“It is hard to find the time which all the members can participate because of the differences time schedules. Some of the students wanted to night time and the others didn’t. A student was available on weekend only. It was so hard!!”

**Theme 2 - Limitation of Location:** The students in the Mobile IM group could participate in their group discussion while they were working or moving to another place. Many students in the Mobile IM group pointed out that this was one of the most favorable aspects of the technology.
“I am working at a café as a part time job. Sometimes when there were a few customers, I can respond to the team members’ opinion. It’s the most strong point of mobile IM because if I use BBS or PC IM, I cannot involve the discussion frequently.”

However, some of the students in the PC IM and BBS groups who used PC computers were limited by their location. Students could access chat rooms or discussion boards only where there was a computer available.

“It is hard to get the chance of using computer at the school because the number of the computer in the lab is not enough. So I can access the discussion board at home... sometimes when I came back home lately or I was so tired, I didn’t want to participate the discussion. It bothered me....”

Theme 3 - Availability for Searching Resources during Discussion: The PC IM group and BBS group reported that they could search the Internet to find any necessary resources related to their discussion topic while they were communicating. They could refer to any references and write their opinions for as long as they wanted without any sudden interruptions by other students. On the other hand, most of the students of Mobile IM felt it was inconvenient to locate necessary references such as textbooks or articles, especially when they participated in their discussion outside of the home or classroom.

“Sometimes I could not remember the detail of a theory which I learned. I really wanted to back up my opinion.... In that case, I want to search the Internet but it is little bit inconvenient through mobile chat. In a mobile chat, multitasking is exceedingly cumbrous.”

Theme 4 - Emotional Closeness: Most of the students in the Mobile IM group claimed that Mobile IM offered a more comfortable and friendly environment where they could talk about private topics as well as their discussion topic using various emotional and social expressions. However BBS and PC IM group students rarely addressed this point as an advantage.

“I usually start to say ‘Hello’ or ‘The weather is great’..... like my team members are besides me. Or sometimes I complain my headache or a lot of papers of the other classes to my team members. I feel free to talk to them about my private stories through mobile chat.”

Theme 5 - Chance for Careful Thought and Reflection: Most of the students in the BBS group commented that the BBS enabled them to post their opinions or responses after having enough time to think about the given discussion topic and to review other team members’ postings. Also, Mobile IM students had enough time to review other team members’ messages and provide thoughtful feedback as compared to offline discussion. On the other hand, some of the team members only focused on typing their opinion without reviewing or considering other members’ postings. Therefore, it was difficult for them to have a more convergent discussion in the PC and mobile IM environments. That point was addressed as a disadvantage of mobile and PC IM groups.

“Some members talk so long when we discuss at offline meeting. In that case, I am sure that I forget what he or she was talking about for the first time. However, in the mobile chat, I can review the full message, so I do not forget what I prepared for the comments”

Theme 6 - Participation of the members: The students in the BBS group reported that some team members did not frequently visit their discussion board, which resulted in delayed responses and disjointed group discussions.

“A group member didn’t visit the discussion board after his first posting. That was it... I was annoyed that we could not discuss anymore. However, there wasn’t any alternative because we had to discuss through BBS only.”

Theme 7 - Inconvenience of Using Communication Media: The most common problem voiced by Mobile IM students was that the relatively small keyboard and screen on their mobile phones constrained them when typing a lengthy opinion or response during their discussion.

“I hate typo but the keyboard of my mobile phone is so small. So, I cannot help mis-typing. Also, the screen is so small. When I read all the discussion, I have to drag the message for a long time. It so irritates me.”
Conclusion and suggestion

Our study contributes by extending the scope of research on mobile learning. Unlike previous research, our study focuses on the effects of mobile learning on collaborative learning processes and outcomes. Social and affective interactions as well as cognitive and metacognitive interactions were also considered as important factors in the collaborative learning processes. Moreover, teamwork that was often ignored as the outcome of collaboration was measured along with taskwork. Based on the results of the study, it is recommended that students use Mobile IM or PC IM in order to facilitate their social and affective interaction at the beginning of their team project when they need to invest in getting to know one another (Lee & Johnson, 2008). Once students have progressed beyond the initial stages of the project, BBS could be the best communication medium to promote students’ cognitive and metacognitive interaction. The results of this study suggest that BBS, PC IM and Mobile IM should be used for different purposes. The BBS and PC IM would be good communication media to improve students’ taskwork while the Mobile IM would be the best choice to facilitate their teamwork. Therefore, understanding the unique characteristics of each communication medium is pivotal to maximize the quality of instruction, and, ultimately, students’ performance.

Future studies are suggested in the following three directions. First, it would be interesting to examine the affective and social aspects of learning as the result of collaborative learning outcomes. In this study, we measured learners’ taskwork and teamwork by focusing primarily on their cognitive development and team effectiveness. However, it will be necessary to examine how much their motivation and attitudes are improved after using Mobile IM for their collaborative learning. Second, we measured learners’ interactions and outcomes that occurred in a one week discussion, but future study is needed to conduct the measurement at least three times to see the change in learner interaction patterns and how their teamwork and taskwork develop over time. According to Fiore, Salas, Cuevas, and Bowers (2003), a team as a cognitive community goes through three coordination phases consisting of pre-process, in-process, and post-process. Depending on each phase of the processes, learners’ interactions and their focus vary. Therefore, it would be a good research topic to examine the change in learner interaction patterns and the development of teamwork and taskwork along with the three coordination phases. Third, from a more practical standpoint, future study also needs to focus on the design of mobile based collaborative learning environments with consideration of the results of this study and provides specific guidelines for effectively and efficiently launching mobile-based collaborative learning in online and offline classrooms. Specifically, it would be interesting to design and develop online instructions using a combination of mobile IM and other online communication tools depending on types of team activities and expected interactions for students’ informal or seamless learning.

References


Mobile Inquiry Learning in Sweden: Development Insights on Interoperability, Extensibility and Sustainability of the LETS GO Software System

Bahtijar Vogel*, Arianit Kurti, Marcelo Milrad, Emil Johansson and Maximilian Müller

Department of Media Technology, Linnaeus University // bahtijar.vogel@lnu.se // arianit.kurti@lnu.se // marcelo.milrad@lnu.se // emil.johansson@lnu.se // maximilian.muller@lnu.se

*Corresponding author

ABSTRACT

This paper presents the overall lifecycle and evolution of a software system we have developed in relation to the Learning Ecology through Science with Global Outcomes (LETS GO) research project. One of the aims of the project is to support “open inquiry learning” using mobile science collaboratories that provide open software tools and resources, and participation frameworks for learner project collaboration, mobile data and media capture, publishing, analysis, and reflection. The primary focus of this paper is to report on our technical development, insights and knowledge gained during the past four years. Technical implementations and the prototypes developed in this project have been tested across several educational trials conducted in Sweden and abroad with more than 400 learners. Insights and knowledge gained from these activities verify that learners’ requirements were adequately addressed while satisfying their needs. The outcomes and results of our efforts provided us with a better understanding with regard to which software engineering processes and approaches can be used to address and support the complex requirements that emerge in novel mobile learning scenarios. Thus, the results discussed in this paper provide deeper insights into the importance of properly addressing issues related to interoperability and extensibility in order to develop software solutions to support mobile learning that are sustainable and endurable over time.

Keywords

Mobile learning, Inquiry-based learning, Software lifecycle, User-centered development, Interoperability, Extensibility, Sustainability

Introduction

Web technologies are enabling Internet applications and services to become easily integrated in interactive systems (Holmberg, Wuenzsche, & Tempero, 2006). Thus, the web is gradually becoming a “central computer” that helps to connect diverse computing and data resources and people (Liang, Croitoru, & Tao, 2005; Giusto, Iera, Morabito, & Atzori, 2010). The evolution of these web developments combined with sensor and interactive technologies provide new possibilities for the implementation and deployment of software applications to support a wide variety of human activities.

Mobile and web technologies and applications provide new possibilities for augmenting learning activities. These are Technology-Enhanced Learning (TEL) activities that can be spatially distributed and can incorporate different physical and environmental sensor data (Wu, Yang, Hwang, & Chu, 2008). There are different mobile, web and sensor-based technologies that provide new perspectives on how learning activities can be embedded in different settings and across contexts (Chang, Wang, & Lin, 2009). One innovative aspect of these new learning landscapes is the combination of learning activities to be conducted across different educational contexts such as schools, nature and science centres/museums, parks, and field trips (Kukulska-Hulme, Sharples, Milrad, Arnedillo-Sanchez, & Vavoula, 2009). In these technology rich and dynamic learning environments learners make use of a wide range of devices and applications and the notions of system interoperability and extensibility, become central in order to successfully fulfill the requirements posed by the different educational activities and the learners. Especially, since these aspects directly influence learners’ satisfaction and experiences with regard to the applications and systems they use and also directly affect how tools and applications are adopted, appropriated and sustained over time.

In our Learning Ecology with Technologies from Science for Global Outcomes (LETS GO) collaborative international project (2008-2012), we have been developing, implementing, studying and scaling up novel ways for fostering secondary school student learning in teams for ecological and environmental sciences (Spikol, Milrad, Maldonado, & Pea, 2009). During the last 4 years we have been working with the design, development and implementation of web and mobile services that integrate geo-sensing, multimedia communication and interactive visualization techniques in specific ecology learning scenarios. Our goal has been to create mobile science inquiry...
collaboratories (Pea, Milrad, Maldonado, Vogel, Kurti, & Spikol, 2012) with teachers, learners and developers and domain scientist on topics related to water and soil quality, ecosystems and biodiversity.

One of the main objectives of the LETS GO project was to develop a robust software system including a wide range of applications and services to support educational activities that promote collaborative scientific inquiry as students formulate questions and hypotheses, and collect, analyze, discuss and compare data while studying problem topics in environmental sciences. All our software solutions have been conceived having in mind how to support all these processes. In this paper we present and discuss the overall lifecycle and evolution of the software system we have developed during the last four years. Our choice to focus on these specific aspects is guided by the challenge of how to address those problems related to the scalability and interoperability of mobile learning applications. Thus, the main question we are trying to answer in this paper can be formulated as following: How can software engineering processes support the functional requirements posed by current mobile learning scenarios and applications? The insights and knowledge gained during our research efforts are closely related to the issues of interoperability, flexibility and extensibility of the LETS GO software system and were identified during iterative user-centred development cycles. Developing sustainable mobile applications that can cope with the changing demands of dynamic learning environments requires new knowledge and approaches. The results presented and discussed in this paper provide some new perspectives in this direction.

The remaining of the paper follows with a presentation of the motivation behind our research efforts, as well as the initial requirements, to continue after with an overview of the LETS GO project and its related activities. The following section presents the details of our design, technical solution and the evolutionary stages that were carried out as a part of this development to continue with the Lessons Learned section where we reflect upon the activities we have conducted during these four years of development. At the end, we provide our main conclusions and discuss possible lines for future research.

Motivation and initial requirements

The initial requirements elicitations with stakeholders emerged from a workshop with teachers involved in the LETS GO project that took place in the fall 2008. Different activities in this workshop helped to identify the need to integrate geo-location and environmental sensing, visualization, and Web 2.0 mashup technologies, as part of a broader educational scenario. These requirements identified the need to support “open inquiry learning” for having access to diverse sensor data, live mapping tools, interactive data visualization and collaboration tools, and additional learning resources. Another requirements related to usability, include low cost, using open standards, multiple application support, and support for different types of collaboration modes and contexts (Spikol, Milrad, Maldonado, & Pea, 2009).

Trying to match these initial requirements brought up a number of challenges that concern software tools for supporting inquiry-learning activities. A survey of the literature and existing approaches to support inquiry science learning conducted at the beginning of the project indicated that there were no existing software solutions that could cope with all these requirements at the same time (Vogel, Spikol, Kurti, & Milrad, 2010). More recently, Sun & Looi (2013) report on a review of different web-based science learning environments for collaborative inquiry. The analysis of the results indicate that even those systems discussed in their paper do not cope with the kind of requirements we are addressing in our work. Already at the early stage of our project, those aspects related to the issues of interoperability and extensibility of the system to be developed were identified as one of the central challenges in terms of “building new technologies or further developing existing technologies to create novel possibilities for supporting human activities” (Tchounikine, 2011). Some of the processes that learners need to be actively involved during inquiry learning activities are to problematize, demand, discover and refine, and apply new knowledge and skills to solve complex problems (Edelson, Gordin, & Pea, 1999). Therefore, our primary focus in this project was to facilitate the integration of proper tools (both hardware and software) and services for supporting inquiry based learning activities.

According to Knapp and Barrie (2001), field trips are important to effectively learn about environmental science and they should be actively promoted. It is suggested that field trips can be helpful to generate relevancy to classroom learning when connected with the outdoor environment. For students, such an approach may raise the interest in and aspirations for science-related careers (Rudmann, 1994). The data collected in such field trips play an important role
for analysis, and hence should be saved and carried back to the classroom. Presenting and analyzing these data using visualization tools may help to increase learners’ understanding of complex subject matter.

Reflecting upon our current knowledge and experiences from the field of TEL, two important issues can be identified for supporting environmental inquiry science learning including outdoors and in classroom activities:

1. Providing technological support (in terms of portable instruments and sensors for data collection and software) for field trips activities that include collecting data, and
2. Providing technological support for classroom activities that include visualizing, exploring, analyzing, discussing and reflecting upon the data collected in the field.

Hence, the system support for these kinds of activities needed to include functionalities for mobile data collection and web-based tools and applications for interactive visualizations. These aspects were also in line with the theoretical aspects of scientific inquiry thinking that suggest that this kind of system support has the potential to increase learners’ engagement and curiosity (Pea, 2002). Thus, based on these different requirements we have developed a variety of software tools and solutions to address these different challenges. A detailed description of this work can be found at Vogel et al. (2010), Vogel et al. (2011) and Vogel (2012).

Figure 1 below provides an initial overview of the educational settings related to the LETS GO project and presents some of the initial requirements. Moreover, this figure maps the key processes of inquiry learning activities related to data collection and interpretation, exploration and reflection, drawing conclusions and communicating the results (Edelson, Gordin, & Pea, 1999; Linn, & Eylon, 2011).
cloud-based and decentralized. From a pure technological perspective, Hoppe (2009) claims that one of the main challenges we are facing involves the need for integration of diverse technological resources in broader educational scenarios. Therefore, these trends once more reaffirm that the issues related to interoperability and extensibility become central for the integration of diverse technological resources for supporting educational activities. In the coming section we provide an overview of the LETS GO activity flow in order to better understand the interplay between the different learning activities and the technological support.

**LETS GO activities and testing**

During the four years of the project more than 400 students have been involved in different type of learning activities. These activities included classroom lessons, field trips and lab work and included data collection in the field, taking images and notes, as well as data visualization and discussions in the classroom (see figure 2).

Usually, the participants in the different activities were either students from K-12 schools in Växjö, Sweden or undergraduate students (the teacher training program at Linnaeus University). As part of the environmental science curriculum, they investigated topics related to soil quality (woodland ecology) and water quality in the surrounding lakes. None of the participants in all these activities had prior knowledge regarding how to use the technologies we developed. In our latest pilot activity that took place at the Potomac River in the USA (September 2012), teachers from both Sweden and USA were involved.

![Figure 2. Different learning activities and the technologies in use](image)

Figure 3 below gives an overview of the learning activity flow and how the different phases of the students’ inquiry process were supported. It should be noticed that these activities where designed according to the different stages of inquiry based-learning as suggested by Edelson, Gordin, & Pea (1999) & Linn & Eylon (2011). Furthermore, these different learning activities have been integrated with their regular curricula in Environmental Science courses at the different schools. A typical LETS GO learning activity usually included workshops for the students to get familiarized with specific subject matter and central concepts and ideas associated with the inquiry learning process. These activities usually comprised six to eight lessons over a period of five weeks starting with the introduction of the inquiry process where basic concepts of the activity were introduced; students discussed the initial questions given to them about a specific topic (e.g., water quality). This activity was followed by the preparation for investigation and experiments to be conducted using different technologies (proves, data loggers, mobile applications for data collection in the classroom). Additionally, users conduct field experiments at a local environment and collect samples for lab analysis (see Figure 2a). The data collected using the mobile data collection tool were geo-tagged content and sensor data (usually pH, dissolved oxygen, temperature, conductivity, moisture, etc. depending on the type of the activity). The learning activity usually ended with a discussion about their findings from the field and lab work and an overall class discussion and reflection by using the web visualization tool (see Figure 2b), which tailored different geo-tagged sensor data and digital content collected using mobile data collection tool. In average, (depending of their course schedule) the entire activity was conducted over the period of four weeks across five lessons units. The logistics, as well as the time period of these activities are illustrated in figure 3 below. As it is presented, each one of these lessons units generated a set of functional requirements that the system should support, namely: sensor support, mobile data collection and interactive data visualization. Furthermore, since these activities (mainly for K-12 students) were part of the regular curricula, students were asked to submit short reports on the outcomes of their efforts and reflections after each unit, concluding with a final test at the end of the activity.
Throughout these activities, we videotaped different sessions for later analysis and some of the researchers from our
group used a systematic observation sheet during field and lab sessions.

Figure 3. Overview of learning inquiry activity flow, the technological support and the learning outcomes

During the lifetime of the project, we actively tested all our developments with school students where we combined
classroom and field trips activities. The user trials (prototype testing) allowed testing the software application
throughout five development iterations on authentic settings and dynamically changing environments, while new
requirements continuously emerged in these activities. These iterations include the release of a prototype and its
active testing with the users/learners. The initial two prototypes have been of throwaway type with a single iteration
stage each. The last prototype has been of evolutionary nature that evolved through three following iterations cycles.
Details regarding these prototypes as well as development iterations are presented in the Table 1.

Table 1 below, provides a detailed overview of the LETS GO field activities conducted since May 2009 until our
latest activity conducted in September 2012. This table provides an overview about learner generated content, and
the records that were stored in our repositories. It provides a summarized view on how many samples/records
learners collected during these activities, as well as number of pictures they stored in our database and server
resources. Furthermore, it also provides a rough overview about the number of users that used our software system
so far, including information about location of the trials, school they belonged, and type of the activity they were
engaged.

Table 1. LETS GO Activities and related data

<table>
<thead>
<tr>
<th>Prototypes</th>
<th>Development</th>
<th>Deployment and Testing</th>
<th>Organisations</th>
<th>= Users</th>
<th>Location</th>
<th>= Records in Database</th>
<th>= Images</th>
<th>Activity</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Prototype</td>
<td>1st Iteration</td>
<td>2009-May</td>
<td>Katedralskolkan</td>
<td>14</td>
<td>Katedralskolkan, Vaxjo, Sweden</td>
<td>24</td>
<td>48</td>
<td>1 indoor</td>
<td>soil quality</td>
</tr>
<tr>
<td>2nd Prototype</td>
<td>2nd Iteration</td>
<td>2010-Feb-March</td>
<td>Katedralskolkan</td>
<td>15</td>
<td>Katedralskolkan and Vaxjo, Sweden</td>
<td>30</td>
<td>60</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>3rd Prototype</td>
<td>3rd Iteration</td>
<td>2010-April</td>
<td>Kronoberg Skola</td>
<td>6</td>
<td>Kronoberg Skola, Vaxjo</td>
<td>139</td>
<td>32</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>4th Prototype</td>
<td>4th Iteration</td>
<td>2010-May</td>
<td>Kronoberg Skola</td>
<td>6</td>
<td>Kronoberg Skola, Vaxjo</td>
<td>139</td>
<td>32</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>5th Prototype</td>
<td>5th Iteration</td>
<td>2010-June</td>
<td>Trelleborg Skolan</td>
<td>25</td>
<td>Trelleborg Skolan, Vaxjo, Sweden</td>
<td>97</td>
<td>192</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>6th Prototype</td>
<td>6th Iteration</td>
<td>2010-June</td>
<td>Trelleborg Skolan</td>
<td>25</td>
<td>Trelleborg Skolan, Vaxjo, Sweden</td>
<td>97</td>
<td>192</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>7th Prototype</td>
<td>7th Iteration</td>
<td>2010-September</td>
<td>Teacher Students EU</td>
<td>12</td>
<td>Teacher Students EU</td>
<td>83</td>
<td>192</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>8th Prototype</td>
<td>8th Iteration</td>
<td>2010-October</td>
<td>Teachers</td>
<td>20</td>
<td>Teachers</td>
<td>83</td>
<td>192</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>9th Prototype</td>
<td>9th Iteration</td>
<td>2010-November</td>
<td>Vaxjo</td>
<td>20</td>
<td>Vaxjo, near Katedralskolkan</td>
<td>83</td>
<td>192</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>10th Prototype</td>
<td>10th Iteration</td>
<td>2011-May</td>
<td>Vaxjo</td>
<td>83</td>
<td>Vaxjo</td>
<td>83</td>
<td>192</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>11th Prototype</td>
<td>11th Iteration</td>
<td>2011-June</td>
<td>Vaxjo</td>
<td>83</td>
<td>Vaxjo</td>
<td>83</td>
<td>192</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>12th Prototype</td>
<td>12th Iteration</td>
<td>2011-July</td>
<td>Vaxjo</td>
<td>83</td>
<td>Vaxjo</td>
<td>83</td>
<td>192</td>
<td>1 indoor</td>
<td>water quality</td>
</tr>
<tr>
<td>13th Prototype</td>
<td>13th Iteration</td>
<td>2011-September</td>
<td>National Geographic Society</td>
<td>15</td>
<td>National Geographic Society</td>
<td>15</td>
<td>50</td>
<td>1 outdoor</td>
<td>water quality</td>
</tr>
</tbody>
</table>

= Total Users: 579
= Total Images: 1386
The evolution of our software system

The LETS GO software system has gone through an evolutionary prototyping approach through the past four years in order to become a stable and robust platform for mobile data collection, aggregation and data visualization. During the prototyping efforts, we extensively used two different development principles: evolutionary prototyping and throwaway prototyping (Sharp, Rogers, & Preece, 2007). The iterative user-centred development cycles of the software system are mapped to our developments, following these two principles of prototyping:

**Throwaway prototyping**—considers creating the basis of a final product, which is eventually thrown away; however, it remains valuable to construct further evolving ideas related to the final product (1st and 2nd prototype).

**Evolutionary prototyping**—considers the evolution of a prototype toward a robust final product (3rd prototype). We applied both these principles throughout our development process.

These two principles of prototyping led to the development of an application that was later used in testing. The design and implementation of the three prototypes was made possible by an evolutionary process through several iterations (as already introduced in Table 1 above). Two initial prototypes have been throwaway type while the last one has been an evolutionary prototype.

Figure 4 presents the timeline overview of the different iterative development cycles and stages of the three software prototypes. During these development efforts, we heavily utilized service-oriented approaches for supporting TEL activities in the context of inquiry learning. These prototypes of a software system evolved from being proprietary applications towards combining several Internet-based services to process and visualize the geo-temporal data collected using mobile data collection tools and web. In the first development stage, we mainly dealt with the integration challenges of different technological resources. In addition, the first software prototype has been mainly implemented using static forms for mobile and desktop enabled visualization features and did not provide real time data representation (and thus was throwaway prototype). During the second development stage, we continued with our integration challenges, however, due to the evolvement of requirements, the challenge of interoperability across diverse technological resources arose. Thus, the second prototype resulted in a combination between more dynamic forms for mobile (XForms), desktop and web technologies and included initial cloud services (Vogel, 2011). This combination enabled data collection in both, online and offline modes and real time data representation. Finally, during the third development stage we continued to address the interoperability issues, thus the last prototype of the visualization completely relied on an Internet based environment (real time and online). The final effort during this process was to conduct a user testing study (Vogel, Kurti, Milrad, & Kerren, 2011). The three development stages made our software system more robust as a product compared with the earlier implementation. A usability study was conducted for assessing the web-based visualization tool since it aggregated and presented the entire data collected using mobile devices. Assessment and testing of the web-based visualization tool was an important issue, in order to identify usability aspects that resulted in a number of concrete suggestions for the further enhancement and improvement of. These suggestions were later used and translated into requirements for further development.

![Figure 4. Cyclic prototyping efforts across a timeline related to requirements and prototype testing, mapped with the integration and interoperability challenges](image-url)
The system architecture and implementation

During the efforts mentioned in the previous sections, the technical developments have evolved and changed rapidly, such as in terms of the design aspects, technology choices and implementation, as well as software & hardware components. Despite the rapid changes of such technologies these developments have not been reflected in the changes of the architecture of our software system (Vogel, Kurti, Milrad, & Mikkonen, 2012). The components identified in the initial architecture have been proven resilient to these dynamic changes and requirements. The results described in this paper are an effort to try to tackle some of the challenges described earlier and are an evolution of the efforts we conducted in one of our previous work (see Vogel, Spikol, Kurti & Milrad, 2010). During our earlier explorations and developments, we have proposed and implemented a system architecture that consists of five different blocks aiming to provide logical divisions between the different resources of our system. Figure 5 below illustrates the component view of the system architecture and its potential for expandability with other technologies and external systems.

In the architecture presented in Figure 5, resources are organized into different building blocks that integrate sensors, mobile data collection units, from the server side the data aggregation components used as data and content storage, and visual representation components that utilize diverse web APIs using external services block of this architecture (Vogel, 2012).

![Figure 5. System architecture including main resources and components](image-url)

The design and implementation of our mobile data collection tool followed with the integration of an open source Java based project into our software system (Anokwa, Hartung, Brunette, Borriello, & Lerer, 2009). This project/solution supported the use of a particular open standard called XForm. XForm is a standard based on a W3C recommendation that is used to build web forms for the easy exchange of data across platforms and devices using XML as the data format. In our case, the design of the forms for data collection used in our mobile applications was developed following the requirements identified from stakeholders, as introduced above. The solution we decided to adopt is based on the Open Data Kit (ODK) which supports various types of data and content inputs, including text, audio, pictures, video, visual codes and GPS and makes it possible to annotate the collected sensor data and content with location metadata (Anokwa, Hartung, Brunette, Borriello, & Lerer, 2009). The use of XForms facilitated data interoperability across a diverse range of devices and applications that compose our system. We have therefore developed several mobile forms that rely on the use of open standards, which provided us with flexibility, fast development and easy adaptation and integration of technological resources for different scenarios (user trials). Figure 6 shows the screen shot of LETS GO data collection tool.

In terms of software development, our major efforts were allocated for the implementation of the web-based visualization tool. The latest version of this tool enables the visualization of different types of geo-tagged content and sensor data collected using the mobile application described above. The web visualization tool utilizes APIs that provide multiple visual representations of the data set available in our repository. These representations allow users to actively interact with graphs, maps, images, and data tables. An initial version of this tool was implemented...
completely in AJAX. Spite a user friendly interface and positive feedback from the users, we experienced some performance problems related to loading and processing huge amounts of data using AJAX (Vogel, 2011).

These latest drawbacks inspired the latest version of our visualization tool that has been entirely developed using Google Web Toolkit (GWT). Our latest version of web based visualization tool called GreenLab (see Figure 7) addressed a lot of the requirements generated during our testing efforts by using iterative cycles with users. This tool has become more stable that the previous AJAX version and it has a lot of features for mapping the data automatically, filtering the data based on different criteria, etc. The data visualization process in Green Lab is divided into two stages; first Green Lab selects the type of data collected during the activity with the mobile devices; and second, it presents and visualizes the collected data. In the first stage, Green Lab loads each data type and presents them as clickable buttons. Each type can contain several forms (XForms), which will be loaded after the user has selected a specific activity. In the second stage, Green Lab presents the data retrieved from each form. The filtering of data in Green Lab is located in a panel and the filtering mechanism triggers all active visualization to update, presenting only a certain part of the data set. Each checkbox as such listens for user clicks, triggering an event to update the views. The filtering panel view contains three kinds of options. The first filtering view is based on Organizations (namely participating schools) that follows with the Groups that belong to that organization. Selecting an organization will first filter the data on selected option and also present all groups for that organization, which also allows filtering specific groups. The Attributes filtering view is constituted from each attributes that can be filtered accordingly, by also selecting multiple attributes. Moreover, a single attribute can be filtered also by their values. The Dates filtering view can be used to choose From-To dates to set the time period the user is interested to investigate.

**Figure 6.** LetsGo mobile data collection based on ODK

**Figure 7.** Green Lab – Web visualization tool

Green Lab contains three different visualization views for presenting the data, which are all resizable. The main visualization view starts up by presenting the data as a table, but has the possibility to switch to bar-, columns-, line-,
or an area-chart by navigating to the icons in the header. The second view contains a Map view, which locates and visualizes all the geo-tagged data collected by the mobile application. The last view contains only a Scatter Plot view, which allows filtering of two kinds of data attributes, by comparing them.

The design and implementation stages carried out during the past four years facilitated the identification of the main features in terms of sustainability of the LETS GO system. The three main salient features that we have identified during our design and implementation efforts can be enumerated as following: (1) Interoperability of the software and hardware components (2) Extensibility of the visual representation forms, and (3) Sustainability of the software solution. A more detailed view of these lines of action based on what we learned during the last year developing the LETS GO system are presented in the following sections.

Lessons learned

Proprietary software solutions are deployed extensively through multiple platforms such as the web, mobile devices and desktop applications. The use of diverse standards brings new challenges when it comes to flexibility, interoperability, customizability and extensibility of different components that are part of software systems. Continuous evolving web and mobile technologies combined with the changes of the environment result on dynamic and complex requirements that become extremely challenging. As indicated earlier, this paper aims to tackle some of these problems by providing deeper development insights into the issues of interoperability, extensibility and sustainability related to the LETS GO software system and its evolution during the last four years. These insights were gained during iterative user-centred development cycles of a software system designed and implemented for fostering collaborative science learning activities.

Interoperability of the software and hardware solutions

The notion of interoperability constitutes one of the most important principles in system integration (Zeng & Quin, 2008) and refers to "the compatibility of two or more systems such that they can exchange information and data and can use the exchanged information and data without any special manipulation" (Taylor & Joudrey, 2003). During the development efforts of the LETS GO software system, we have in practice been able to tackle two out of four interoperability categories as introduced by Sheth (1998), namely:

- Syntactic interoperability: differences in data formatting.
- System interoperability: heterogeneous systems and applications.

Despite considerable research efforts, achieving interoperability of various sensor readings and mobile devices and various web services has remained an open issue. In connection to syntactic interoperability our focus was related toward making use of open standards for data exchange such as XML, XForms and JSON. While from a system interoperability perspective, the focus was on making use of open software tools with cloud-based services for matching our requirements.

The features of the cloud environment and services made our system more flexible. Initially, traditional desktop-based integrated development environments were employed and the development then gradually moved towards a mashup-pattern that combined different service-oriented approaches. One identified issue was that the rapid speed and evolution of software and web technologies affected the development process and the application itself. From heavily using Internet based services, such as cloud environments and due to the problems encountered from such services, we started to deploy all our developments into our local repositories and environments. By migrating the LETS GO system to our local environment we mitigated the risks and uncertainties of the cloud development environments. Furthermore, this created new opportunities to closely address the interoperability issues of the software components comprising the LETS GO system. During the last year of development we specifically addressed interoperability issues in the “Mobile Data Collection” component and “Data Aggregation” component.

The interoperability issues in the “Mobile Data Collection” component dealt with both hardware and software issues. Since the current mobile data collection tool was tailored only to Android smartphones and having in mind the emergence of HTML5 and CSS3 as well as multiple mobile cross-platform frameworks (such as PhoneGap, Titanium etc.), we have started expanding the mobile collection tool toward iOS and Windows Mobile platforms.
The purpose of this effort will be to widen the base of the eligible devices to be used as mobile data collection tool (so we do not pose any limitations for the use of the mobile application in schools) as well as making them fully interoperable with the rest of the LETS GO system.

Interoperability issues in the “Data Aggregation” component have been addressed from the perspective of data export capabilities. The idea for such development was to make the LETS GO system open and interoperable with similar tools developed in other research projects. The Data Export component we have developed can be described as a middleware that prepares the surveys data stored on the ODK Aggregate server. The main application is a Java servlet that processes calls from clients and replies either a list of available “SurveyTypes,” links/URLs to all the stored forms and their submissions, or the actual data of the submissions. These calls can be processed in the form of a Web API, where the data is responded as JSON. This enables that all the data aggregated in the ODK aggregate becomes available for export using this servlet and in a JSON format. In the current version of the DataExportServlet all the data is read and exported as the clients upload it and stored in the database.

Beside these efforts we have also been working on the development of the form rules as guidelines for design of the XForms for mobile data collection in a form of naming conventions. These naming conventions developed enable mapping of the collected data dynamically to the visualization tool by the form designed based on our guidelines. All these developments enabled our LETS GO system to be fully interoperable with different mobile devices as well as new visualization tools and services.

Extensibility of the software and system architecture

The massive use of mobile and web technologies for data collection purposes produces vast amounts of data. This is another challenging task that requires attention while trying to make sense of all the data generated by users. Therefore, as the amount of available data continues to grow, conceptualizing and developing new interactive tools for visualization becomes an important task to tackle these challenges, for, e.g., seeking new ways of presenting and sorting appropriate and relevant data, or managing and analyzing information (Ackerman & Guiz, 2011). Different visual representations can provide different insights to users by enabling them to observe data in context, to analyze these data and to draw different conclusions by using different analytical approaches (Eiselle, & Weiskopf, 2009; Sedig, Liang, & Morey, 2009). The extension of web-based visualization approaches, along with new forms of interactive collaborative technologies, is constantly growing (Sedig, Liang, & Morey, 2009). Lately, TEL researchers have been taking advantage of different interactive visualization techniques and tools (Linn, & Eylon, 2011). Research in this area indicates that visualizations have the potential to improve learning outcomes, especially related to inquiry science learning (Johnson, Levine, Smith, & Stone, 2010; Edelson, Gordin, & Pea, 1999; Linn, & Eylon, 2011; Pea, 2002). Moreover, interactive visualizations support and increase students’ engagement in scientific inquiry (Linn, & Eylon, 2011; Pea, 2002). In the scope of our work, “learning through collaborative visualization” refers to developments of “scientific knowledge that is mediated by scientific visualization tools in a collaborative learning context” (Pea, 2002). The latest version of the LETS GO system allows for integrating new interaction features provide by multitouch enabled devices and gesture based interaction in a way that we can expand the interaction modes in which learners work with the visualizations.

Recent developments of our LETS GO system include the implementations of two prototypes using gesture based interaction supported by the use of the Microsoft Kinect (Vogel, Pettersson, O., Kurti, & Huck, 2012) and touch enabled interactions facilitated by the use of the Samsung SUR-40 tabletop computing surface (Müller, 2012). Both these prototypes are fully functional and make use of the data already stored at the LETS GO repository. The initial benefit seems to be the fact that these two new interactions paradigms promote collaboration among users while reflecting upon collected data. Figure 8 below illustrates the Natural User Interface (NUI) for the Green Lab application.

NUI Green Lab is a visually driven explorative interactive visualization tool. The tool focuses on a graspable presentation of the geo-tagged environmental data, collected during outdoor activities using mobile devices, in form of digital maps, charts, and images. The application provides a multi-user interface facilitating the synchronous collocated collaboration of at least two users. The interaction makes use of multi-touch interactions on the SUR-40 tabletop computing system and in-air gestures facilitated by the Microsoft Kinect depth sensor as direct input methods. Furthermore, the interface consists of freely movable digital items allowing the users to set up dedicated

![Figure 8: Natural User Interface (NUI) for Green Lab](image-url)
workspaces and explore datasets on their own. Taking all this into account, the main goal was to provide users, such as students and teachers, a prototype supporting and extending collaborative science learning activities. Therefore, an initial usability study was performed upon these two extended developments (multi-touch and in-air gesture). The initial analysis revealed that the study participants achieved overall better results in the multi-touch scenario compared to the in-air gesture scenario. This led to the conclusion that the in-air scenario is not suitable for complex productive workflows, while the multi-touch interaction is. A distinct advantage in this case could be identified when it comes to the visualization of big amounts of data where multiple users could actively be engaged. All the participants approved the possibilities of the collaborative application and liked the visually driven data visualization and exploration as an offset to traditional workstations with single-user input. This overall finding highlights the need for an integration of new interaction technologies and scenarios in collaborative interactive data visualization, and especially in scenarios related to environmental science learning.

The latest activity we recently carried out with regard to “extensibility” was connected to National Geographic Society’s (NGS) GIS tool, which has been designed to support geographic investigations and encourage collaboration between young citizens and researchers. Our software system was successfully integrated with NGS’ GIS platform called FieldScope. FieldScope has been designed to support geographic explorations and to promote citizen science practices in real-world issues. One of the main drawbacks of NGS FieldScope was the lack of uploading data onsite where the data was actually collected. Thus, the main idea was to extend NGS FieldScope by using our mobile data collection tool. This illustrates the notion of extensibility of our software system in the sense of how sensor data and observations collected using our mobile application combined with the Export function from the Data Aggregation component were used to visualize these data sets in other tool such as NGS FieldScope. The activities we conducted validate the flexibility for data exchange and integration that our software system offers. Our software system has been conceived and based upon the notion of an open and extensible architecture (Vogel, 2012). Figure 9 below depicts the screenshot of NGS’ FieldScope that visualizes water quality data collected using the LETS GO system.
Sustainability of our solution

During the last year of development, our efforts were focused into making the current system a sustainable one, so it can be widely used even after the end of the project. Having in mind the problems we experience with cloud environments (as introduced previously), especially on changes on Google App Engine regarding authentication and limitation of the services, we decided to migrate the LETS GO system (i.e., aggregation server) into our own local environment. The entire environment is based on open source software and open standards. Furthermore, by having full control over this environment we foresee a long maintenance of this software solution and its application across different domains that require mobile data collection and visualization activities. The next step will be making our current solution even more accessible, open and usable for a long period of time.

The user testing study described in earlier sections provided us with additional requirements (Vogel, Kurti, Milrad & Kerren, 2011) that were implemented during the two additional evolutionary prototyping efforts (4th and 5th development iterations, introduced in Table 1). These two additional development efforts in total make five development iterations that made our software system extensible and sustainable while new requirements continuously emerged in these activities. These entire processes made it possible to verify that user requirements were adequately addressed while satisfying their needs. Figure 10 below provides details about our continued evolutionary prototyping efforts (as a continuation from Figure 4 introduced above) across a timeline related to requirements and prototype testing, mapped with the extensibility and sustainability aspects that were considered during these two last iterations. The fourth development stage addressed the extensibility challenges as described above. The fifth development stage made us think to make our system more sustainable towards providing an open platform comprised of a rich set of tools that offer flexible mobile and web based applications that can be deployed by users to support data collection, visualization and collaboration.

The LETS GO system has evolved over a four years period from a prototyping system to become a sustainable one with the possibilities to rapidly be adapted with new features and extensions by taking into consideration the rapid evolution of software and web technologies. In addition, we want to emphasize that the modular design and capabilities of such sustainable system have been conceived with the intention to reduce total platform replacements, where the replacement of a certain component or service as well as the extension of the system with new functionalities becomes increasingly feasible into our solution.

Conclusions and future work

The development lifecycle presented in this paper enabled us to gain valuable insights related to different aspects of mobile and web engineering while developing a system for supporting inquiry based learning activities. During the four years of development efforts, three software prototypes were implemented utilizing service-oriented approaches.
that include mobile, web and interactive visualization modules. The main challenges we identified during these efforts were related to integration, interoperability, extensibility and sustainability of our software system while fostering collaborative science learning activities in the field of TEL. These efforts have been tested with more than 400 users in connection to several trials that took place during this period.

The LETS GO educational activities and tools enabled students to learn in a variety of ways that encompass indoor and outdoor activities across locations and time with the supports of sensor technologies, mobile devices and web-based tools. The experiences and knowledge gained during these years enabled us to develop the LETS GO system to a sustainable and robust platform for mobile data collection, visualization and collaboration. The user trials allowed testing the software applications throughout five development iterations on authentic settings, while new requirements continuously emerged in these activities. Reflecting upon our latest development efforts and the results presented in this paper the main findings of our research are discussed in the lines below.

Collaborative technologies, which were used in the iterations of this project, facilitated the adoption of a learner/user-centred approach. The learner/user-centred approach has been suggested by Bonk and Cunningham (1998), where they emphasized “the need to anchor learning into real-world or authentic contexts that make learning meaningful and purposeful”. Sensor kits, mobile devices, web services and interactive technologies nowadays provide us with a vast amount of opportunities of embedding learning activities into real world settings. In these environments the real challenge is the need of matching the dynamic requirements that are generated during learning activities. Hence the main contributions that this paper addresses are the development insights while integrating the technological resources and support for successful implementation of the educational activities related to environmental science learning in authentic settings.

Based on the insights gained from our research efforts we consider that by utilizing an extensive prototyping approach, the discussion of ideas, designs, requirements and implementation possibilities with users/learners the development becomes more easily manageable and understandable. For testing the technical feasibility and understanding whether the technology and implementation behaved as expected, agile development approaches based on prototyping (by combining throwaway and evolutionary approaches) were utilized. They offered an easy and communicative way to test it on real time activities and in authentic settings. Furthermore, this approach helps to find a balance between the design and implementation stages by considering the rapid evolution of technologies.

The integration of diverse heterogeneous device environments where learning activities take place must be based on solutions that promote data exchange, integration and reuse. In our research, we have identified that using open standards technologies for data exchange, promotes systems interoperability and extensibility with new features. Clear cases of such approaches have been the extensibility with interactive technologies and services such as NUI elements and NGS Fieldscope. The initial benefit of this approach seems to be the fact that these new interactions paradigms promote collaboration among users while reflecting upon collected data during outdoors activities.

Moreover, the rapid technological changes affect the flow of learning processes and educational organizations. For enabling rapid changes to be smoothly reflected in everyday activities in this area, there must be well-defined processes to ensure the continual refinement of the applications developed. Facilitating the communication between research projects/researchers and developers on the one hand and research projects/researchers and educators on the other hand, are key factors for the success of these interventions and their sustainability. This approach would enable implemented technologies and applications be closely integrated into everyday educational practices, thus maximizing the benefits in terms of the long-term goals, costs, time, and to satisfy learners/educational institutions with their system. A systematic view on those aspects and their implication for developing sustainable software solutions to support mobile learning could lead to a number of potential benefits:

• Standard based systems
• Constant interaction with users/learners
• Incremental development
• Reduced time and costs
• Expandability
• Flexible change of technologies
• Higher usability
• Easy maintenance and sustainability
In summary, all these identified research insights and benefits were gained during iterative user-centred development cycles of a software system for fostering collaborative science learning activities. Moreover, they provided solid foundations in terms of the possibilities of tackling the requirements for supporting inquiry learning in a flexible manner. From a system perspective, these requirements are best fulfilled by using service-oriented approaches that facilitate interoperability through utilizing open source and open standards and by following the evolutionary approach of prototyping. These findings are directly related to software engineering processes aiming to address the requirements posed by mobile learning scenarios and applications. The issues of interoperability and extensibility of the software solution are directly connected with the possibilities of dynamic reconfiguration of learning spaces to respond to learners’ contextual needs. A sustainable design of technological support to meet these needs it requires to be closely developed and deployed in close iterations with different stakeholders (including also teachers and students). This approach may help teachers and learners to overcome some of the complexity of the learning activities and furthermore it may promote the seamless integration of physical and digital learning resources.

References


Informal Participation in Science in the UK: Identification, Location and Mobility with iSpot

Eileen Scanlon*, Will Woods and Doug Clow

Institute of Educational Technology, The Open University, UK // Eileen.Scanlon@open.ac.uk // Will.Woods@open.ac.uk // Doug.Clow@open.ac.uk

*Corresponding author

ABSTRACT

Informal participation in science is being recognized as an important way of developing science learning both for children and adults. Mobile learning has particular properties that have potential in informal science settings, particularly outside traditional educational settings. Mobile technologies provide new opportunities for learners to engage with science on the move. This paper reviews the impact of participation in informal science settings on some members of the public using the experiences of the iSpot project as a case study. iSpot aims to create and inspire a new generation of nature lovers by getting people to explore, study, enjoy, and protect their local environment. It facilitates an inquiry learning approach to identification of wildlife with support provided by a community developing around the resource. The iSpot project described here provides evidence of the ways in which informal participation in science can be enhanced by the use of technology. We draw on the findings of two case studies within the project - iSpot Mobile and iSpot Local. These demonstrate particular ways in which location-based activity and mobile learning can be developed and have an impact on the informal learning of science.

Keywords

Informal learning, Participation, Science learning, Mobile learning

Introduction

This paper discusses informal science learning in mobile contexts, and the theoretical framing and development processes used in the creation of iSpot (http://www.ispot.org.uk). It analyses two projects related to iSpot – iSpot Local, and iSpot Mobile – which have developed particular approaches to the support of informal participation in science.

Mobile learning

Informal participation in science is being recognized as an important way of developing science learning both for children and adults (see e.g., Bell et al., 2009). Informal learning (see Trinder et al., 2008) has become an important area of interest for education researchers in recent years. Livingstone (2001) has documented the informal learning opportunities used by adults and the issues which arise in studying such settings. Mobile learning has particular properties that have potential for productive activity in informal science settings, particularly outside traditional educational settings. Sharples et al. (2009) define mobile learning as “the processes (personal and public) of coming to know through exploration and conversation across multiple contexts, amongst people and interactive technologies” (p. 5). Sharples provides some examples of this process including MyArtspace where school pupils use mobile phones to support learning on fieldtrips to museums. In particular, mobile technologies provide new opportunities for learners to engage with science learning. Dierking et al. (2003) have a view of learning as a cumulative process involving connections and reinforcement among a variety of learning experiences and describe informal science education as “science learning which is strongly socioculturally mediated and occurs across a wide range of physical contexts” (p. 109).

Here we discuss the impact of participation in informal science settings where mobility is an asset. The National Science Foundation describe informal learning as follows: “Informal learning happens throughout people's lives in a highly personalized manner based on their particular needs, interests, and past experiences. This type of multi-faceted learning is voluntary, self-directed, and often mediated within a social context…; it provides an experiential base and motivation for further activity and subsequent learning.” (NSF, 2006, Section I, Introduction.)
Increasingly it is recognised that mobile technology can play a part in Citizen Science activities (discussed further below). See e.g., Robson (2012) describing the use of mobile phones in CreekWatch.

It is important to emphasise with Sharples et al. (2009) that in mobile learning what is mobile is the learner. This is important for the topic of this paper: mobile learning of science in informal settings. In the iSpot case studies which follow, the learner is always mobile, sometimes accessing a website from a field location, sometimes using a mobile device but always engaged in location-based learning. Mobile learning in science settings has been studied both in formal and informal settings. There are a range of relevant studies in formal learning (e.g., Littleton, Scanlon, & Sharples, 2012; Chen, Kao, Yu, & Sheu, 2004) but fewer in informal settings. Early examples of studies which demonstrate the potential of mobile learning in informal science settings include that of Clough (2009). She describes developing mobile support for nature trails, and researching the use of mobile technology with GPS in the geocaching community. It is a challenge for learning scientists to develop and study learning in such completely informal settings.

Approach to development, theoretical framing of the design and methodological challenges

In this section we set out the development of the design of the iSpot project. A core group (Jonathan Silvertown, Martin Harvey, Richard Greenwood and Doug Clow) led the creation of iSpot, the iSpot website, and generated its initial design by informal discussion, based on the expertise they brought to the project, which included field biology, citizen science, online learning, and software development. Some iSpot team members were driven by theories of participatory design. An initial motivation was the exploration of applications of geographically referenced teaching and learning. Next we compare the features and intentions of the work with theoretical perspectives from research on learning.

iSpot supports a community of practice (Lave & Wenger, 1991) where members learn from legitimate peripheral participation (Wenger, 1998) and develop their expertise through a process close to apprenticeship. A central theoretical design principle for work on communities of practice is the support of different modes of participation. Preece and Shneiderman (2009) set out a 'Reader to Leader' framework, categorising successive levels of social participation in online communities as reading, contributing, collaborating and leading. Other work suggests that a developmental model is not a good fit with observed activity in online learning sites: rather, different users participate in different ways at different times (as described in the 'Fairy Rings' model see Clow and Makriyannis, 2011).

One way to consider a contribution on the iSpot website is as a shared social object (see e.g., Knorr-Cetina, 2001) which can structure this participation, and scaffold participation in the community of practice. iSpot also reflects the constructivist notion of authentic learning activities (Jonassen, 1999) together with what Scardamalia and Bereiter (2006) describe as knowledge building: The learning activity is not only akin to scientific activity, it initiates learners into the knowledge-creating culture and enables them to actively contribute to scientific knowledge.

The development of a system such as iSpot needs to combine, in a cyclic approach, research, pedagogical design, and technology development. Accounts of socio-cognitive software design (Sharples, Taylor & Vavoula, 2007; McAndrew, Taylor and Clow, 2010) are influential in developing such processes, as are principles of Agile software development (http://agilemanifesto.org). Substantial engagement and envisioning activities with stakeholders were conducted, followed by deployment of the system to gather feedback from users.

There are a number of definitions of design-based research: our approach was in line with Barab and Squires’s description: “Design-based research [...] was introduced with the expectation that researchers would systemically adjust various aspects of the designed context so that each adjustment served as a type of experimentation that allowed the researchers to test and generate theory in naturalistic contexts” (Barab & Squire, 2004, p. 3). In some aspects of our project, particularly the development of the mobile app this included iterative cycles of designing (both pedagogy and technology), running an inquiry, and then evaluation and analysis that fed into the next cycle. Thus some of the key findings of the research become embedded within the system: not just in the design of the software, but in how it is used by the growing and developing community of practice.

So the approach taken in the design of iSpot was the co-design of technology and pedagogy i.e. to design the educational activities and technology together, drawing on a participatory design approach (see e.g., Penuel, Roschelle & Schetman, 2007).
There are methodological and practical challenges associated with developing an understanding of how learning takes place in the communities which use iSpot. The learning episodes which involve a user can be relatively short and informal. An important perspective on learning that comes from the public understanding of science movement is to think more broadly about the impact of engagement. Relatively simple models of learning, such as the deficit model used at first in work on the public understanding of science, were replaced by an investigation of the potential outcomes, including increased awareness and impact on attitudes, as well as engagement and participation. Groups enabled by technology will form round particular interests and issues suggesting a need to assess how expertise can develop in these groups. There is a complexity to examining such learning settings as iSpot. We need to look more broadly at them, in terms of new data and analysis methods (Scanlon, 2012, July).

Citizen science

In considering the learning which takes place through participatory science enabled by the use of mobile technology in field settings, it is necessary to look for some different ways of examining those learning settings and the use of mobile technology. Dron and Anderson (2007) describe how online communities enable different types of participation in the form of groups, networks and collectives.

iSpot may be described in Dron and Anderson’s terms as a network. It also can be seen as an example of citizen science. Wiggins and Crowston (2011) provide a typology of citizen science projects where members of the public work in combination with researchers. Hand (2010) and Newman et al. (2010) caution that additional verification may be necessary on projects which involve citizen scientists. Rotman et al. (2012) surveyed volunteers on ecological science projects to find out their motivations for participation, and many cited their desire to increase their scientific knowledge.

iSpot

This section describes iSpot, the project at the heart of this paper. This paper uses the iSpot project to examine the ways in which informal participation in science can be enhanced by the use of technology, and in particular ways in which location based activity and mobile learning are developed in the project. iSpot allows an inquiry learning approach to the identification of wildlife with support provided as part of a community of practice. It is important to note however that in what follows we are drawing examples from approaches taken in the particular case studies, rather than describing the whole cycle of development in the iSpot project or all the particular design decisions taken to develop its website.

iSpot (McAndrew et al., 2010; Woods & Scanlon, 2012) aims to create and inspire a new generation of nature lovers by getting people to explore, study, enjoy, and protect their local environment. The iSpot web site (home page shown in Figure 1), launched in June 2009, allows users to post observations of animals and plants on the site, and the iSpot community helps to identify them reliably. As a web-based system was used, this allows users to access and learn 24/7 and at anyplace with Internet access. These observations constitute the ‘shared social object’. Support is provided for identification partly by online resources but more fundamentally by the community of practice active on the site. The site connects together informal novice learners with experts in a wide range of natural history fields, including over 100 who are representatives of natural history organisations. Learning the name of an organism you have observed is the first step in learning more about it. Furthermore, the process of recording observations of species - including the name of the species, the location and the time of the observation - is the fundamental unit of activity in biodiversity monitoring and research. Indeed, selected observations from iSpot users are now used as part of formal biodiversity monitoring. Thus iSpot enables learners to engage in Scardamalia and Berieter (2006)'s knowledge building: they contribute to new knowledge, as a community activity.

A key feature of iSpot is its sophisticated but easy-to-use reputation system, which provides an indication of each user's expertise on the site (see Clow and Makriyannis, 2011). Unusually among online reputation systems, as well as providing an indication of “social” reputation on the site, the iSpot reputation system includes elements designed to provide sound indicators of the expertise – or learning – displayed through activity on the site. The reputation system structures and makes manifest expertise, facilitating learners' development within the community of practice.
iSpot findings on participation and learning

The impact of iSpot can be seen through its wide reach. Currently it has over 31,000 registered users who have added more than 200,000 observations with over 340,000 images, identifying more than 6,900 different species. The project has identified two species previously unrecorded in the UK: a bee-fly (Systoechus ctenopterus) and euonymus leaf notcher moth (Pryeria sinica). Further empirical analysis of learning activity on iSpot is underway, but some initial findings are presented here.

Qualitative analysis shows clear examples of users who start as complete novices, but come to fairly sophisticated understanding of identification. There is also quantitative evidence of users learning. For instance, analysis of a sample of 407 users as they progressed through submitting and identifying their first fifty observations within iSpot is strongly suggestive of learning. As shown in Figure 2, users showed improvement in their ability to identify other people’s observations over the period that they submitted observations: As users progress from their first to their 50th observation posted on iSpot, they have a bigger percentage of correct identifications that is they are more likely to identify what they have seen for themselves.

Figure 2. How people improve in identifications from repeated use of iSpot
The crowdsourced identification model within iSpot, rewarding improvement in ability to identify observations, provides some of evidence that people are learning and improving their understanding of nature through iSpot. However a person may gain reputation through identifying very common species and without expanding their knowledge of other species.

In order to get a better understanding of how and whether people learn from using iSpot we require empirical evidence of improvement in people’s ability to identify a greater variety of observations as their reputation improves. We designed the iSpot intelligent quiz to test this knowledge. The quiz was launched in July 2013, since then around 350 people per week have taken one or more quizzes, so an average of around 50 people per day. The quiz is tailored to the level and subject area that people request when they start a new quiz on iSpot. The reputation level that iSpot provides is a good indicator of the level that people should take but there is no restriction on the level so, for example, a level five expert could take a level 1 quiz and vice versa. The data from the weekly logs shows however the people are averaging about 7 out of ten for quizzes across the skills levels which suggests that people are naturally finding a level which challenges them.

Figure 3. Screenshot of the iSpot intelligent quiz

The quiz has a number of different types of question that test a range of knowledge within a specific domain, some questions are multiple choice and others are about entering the correct name or type of observation. The data collected so far indicates that people who use iSpot are gaining knowledge about nature.

Face-to-face outreach work has reached over 55,000 beneficiaries, over 10,000 from hard-to-reach groups, whilst over 800 participants have used iSpot at local “bioblitz” events, including schools, local government and voluntary sector organisations.

This account of iSpot provides the framing for the description of two specific projects linked to iSpot that particularly explore mobile and ubiquitous learning: iSpot Local and iSpot mobile.

### iSpot mobile

The first case study linked to iSpot is iSpot Mobile. The iSpot website was already available to be viewed on mobile phones. However since people are outdoors making observations, there was both a need and an opportunity to use mobile phones with digital cameras to make observations and interact with the iSpot community.
The iSpot mobile design approach

A lightweight contextual design approach to establishing the requirements for the mobile app was taken based on the user-centred design process developed by Beyer et al. (1998), exploring the types of users, the scientific context of nature study, the environment which they would be exploring and the learning outcomes to be achieved. We defined the main purpose of developing the mobile app as allowing users to create and upload observations (a combination of photo, identification, and location) to iSpot using their mobile device and to become part of the iSpot community using tools for sharing information. The secondary purpose was to enable iSpot website functionality on a mobile device in a native format and using the enhanced capabilities of a multi-touch mobile phone. For example the ability to pinch to zoom on images to see greater detail and the ability to use the devices to interact with the iSpot community whilst on the move and to enhance their experience through utilising the geo-location services available within mobile devices.

A core group consisting of Jonathan Silvertown, Martin Harvey, Will Woods and Richard Greenwood produced the initial design. A light-touch user-centred design approach was used for app development, beginning with a storyboarding process using experiences from users of the current iSpot website. Specifically, data was collected from a small selected group of experienced iSpot website ‘volunteer’ users whose practice was monitored through interviews and forum discussion, taken alongside usage data from the website, and feedback from the core group to establish common patterns of use. These were converted into stories to build a coherent functional specification. For example, one user said “I am running an inquiry based learning project and I want to [use the iSpot mobile app to] develop scenarios around ecosystems that I’m observing, for example birds in my garden, population of bugs in my flower bed, fauna in my pond ...”

The user-centred design approach that the team adopted involved gathering feedback about how people engaged with early prototypes of the environment to inform later iterations. Twenty people volunteered. A small number of volunteers from the existing iSpot community were also invited to participate, including iSpot “mentors” (associates who work with iSpot to assist others in identifying observations). A series of usability and accessibility testing cycles were conducted during the course of the app development. The feedback was gathered and interpreted by the project team to help improve the functionality and design of later iterations.

First iteration

The first iteration of the app started in October 2011 and took a total of ten weeks including development, bug fixing and testing. An initial issue was that Android devices have all manner of shapes and sizes of screen and this made the display of the images a challenge. Figure 4 shows a screenshot of the observation list.

Testing took place over a two-week period which included an evaluation conducted by a usability expert. The results indicated that the app was missing some critical functionality and had a number of bugs.

Figure 4. Screenshot of original iSpot app design
The application was also provided to a group of ten experienced mobile users. For example, one experienced iSpot user suggested a process of checking and validating an observation using the mobile app, producing the following scenario: “What iSpot offers is an authoritative resource for helping people learn identification. The new app could be like having an expert out in the field with you which is, I'm sure you'll agree, the best way to learn identification; in the field not through photographs.” To test these scenarios, users were asked to go out and take observations in naturalistic settings and then gain identifications from the iSpot community and then to provide feedback on this experience.

From the feedback it was clear that people were generally enthusiastic about the functionality of the app but they were less positive about the interface design. For example, here is a quote from notes taken during an interview with one of the testers:

“She thought she had to put something in the scientific name or the common name and did not realise that she could leave these blank (she knew it was a ladybird but there was not the option to say just ladybird so she selected one of the named ladybirds, a 10 spot one, even though she knew it was wrong just to get to the next screen and submit the observation).”

The iSpot service is distinctive from competitors as it references species dictionaries and because observations are identified by the iSpot community, often within a very short time of being observed and uploaded: half of all un-named observations are identified within an hour of appearing on the site. The app therefore provides these unique services to mobile users, allowing them to have observations identified and potentially to identify and agree with the identification of other people’s observations.

The evaluation process established that the service created for iSpot Mobile largely mimicked the iSpot website navigation and the design felt quite sterile. The team concluded that the app should therefore be completely redesigned around a navigation and layout more suitable for a mobile app, increasing the interactivity and social elements.

Second iteration

In January 2012, a second iteration of application specification, design and development took place. A mobile interface designer worked alongside the developer to implement a set of improvements to the interface.

This design iteration involved providing a big button menu screen as the ‘home’ screen to get into the main app functionality (Figure 5). The designer created a stylised logo and incorporated design features of the iSpot website to improve the app and make it feel more nature related by using grass and wildlife within the layout.

Figure 5. Second iteration of iSpot app
The redesigned tool the users directly to the observations. We made the observation thumbnail images larger to increase usability and aesthetic appeal (see Figure 6). To avoid removing valuable screen ‘real estate’ on what is a small screen we explored using a dynamic menu which users could click on or swipe to view and which provided all the functions within the application, allowing extensibility using horizontal swipe to access menu choices.

![Figure 6.](image-url) The image-centric design adopted for the beta release of the iSpot app

Further testing was conducted with another group of ten mobile proficient users. This interface received positive feedback from users, including the design, with comments such as:
- “Pull down icon menu intuitive once you try it for the first time”
- “Tried taking photo of pot plant and identifying it. Intuitive interface. Easy to add details. Recognised my location. Though somewhat cramped with keyboard. Pleased to see my first observation appear on iSpot.”
- “Overall I have found the app to be extremely stable, easy to navigate and fairly intuitive.”

However, there were still concerns that the navigation was not providing rich interaction and direct engagement, and that this interface design was not scalable, i.e., as functions were added how would they be incorporated into the fixed four button menu?

As a consequence of the positive feedback from both user testing and technical testing the team felt in a position to move towards releasing the beta version of the app to the public. The Android iSpot application “stable beta” was released to the public via the Google Android app store (Google Play) on 8th June 2012.

**Third iteration**

The third iteration of development began in August 2012. This iteration incorporated improvements to the application through the feedback gained from the testing processes, through user feedback from the beta release, and through use of enhanced reference material from Google on designing for the Android Platform (http://developer.android.com/design/index.html). The beta app on the Google Play store received positive feedback from the public.

The third iteration included enhancements to the geo-location services to provide “around here” information about observations within a specific locale, i.e., within a 1 kilometre radius of the current location using the GPS capability of the device. Users can also scroll to move the map location and receive information about observations within a 1 kilometre radius of any location. There are enhancements to the social and community aspects of the application, in particular allowing users to identify other people’s observations as well as comment on them. Finally, there are improvements to the discovery and filtering services, to filter on species type, to allow users to quickly find out information related to a particular observation, and to create their own individual journeys of self-discovery.
The full release, as a consequence of the testing and evaluation, provides a richer and more interactive experience with an improved user interface, including a contextual “active menu” and larger images, as shown in the sample screenshots below (Figure 7).

![Sample screens from current iSpot app development showing (1) “active menu” and text overlays on images (2) The slide out navigation panel (3) The post comment and post ID capability](image)

**Figure 7.** Sample screens from current iSpot app development showing (1) “active menu” and text overlays on images (2) The slide out navigation panel (3) The post comment and post ID capability

### iSpot mobile testing and evaluation

A further round of comprehensive testing was conducted prior to release using the state-of-the-art mobile eye tracking and mobile data capture facilities available within the Open University Jennie Lee Research Labs (http://jennielee.open.ac.uk). The app was judged to be more robust, fully featured and a better user experience. For example comments included:

“[The]’Around here’ map showing locations of observations in my immediate vicinity seems clear …and easy to use”

After the testing and feedback, the version 1.0 product was released to the Google app store (Google Play) in December 2012. It is achieving over 1000 installations per month and currently has a user rating of 3.8 out of 5 (27 September 2013).

As learners become more mobile, the mobile apps may become the default way of engaging with iSpot and establishing participatory science learning journeys. The app may prove particularly suitable for individuals or groups engaging in local community bioblitzes. The iSpot team expect to use the app to support local group learning activities of this type in the future.

### iSpot local

The iSpot Local project extended the iSpot approach to investigate the potential of using hyper-local events to frame the learning activity, moving it from a largely virtualised activity (on the iSpot website) to a grounded, community, mobile setting – including beyond the reach of electronic networks. This built on and extended the existing community of practice and knowledge building approach.

### Bioblitzes

The key mediating event in iSpot Local was the bioblitz, a survey of the wildlife at a particular site at a particular point in time - say an afternoon, or a day. The general public, supported by a team of experts, try to identify and
record as many different organisms as they can within the time. This can generate real scientific data (knowledge building) as well as engaging the public in the scientific process - and the site itself - through active participation and learning within a community of practice. However, in traditional bioblitzes, it can be difficult to manage the data generated by the public, and identifying the species observed is problematic.

iSpot Local addressed these challenges by coupling the wider perspective and observational recording abilities of iSpot with hyper-local engagement with community stakeholders and effective practical management of the bioblitz events. The basic activity of iSpot Local is set out in Figure 8, a cartoon developed to explain the bioblitz to participants.

Six bioblitzes were organised across the South West of England, at a range of sites from schools to more public sites. The IT facilities available on site ranged from a high-speed wifi network and room full of dedicated computers (at a school) to a nature reserve with no power, no network, and negligible mobile phone voice signal. The team used a hybrid, flexible approach to technology to maximise the benefits given the nature of the site, typically using a set of laptops in a marquee to log photographs and observations, which were uploaded to iSpot later if connectivity was limited on site.

An important feature of iSpot Local was the way in which mobile access - mediated, supported and contextualised - enabled the hyper-local (the individual bioblitz) to connect to the worldwide (the international community network of experts and enthusiastic amateurs on iSpot).

![Owl and Badger go iSpoting](image)

**Figure 8.** Cartoon produced to help explain the iSpot Local approach

**Development approach**

Engagement with a wide range of stakeholders was critical to the success of the project. The funded project partners were the UK Open University, Ambios Ltd (a small not-for-profit company promoting environmental understanding) and Learning South West (a membership organisation coordinating learning and skills and youth work, with members including local authorities, colleges, private training providers and voluntary sector organisations). In line with the design approach, the project partners engaged extensively with many other stakeholder organisations including adult educators, family learning specialists, local government, volunteers, technical specialists and natural history experts.

The project developed and validated a three-phase model for ensuring effective participation including pre-bioblitz work, the bioblitz itself, and post-bioblitz activities. Thus the participation in the on-site activities was scaffolded and embedded through framing and linking activities, enabling the learner's participation in the community of practice.
In addition, as part of the iterative approach to development, some technical development was carried out to create a module to enable observations from a bioblitz to be embedded within a community website, part of which is shown in Figure 9 below.

![Figure 9. iSpot Local map showing observations at one of the bioblitz sites](image)

**iSpot local evaluation**

Evaluation started at the beginning of the project, with the production of an initial Evaluation, Dissemination and Mainstreaming Action Plan. As described above, the design-based research approach meant that much of the outcomes are embedded within the practice developed as the project continued.

The bioblitz sites were generally open for participants to come and go as they pleased, and the participants were very diverse, from primary school-age children through to retired-age adults in their 80s, and with previous experience of nature ranging from negligible to expert naturalists.

To engage with this diversity, a range of evaluation methods were employed to supplement analysis of the online activity, including observation, registration cards and evaluation cards during the bioblitz, and follow-up discussions with selected participants and stakeholders (e.g., school teachers).

Participation levels recorded through Registration Cards were high at the events, with significant data uploaded to iSpot. There was also a high level of activity by the wider iSpot community, with tentative identifications arising from an iSpot Local bioblitz rapidly translated into confirmed identifications on the website, and high numbers of others indicating agreement and posting further comments.

In total, 820 people participated directly in the six iSpot Local bioblitzes, making more than 1,800 observations in the course of the bioblitzes. On iSpot, these observations received over 2,000 identifications (some observations had more than one identification), from the bioblitz participants and the wider iSpot community. These identifications in turn received over 3,000 agreements. Most participants (74%) were children under 18 and the rest were adults. The gender ratio was roughly equal (53% female, 47% male). As a result of this engagement and participation, the iSpot reputation system was able to confirm over 1,250 observations as having a “Likely ID”-confirmed by sufficient expertise. This is clear evidence of the participants taking part in genuine “knowledge building”: despite the diversity of their initial expertise, they were able to jointly contribute to new knowledge.

The feedback from the participants shows further evidence of the participants' engagement and learning—for example, the feedback from the evaluation cards included a parent reporting that the best bit was “Watching my children get so involved, questioning and learning about the world around us that we don’t always stop to appreciate.” Another participant reported that they gained “A better appreciation of just how much wildlife lives alongside us in the school field. Brilliant experts, really approachable.”
Engagement with iSpot Local motivated many participants to engage further with learning about nature—for instance “Examining and cataloguing and drawing the wild flowers in the lane to Granny's house” and ‘Have looked at iSpot site and held our own mini bioblitz in the garden.” The diversity of the participants was reflected in the diversity of outcomes from the activities: at one site, the volunteers engaged in conservation work planned to run repeat events annually to track the effects of their work; at another (a school), follow-up learning events targeted to the children’s interests and the curriculum were developed.

The use of bioblitz events coupled with the iSpot website, in the context of a wider learning community, shows the potential for the iSpot website to support a vision of mobile, ubiquitous and lifelong learning at many levels, harnessing the power and range of a global, broad network of expertise with local concerns and knowledge.

The individual learner, located in a particular environment, was connected to multiple potential sources of learning, ranging from informal personal contact with experts through to technology-mediated access to explicit learning resources and relevant formal education opportunities. This rich environment structured their apprenticeship within a community of practice, and enabled them to engage in knowledge building.

Conclusions and lessons learned

The paper has drawn on an evaluation of the iSpot Local and iSpot Mobile projects to consider evidence on the impact of participation, and on which features can be identified as important in the design of such community projects. In particular, explicit attention was paid to how learners can be supported to be members of a community of practice, with participation structured around a shared social object, engaging in knowledge building as active contributors to knowledge, and the iterative, integrated approach to development have all proven valuable.

The overall experience of the iSpot project with its analysis of the improvement in identification as users become more experienced provides some evidence of knowledge development. However it is also possible that the users are becoming more proficient with the system so there is room for further investigation of how learning taking place with the system.

We know from the analytics of people’s progress through iSpot that they appear to be improving their identification knowledge as they become more experienced users of iSpot. The new quiz service within iSpot, also available on mobile, will provide further evidence to assess whether learning is taking place.

iSpot Local and iSpot Mobile are two elements within a comprehensive roadmap for iSpot development, with a full range of objectives to be achieved within the project through to the end of 2014. These include the internationalisation of the service, extensions to support integration with other systems and services (Facebook, mobile, species dictionaries), improvements to the service robustness, personalisation and the ability to support local and regional content to create an adaptive user-centred service and services to further test the learning that is taking place through analytical tools and intelligent quiz to track how users are increasing in their ability to identify and understand nature.

The iSpot Local project and the iSpot Mobile app are examples of mobile learning: the learners access iSpot in a range of contexts, settings and locations as appropriate to their individual situation. To a degree they are also an instantiation of the vision of context-aware ubiquitous learning (Hwang et al, 2008): the location of an observation is a crucial piece of information on iSpot, and it is possible to use location-aware sensors to capture this data automatically from the learner’s context.

Each case study illustrated a different facet of informal participation in science and contribution to knowledge in this area. The iSpot Mobile case study demonstrated the impact of a design based research approach to the development of such systems. The formative feedback also provided us with information on the processes by which the mobile app would facilitate learning. The iSpot Local case study showed how online communities of practice can be extended and connected to physical locations, providing more contextual opportunities for knowledge building, and co-creating new scientific data.
This experience and analysis demonstrates some of the potential for mobile and ubiquitous technologies to support learning in informal contexts but certain issues remain. These new developments include an iSpot site created for Southern Africa, managed by the South African National Biodiversity Institute so is showing evidence of how the initiative can be translated into new settings. Also, iSpot has linked up with Treezilla an ambitious Citizen Science project to map all of Britain's trees and record vital data about tree disease and the environmental benefits that trees provide, developing a mobile app for use as part of the Open Science Laboratory initiative. The sustainability plan involves moving the infrastructure to managed cloud-hosting over the next twelve month period and for moving support for the application and services to the central IT department over the next two years, to be completed by July 2015. This underwriting of such a research system demonstrates the understanding of the importance of iSpot to the Open University and to the growing community that it supports. The iSpot team led by Jonathan Silvertown works in partnership with nature organisations (currently more than 100), the iSpot community, and other stakeholders, to enhance iSpot and to help further assure its longer-term future.

Acknowledgements

The authors are grateful to all those involved in iSpot, including Jonathan Silvertown, the originator and overall project leader, the rest of the iSpot team, the many groups and organisations supporting iSpot, and all participants in iSpot. The analysis of learning in iSpot and Figure 2 summarizing data were provided by Jonathan Silvertown and Martin Harvey. iSpot was funded by the UK National Lottery through the Big Lottery Fund for England between 2007 and 2012, as part of the Open Air Laboratories project (www.opalexplornature.com) and the British Ecological Society and the Wolfson and Garson Weston foundations. The iSpot Local project was funded by JISC through the eContent Programme.

References


Implementing Mobile Learning Curricula in Schools: A Programme of Research from Innovation to Scaling

Chee-Kit Looi* and Lung-Hsiang Wong
National Institute of Education, Nanyang Technological University, Singapore // cheekit.looi@nie.edu.sg // lunghsiang.wong@nie.edu.sg

*Corresponding author

ABSTRACT

Many countries, regions and education districts in the world have experimented with models of one-device-per-student as an enabler of new or effective pedagogies supported by mobile technologies. Researchers have also designed innovations or interventions for possible adoption by schools or for informal learning. Of critical interest to the community is the question of how the more successful of these top-down or bottom-up models or innovations can proliferate to more usage, adoption and adaptation across levels of the education system. This paper describes a research programme that demonstrates how to make successful research innovations count in practice and that delineates what types of educational R&D involving scaling need to take place to make the critical link to impacting practice. We do this in the context of one such curricular innovation in a Singapore school that moves through the various phases to where the innovation is becoming an integral part of routine classroom practices.

Keywords
Mobilized curriculum, Translation and scaling up, Seamless learning model, Design-based research, Science education

Introduction

The literature on educational technology research is packed with examples of pilot studies and proofs-of-concepts. It is rarer, in fact, in the literature, to see a project move through the various phases to where the innovation actually has become an integral part of routine classroom practices. In our collaboration work with a primary school in Singapore, we have developed a viable innovation model (the Seamless Learning Model or SLM in short) by working with a class of primary school students over a period of two school years. The innovation involves the transformation of the existing science curriculum into an inquiry-based one which leverages the affordances of mobile technologies. Because SLM demonstrated increased student achievement, the school has decided to scale-up the roll-out of the transformed curriculum to more classes and more subjects in the coming years, thus providing the opportunity to study an innovation as it scales up.

In this article, we trace the journey of this research programme that started with the co-design of a 1:1 (one-mobile-device-per-student) mobilized curricular innovation for a primary school in Singapore, leading to the establishment of efficacy findings by researchers, the decision to scale-up by the school, and the plans for scaling-up. In doing so, we elucidate a multi-term research agenda on an 1:1 mobilized curriculum that studies scaling beyond the initial proof-of-concept to broad and deep usage in the context of a research-based innovation in a school in Singapore. Such an agenda articulates the research questions and posits the scaling research framework and approaches involved. We hope this can provide an existential example of how researchers can do research that addresses the multi-term, multi-pronged, multi-level and systemic aspects of school-based innovations, and that ultimately benefits schools and yet at the same time, derive and refine scientifically and empirically theoretical frameworks, design principles, resources and strategies for learning.

Our research approach: Design-based research and our roles as meso-level mediators

With the goal of working towards scalable and sustainable classroom practices, we took a design-based research (DBR) approach to address complex problems in real classroom contexts in collaboration with practitioners, and to integrate design principles with technological affordances to create solutions to real needs of teaching and learning. The goal of design research is to conduct rigorous and reflective inquiry to test and refine innovative learning environments as well as to refine new learning-design principles (Brown, 1992; Collins, 1992). DBR is iterative as researchers strive to engage in design, work with teachers to enact the design in classroom settings, do research on
the contextualized learning processes, develop or refine theories of learning, engage in iterative re-design, and thereby continue the cycle of design and implementation. It is characterized as being interventionist, iterative, process-oriented, utility-oriented and theory-oriented (van den Akker, Gravemeijer, McKenney, & Nieveen, 2006). It is distributed across teachers and is ongoing, as opposed to a completed trajectory that we as researchers can foresee and oversee. DBR can result in greater understanding of a learning ecology by designing its elements and by anticipating how these elements function together to support learning (Cobb, Confrey, diSessa, Lehrer, & Schauoble, 2003).

Cognizant of the multiple level constraints that act on teachers adopting new curricular innovations in the classroom, we recognize the complex interplay of multiple dimensions of education reforms. Thus, we approach our programme of research from a systemic change perspective that recognizes the micro, meso, and macro levels of educational systems (Looi, 2011; Looi, So, Toh, & Chen, 2011). The policy imperatives governing Singapore’s educational landscape constitute the macro-level factors, and the contextualized classroom-based work and interactions as micro-level factors. By meso levels, we adopted the view of Jones, Dirchinck-Holmfeld, and Lindstrom (2006) where they define: “meso is an element of a relational perspective in which the levels are not abstract universal properties but descriptive of the relationships between separable elements of a social setting” (p. 37). Meso-level agencies can be perceived as the “recontextualizers” or “constructors of pedagogic discourse who de-locate and re-locate discourse, moving it from its original site to a pedagogic site” (Jephcote & Davies, 2004, p. 549).

<table>
<thead>
<tr>
<th><strong>Macro-level actors:</strong></th>
<th>Policymakers or other actors who set the climate or policies for educational reforms in schools and in learning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Macro-level environment:</strong></td>
<td>As seen in national plans or Masterplans where a conducive macro-environment for innovative practices is enabled by governance practices through:</td>
</tr>
<tr>
<td></td>
<td>- Setting up the infrastructures</td>
</tr>
<tr>
<td></td>
<td>- Creating readiness</td>
</tr>
<tr>
<td></td>
<td>- Phasing changes</td>
</tr>
<tr>
<td></td>
<td>- Institutionalizing and undergoing creative renewals</td>
</tr>
<tr>
<td></td>
<td>- Providing resources</td>
</tr>
<tr>
<td><strong>Macro-level emphases:</strong></td>
<td>Setting broad educational outcomes</td>
</tr>
<tr>
<td></td>
<td>- Scanning trends and directions, and reviewing research at meso-levels and micro-levels that inform pedagogical and technological practices</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Meso-level actors:</strong></th>
<th>Researchers as recontextualizers who moved discourse from original to pedagogic site</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Meso-level environment:</strong></td>
<td>The socio-cultural factors that make up the school’s learning ecology such as the classroom setting situating between individual activities, small groups and larger communities.</td>
</tr>
<tr>
<td><strong>Meso-level emphases</strong></td>
<td>Interpreting and operationalizing macro-level emphasis by:</td>
</tr>
<tr>
<td></td>
<td>- Effecting the desired epistemological and socio-cultural changes via design research</td>
</tr>
<tr>
<td></td>
<td>- Mapping to effective classroom orchestration and implementation that seeks to achieve the desired micro-level interactions and outcomes, via design research</td>
</tr>
<tr>
<td></td>
<td>- Considering systemic forces and mediating inter-related tensions to lead to sustainability and scalability</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Micro-level actors:</strong></th>
<th>Individuals such as students and teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Micro-level environment:</strong></td>
<td>Interactions or discourse within small group and classroom settings</td>
</tr>
<tr>
<td><strong>Micro-level emphases</strong></td>
<td>Informing macro and meso-level emphases by:</td>
</tr>
<tr>
<td></td>
<td>- Studying contextualized group or classroom-based interactions in an in-depth manner</td>
</tr>
<tr>
<td></td>
<td>- Eliciting feedback from participants</td>
</tr>
</tbody>
</table>

Figure 1. A systemic framework for enabling innovative practices via the alignment of macro, meso and micro levels (adapted from Looi et al., 2011, p. 11)

The socio-cultural factors of the school’s learning ecology constitute the meso-level environment. As researchers from the university, we serve as meso-level actors who work in that environment to recontextualize pedagogic discourse. This re-contextualization process is a “meso-level” mechanism. The orchestration of efforts from all
actors will contribute explanatory power to the sustainability of an intervention. By approaching this pedagogy-driven reform at the macro, meso and micro levels, we seek the alignment of systemic forces at work to provide a buttress for sustainability. Thus we, as researchers as the meso-level actors, help the school practitioners understand and interpret policy imperatives and actualize them into classroom teaching and learning practices in ways that are informed by research and theories. Figure 1 shows the roles of actors in these three levels in enabling and sustaining innovation.

The curricular innovation informed by the seamless learning notion

Through a DBR approach and a lens of researchers serving as meso-level mediators to work with the teachers, we first conducted several research cycles in the three-year research project entitled “Leveraging Mobile Technology for Sustainable Seamless Learning in Singapore School.” The work helped us to develop a viable innovation model (the Seamless Learning Model or SLM in short) by working with a class of primary school students over a period of two school years. The innovation involved the transformation of the existing science curriculum into an inquiry-based one which leverages the affordances of mobile technologies. The project was framed in the broader context of constructing “seamless learning” environments to bridge different learning contexts (such as between formal and informal learning settings, individual and social settings, and learning in physical and digital realms), mediated by mobile devices in 1:1, 24x7 basis (Chan et al., 2006; Wong & Looi, 2011). The basic rationale is that it is not feasible to equip students with all the skills and knowledge they need for lifelong learning solely through formal learning (or any other single learning space); henceforth, student learning should move beyond the acquisition of content knowledge to develop the capacity to learn seamlessly (Chen, Seow, So, Toh, & Looi, 2010). Nevertheless, as such a learning approach is a tall order for students who are more accustomed to the present instructivist-dominated education system. In this regard, we envisaged the design and enactment of long-term seamless learning curriculum where teachers engaged learners in an ongoing enculturation process (Wong, 2013b) in nurturing their habit-of-mind in seamless learning.

We first describe the underlying intent and the guiding principles of this curricular innovation enabled by mobile technologies. These were what shaped the curricular innovation, and they had been iterated and improved progressively through a process of DBR by working with teachers over three years. The process was marked by in-situ iterative and collaborative cycles of co-design by researchers and the teacher, leading to enactment by the teacher in one experimental class with data collection. The researchers observed the classes, and provided feedback to the teacher to improve the design and subsequent enactments. The class was a mixed-ability class.

The design of the learning units in the mobilized curricula for science is designed based on these design principles (Zhang et al., 2010):

- Design student-centered learning activities (to promote engagement and self-directed learning)
- Make students’ thinking process visualizable (so that they can be shared and subject to further refinement)
- Incorporate different learning modalities (to personalize learning)
- Design for holistic and authentic learning (make science learning meaningful)
- Facilitate social knowledge building (to promote collaborative learning)
- Ensure that the teacher plays the role of facilitator (to move away from didactic teaching)
- Provide an environment to integrate all learning activities (students have a hub to launch or continue their learning activities)
- Assess formatively (through the learning activities, students can receive feedback for their own ideas from peers or the teacher)
- Extending classroom learning activities beyond school hours and premises (to support the notion of seamless learning).

In the co-design process, researchers worked with teachers to revise and mobilize two years’ worth of the national curriculum for Primary 3 and 4 Science by considering the opportunities afforded by ubiquitous access to mobile devices. Activities were designed which seeks to extend learning activities beyond the classroom. To support the continuous and long-term learning activities, 34 students from the experimental class were each assigned a smartphone with 24x7 access in order to mediate a variety of learning activities such as in-class small-group activities, field trips, data collection and geo-tagging in the neighbourhood, home-based experiments involving parents, online information search and peer discussions, and digital student artifact creation, among others.
The key epistemological design commitments of the curricular innovation are: learning as drawing connections between ideas, and learning as connecting science to everyday lives, across multiple learning spaces. The curricular commitment is seamless learning, and inquiry-based facilitation and learning. The technological commitments include: technology for construction, technology for communication, and technology for searching information anywhere anytime.

Concerning the curricular commitment, in science teaching, the Ministry of Education of Singapore has advocated the use of the BSCS 5E Instructional Model (Bybee, 2002). This 5Es model consists of the following phases: engagement, exploration, explanation, elaboration, and evaluation. Each phase has a specific function and contributes to the teacher’s coherent instruction and to the learners’ formulation of a better understanding of scientific knowledge, attitudes, and skills. The model is used to sequence the learning activities in a science lesson or over a series of science lessons.

The designed curriculum was developed with the use of software apps on the GoKnow™ MLE (Mobile Learning Environment) that runs on a Microsoft Windows Mobile operating system. The GoKnow MLE enables teachers to create differentiated lessons easily via its online learning management system, GoManage, and it enables students to easily personalize their learning experiences (Looi et al., 2009). MLE supports teachers in creating complete, coordinated, curriculum-based lessons that employ multiple media and applications (e.g., text, graphical, spreadsheet, animations, and the like). It is an environment in which students engage in the specified learning activities and create various artefacts. It includes software tools such as:

- KWL (what do I already Know? what do I Want to know? What have I Learned?) to allow students to learn in a self-regulated way,
- Stop Watch that supports timing of events,
- Sketchy™ as an animation/drawing tool, and
- Picomap™ that allows students to create, share, and explore concept maps.

Typically, the lesson is designed to provide opportunities for the student in the Explore, Explain and Elaborate phases of 5E to use the software tools to do their science inquiry. In the Engage phase, the teacher motivates the inquiry by doing some classroom science demonstration or posing some science questions. In the Evaluate phase, the teacher or student peers review the work done on the software tools to detect and correct students’ developing conceptions of the science concepts.

We designed a total of twelve MLE units in the two years of intervention. Note that we did not design the whole curriculum in one go before the intervention commenced. During and after each design-enactment cycle for a MLE unit, we reflected upon the lessons and apply such understanding to inform the design of the next MLE unit. In addition to offering a logical flow for learning the domain knowledge, we had progressively incorporated various types of inquiry/seamless learning activities, from simpler to more demanding ones. This was to facilitate the students’ gradual changes in their habits of mind moving towards learning seamlessly. For example, while the earlier activities involve the students expressing their understanding using the software tools or capturing artifacts outside of the classroom and relating them to the curriculum concepts, the latter activities involve some form of parental involvement in the students’ learning.

Through a process of considering the space of activities enabled by the affordances of the software tools and the smartphones, and the kinds of activities that would be beneficial for primary school science students, we develop a categorization of the 10 major types of smartphone-mediated activities as shown in Table 1. Table 2 summarizes the essential information, including what smartphone-mediated activities were incorporated, of the twelve MLE units.

<table>
<thead>
<tr>
<th>Activity ID</th>
<th>Activity Type</th>
<th>Mobile Affordances</th>
</tr>
</thead>
<tbody>
<tr>
<td>KWL</td>
<td>Self-regulation of learning progress</td>
<td>KWL</td>
</tr>
<tr>
<td>Anim</td>
<td>Animation creation</td>
<td>Sketchy</td>
</tr>
<tr>
<td>Ph</td>
<td>Photo taking</td>
<td>Built-in camera</td>
</tr>
<tr>
<td>CM</td>
<td>Concept mapping</td>
<td>PicoMap</td>
</tr>
<tr>
<td>Dsc</td>
<td>Online artifact sharing and discussion</td>
<td>Blog / Mobile Forum</td>
</tr>
<tr>
<td>Trp</td>
<td>Field trip</td>
<td>Video, photo &amp; note taking tools</td>
</tr>
<tr>
<td>Exp</td>
<td>Scientific experiments</td>
<td>Video, photo &amp; note taking tools</td>
</tr>
</tbody>
</table>
Table 2. List of MLE units designed for the mobilized curriculum (Wong, 2013b)

<table>
<thead>
<tr>
<th>Level &amp; time period</th>
<th>Topic</th>
<th>Anim</th>
<th>KWL</th>
<th>Ph</th>
<th>CM</th>
<th>Dsc</th>
<th>Trp</th>
<th>Exp</th>
<th>Par</th>
<th>Web</th>
<th>Col</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb 2009</td>
<td>Classification for living &amp; non-living things</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb &amp; Mar 2009</td>
<td>Classification of animals</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar &amp; Apr 2009</td>
<td>Plant</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar &amp; Apr 2009</td>
<td>Plants &amp; their parts</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar &amp; Apr 2009</td>
<td>Fungi</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr &amp; May 2009</td>
<td>Materials</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug &amp; Sep 2009</td>
<td>Body systems</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan &amp; Feb 2010</td>
<td>Cycles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb &amp; Mar 2010</td>
<td>Matter</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 2010</td>
<td>Light &amp; shadow</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr &amp; May 2010</td>
<td>Heat &amp; temperature</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul 2010</td>
<td>Magnet</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The first seven units constituted the Primary 3 mobilized curriculum; the last five belonged to the Primary 4 curriculum.

**Demonstration of efficacy of innovation**

The curricular innovation involves designing a coherent and sustainable classroom program. The unique research context is that we worked with a school with 9 classes in the cohort of Primary Grade 3 taking the science subject. The experimental class followed the same class schedule and assessment schemes as the rest of the classes. What evidence can we demonstrate in establishing the efficacy of the innovation? Our analysis of the science examination scores after the one year intervention (with the experimental class using the mobilized curriculum to replace the traditional curriculum) shows that amongst the 6 mixed-ability classes in Primary (Grade) 3 in the school, the experimental class performed better than other classes as measured by traditional assessments in the science subject (see, Looi, Zhang, et al., 2011). With mobilized lessons, students were found to learn science in personal, deep, and engaging ways as well as developed positive attitudes towards mobile learning. We feel that this result is a very worthwhile contribution to the field, as much research work on mobile learning focus only on units of at most a few weeks duration or they are add-on activities to some existing curriculum.

The use of the mobilized technologies provides many leverage points for the researchers and teachers to co-design a new curriculum that focus on inquiry learning. Once designed, the curriculum can be enacted by science teachers, and it is important for the teachers to understand the design principles behind a mobilized curriculum for inquiry learning and how to implement in the way to harness the best learning outcomes for students. We see a shift in the teacher’s attitudes and behaviors towards science teaching, from a style that sees her pre-occupied with just covering the curriculum to one that allows her to watch over and facilitate students’ work on the inquiry activities on their handhelds.
With the mobilized lessons, we observe students engaging in science learning in personal and engaged ways. They demonstrated their understanding of science phenomenon in multimodal ways and did self-directed learning by doing online search and exploration on questions related to the curriculum topics. They engaged in instructional activities that involve their parents, as in our mobilized lesson for the body systems. This lies in contrast to the more “traditional” way of learning, in which students learned science from the didactic instruction of the teacher or from the textbook.

Providing the basis: Explicating and theorizing the SLM model

The pilot has shown that Seamless Learning (SL) is a viable pedagogical model. SLM harnesses the portability and versatility of mobile devices to promote a pedagogical shift from didactic teacher-centered to participatory student-centered learning (Facer et al., 2004). With appropriate learning design, the mobile technology facilitated the transformation of classroom learning activities into a more student-centered, personalized, and social learning process for the students where they need to process and associate their experiences or the information received in the informal contexts with the knowledge that they have acquired or constructed in the classroom, reflect upon any discrepancy, and apply the knowledge to solve real-life.

As design-based researchers, we also seek to work towards theories for how to get students to learn as self-directed seamless learners (Looi et al., 2010; So, Kim, & Looi, 2008; Wong, 2013b). We explore how the theories and methodology of self-regulated learning (SRL), an active area in contemporary educational psychology, are inherently suited to address the issues originating from the defining characteristics of mobile learning: enabling student-centered, personal, and ubiquitous learning (Sha, Looi, Chen, & Zhang, 2012). These characteristics provide some of the conditions for learners to learn anywhere and anytime, and to be motivated and able to do so in a strategic way, namely, self-regulate their own learning. We propose an analytic SRL model of mobile learning as a conceptual framework for designing and analyzing mobile learning, in which the notion of self-regulation as agency is at the core (Figure 2). The rationale behind this model is built on our recognition of the gaps in the current conceptualization of the mechanism and processes of mobile learning, and the inherent relationship between mobile learning and SRL.

![An analytic SRL model of mobile learning (Sha et al., 2012)](image)

At the center of the model is the notion of self-regulation as agency, referring to the learner characteristics that function as internal driving forces initiating and sustaining a self-regulated mobile learning process. The key personal factors include domain knowledge, prior experiences, motivation, and metacognitive awareness, epistemological beliefs, and so on. Self-regulated mobile learning processes can be understood and analyzed by means of SRL theories and methodologies (e.g., self-report survey, trace analysis). Mobile learning processes that are regarded as manifestation/exercises of agency (fundamentally composed of motivation and metacognition) can be understood, analyzed, and assessed from the theories and methodologies of SRL. Second, mobile learning activities are supposed
to be mediated by mobile technologies and devices, which presumably function as social, cognitive, and metacognitive tools. The existing studies in mobile learning largely focus on the social and cognitive functions but ignore the metacognitive function.

**Methodological contributions: Studying students on the move**

SLM advocates continuous learning by the students inside and outside of the classroom. The curriculum incorporates components where they use their personal mobile devices to do planned activities at home or outside of the classroom, and where they also learn in emergent and unintended ways. The challenge for us researchers is how to study “seamless learning” as synergistic and continuous learning across multiple spaces and time scales which involves theoretical and methodological challenges (Toh, So, Seow, Chen, & Looi, 2013). With students constantly on the move and many of their interactions happening in informal settings, researchers face significant challenges for investigating and documenting emergent forms of learning and participation across multiple contexts. The challenges of collecting data for students on-the-move include making sense of the voluminous data collected, and navigating the complexity of ethical issues involved in the intrusive and non-intrusive data collection methods.

To fill the gap in mobile learning research, we innovated on a methodological approach for researching digital kids’ learning supported by mobile technologies. The specific focus on methodological issues is significant with recent trends in mobile learning research’s focus on learning beyond school settings. Building on earlier work on methodological issues (e.g., Hsi, 2007; Martin, 2004; Vavoula & Sharples, 2009), we adopt and adapt the method of cooperative inquiry to study the sense-making endeavors of digital students who are always on-the-move and learning across multiple settings. We do so by working with parents who would cooperate to collect data such as by using video-recording and taking photos of their children doing things on their mobile devices.

**Moving beyond the Proof-of-concept: From innovation to scaling**

Is the curricular innovation ready for scaling and how do we know to what extent scaling has taken place? One of the most cited literature on scaling is that of Coburn (2003) who defined scale as encompassing four interrelated dimensions:

- **Depth:** Depth refers to deep and consequential change in classroom practice, altering teachers’ beliefs, norms of social interaction, and pedagogical principles as enacted in the curriculum.
- **Sustainability:** Sustainability involves maintaining these consequential changes over substantial periods of time.
- **Spread:** Spread is based on the diffusion of the innovation to large numbers of classrooms and schools which will adopt and adapt the innovations.
- **Shift in reform ownership:** Shift requires districts, schools, and teachers to assume ownership of the innovation, deepening, sustaining, and spreading its impact.

Building on this work, Clarke and Dede (2009) added a fifth dimension, namely, evolution, in which the innovation, as revised by its adapters, is influential in reshaping the thinking of its designers and creating a community of practice that evolves the innovation.

Our 3-year design research study in the primary school has helped achieve some successes in dimensions of scale by Coburn (2003) in four ways:

- **Depth:** The intervention has created positive learning gains for the students of the 2 classes and positive changes in attitudes and knowledge of the 2 teachers (Looi, Zhang et al., 2011).
- **Sustainability:** Clearly changes have occurred in the school with evidences from research analysis (Looi, Zhang, et al., 2011; Zhang et al., 2010) during the two years of intervention, and from interviews with the stakeholders (school leaders and teachers).
- **Spread:** In 2009, we worked with 1 teacher and 1 P3 class. In 2010, we worked with 2 teachers and 2 P4 classes. The school is spreading the mobilized curriculum for science to all P3 classes in 2012, and to all P4 classes in 2012.
- **Shift in reform ownership:** The school has taken over ownership by driving the spread of the mobilized curriculum for science to all P3 classes in 2012.
When the curricular innovation using mobile devices has been developed and studied in the context of one class, and the empirical evaluation of the mobilized curriculum has shown its potential for learning effectiveness, the school leaders decided that it was a worthwhile innovation and, in consultation with the researchers, would like to scale up the innovation.

We envisage the following four levels of scaling in the life-cycle of educational research and development work that lead to practical and policy implications for how to scale up in a school context:

**Level 1:** Developing the intervention as an innovation through a pilot in one or two classrooms

**Level 2:** Grade-level scale—spreading to more classes in a grade level and eventually to a whole grade level

**Level 3:** School-level scale—spreading the innovation to a whole school

**Level 4:** District or country-level scale—in which such work will have policy implications for the educational authorities.

We note that many research interventions do not survive the first level, or they merely go through many research funding cycles to iterate and re-design the intervention, and stop at that. Our interest is those research innovations which have been shown to have demonstrated learning efficacies, and there is purpose and commitment from all the relevant stakeholders to scale-up the intervention. We know it is rare for bottom-up research interventions to move from Level 1 to Level 2, let alone to the other levels. Thus it is critically important to study and learn from the few instances out there that actually move up the levels, and this is towards exemplifying and informing how research can really bridge the research-practice gap by through such scaling studies.

To make SLM an integral part of classroom practices, what would then be useful for the school and for the Ministry of Education is to “ruggedize” the innovation for sustainability to retain substantial efficacy in diverse or even relatively barren contexts. To ruggedize the innovation, robust-design strategies are needed that will enable the innovation to be used in multiple settings (Clarke & Dede, 2009). The school has moved onto the Level 2 and has plans to share the model with other schools.

So far, we have created a successful story of research-informed technology-enabled practices in the school, but we want to study further how to sustain such successful practices over time after an initial influx of resources and other forms of external support. For sustained changes, we would reemphasize the importance of meso-level mechanisms that support teacher capacity building and reinforce school leadership and culture. Barab & Luehmann (2003) discuss issues of sustainability and local adaptation as crucial for scale. They describe the teacher’s role in local adaptation as identifying local needs, critiquing the innovation in the light of those needs, visualizing possible scenarios of implementation, and finally making plans or decisions concerning the implementation. Teachers will be ultimately involved in the adoption, customization, and implementation process, and they are continually remaking and contextualizing the innovation in terms of their local context. Barab and Luehmann (2003) argues that instead of the equation

\[
\text{Designed Curriculum} = \text{Implemented Experience}
\]

it should be

\[
\text{Teachers Perceptions} + \text{Designed Curriculum} + \text{Classroom Culture} = \text{Implemented Experience}
\]

For example, stark differences would exist between the customization and implementation of an innovation by a subject-matter focused teacher who is very concerned about students’ understanding of the content compared to a logistically focused, teacher who is most concerned that the activities run smoothly and in the appropriated time frame. Our current phase of research studies how different teachers would locally adapt the innovation and what the resulting implemented experiences for the students and for the teacher would be.

The critical research questions are:

1. How to adapt and to “ruggedize” the innovation for sustainability to retain substantial efficacy in diverse contexts of more classes in the grade level and more teachers, in which some of the conditions for success are absent or attenuated? This includes designing a more device-independent curriculum intervention for mobile learning. By clarifying the design principles and the design affordances, it is hoped that the new curriculum development model is more generic and less dependent on the devices the students use.

2. What is the impact on depth, sustainability, spread, shift of ownership and evolution when a school takes over ownership and scales up an innovation developed from one class to more classes and eventually to a whole grade level?
3. What are the strategies for curriculum development and professional development models needed to support the spread of a mobilized curriculum to one whole level?
4. What are the strategies for organizational, technological and institutional changes needed for such a sustainable and scalable translation of research into practice?

A key issue in scaling is how to explicate the curriculum for the typical teacher. While during a pilot, one can leave both the content and the instructional strategies “looser” and “less-defined,” that strategy will not work when all teachers read, use and interpret the same curriculum. The curriculum needs to be specified (Cohen & Ball, 1999). The goal is to make the content and the instructional strategies explicit – make the curriculum transparent for the teacher to enact initially and to adapt subsequently.

Another key issue relates to ongoing professional development. Teachers need to be provided with continuous support as they transition to a new curriculum, providing the performance support (Cohen & Ball, 1999). In contrast to a pilot, where the teachers are typically highly motivated and are top-notch teachers, as an innovation goes to scale, all teachers must be brought up to speed. Some of those teachers will be motivated and some less so; some teachers will be highly competent and some less so. Thus, professional development that is ongoing, continuous must be put in place to help all the teachers, especially the weaker ones, understand how to rollout the innovation.

The teachers need more collaborative work sessions so they can help each other with suggestions on instructional strategies and with tweaks to the curriculum. Those additional collaborative work sessions are critically important. If teachers feel isolated, they will not enact the innovation. They need to form their own learning community for mutual support and rely less on the research team. Moreover, the fundamental issue is to shift the teachers’ epistemological and educational beliefs from being a transmissionist and behaviorist to being a socio-constructivist – without the shift, all the curricular mobilization efforts will become a mere formality.

Aligned with the issue of teachers adopting and adapting a new pedagogy, there is a need to make the formative evaluation techniques explicit and to show teachers how to use the formative assessments in order to tailor their instructional practices to the specific needs of the individual students. In order to go to scale, all teachers need to use the same formative assessment techniques and thus these techniques must be made explicit and teachers must be given support as they learn how to administer and use those formative assessments.

The curricular innovation is enabled by mobile technologies. A more sustainable approach is to adapt the curriculum for a more generic mobile technology. This makes the curriculum and the formative assessments work with a broader range of mobile technologies and applications. While initially the innovation used smartphones, the goal is to create materials – curriculum, instructional strategies, and formative assessments that are mobile technology agnostic. Mobile technologies are changing very quickly; thus, we do not want our learning resources to be tied to a specific mobile technology. Moving towards the blending of mobile and cloud computing technologies (Wong, 2012) is a direction in our agenda for scaling up. The new direction will not only offer a feasible solution to the above-stated challenge of changing technologies, but also has the potential to open up new opportunities for developing more advanced affordances to mediate a wider range of seamless learning activities.

What have we learned about scaling so far?

At the school, we have moved from serving 80 students (2 teachers) to serving 320 students (6 teachers). That almost order-of-magnitude jump in who is supported requires R&D if that transition is to be effective. What we learned in the pilot was critically important, e.g., that SLM can be an effective pedagogical model in supporting students as they engage in inquiry-based learning. However, in going to scale, the issue is no longer one of efficacy but one of infrastructure – how do we take an innovation that was hand-crafted and make it more rugged, more robust, more stand-alone, more transparent.

By the end of academic year 2012, all teachers of P3 have enacted the curriculum. What have we scaled up to enable this spread to all teachers and all classes in P3? The scale-up comprises these multiple dimensions:
1. Mobilized curricula (to lead students to self-directed learning and to bridging informal learning spaces)
2. Teacher facilitation skills
3. Teacher readiness
4. Student readiness including hardware and software training
5. Technology infrastructure, e.g. WiFi and 3G Connectivity; availability of mobile devices in 1:1, 24x7 basis

On assimilating the curriculum into the classroom culture (Squire, MaKinster, Barnett, Luehmann, & Barab, 2003), we summarize the challenges the teachers faced in doing a new curricular innovation that builds new skills and competencies beyond what is usually assessed in the standardized assessments used in the school. In the researchers’ interviews with the teachers, the teachers raised these concerns:

• They are not sure if they are conducting the designed curriculum lessons in the right way. Some teachers expressed doubts on the students’ ability for doing self-directed learning.
• Some teachers expressed that they needed help in developing and practicing questioning skills in the classroom.

What support was provided to the teachers to help them to address these challenges? The response to the first concern was for the researchers to provide support to the teachers by giving personalized feedback after each lesson enactment, and by helping the teachers to adapt the lessons for different ability students. The response to the second concern was for teacher sharing facilitated by the researchers in the weekly curricular design meetings to discuss questioning skills via modelling by a teacher or researcher, and by lesson study discussions on a recorded classroom sessions. While researchers provided these initial scaffolding, the plan was to build up the capacity of this group of primary 3 science teachers so that they can sustain the innovation in the coming years with fading from the researchers in the subsequent year.

In levelling up the capacity of the teachers, some new activities were planned. Arrangements were made for teachers to observe their peers conducting the lesson activities. There were some discussions for each teacher to co-teach with another teacher, but this was later not followed up on because of time constraints and time tabling issues (for example, when the teachers are teaching concurrently thus making it difficult to visit other classes for co-teaching). The researchers decided to edit short clips of videos which they found out to exhibit good teaching and learning scenarios to be shared with other teachers.

**Scaling to Chinese and English language learning**

In previous sections, we reported the outputs of our past research – the explication of the SLM, the curriculum development principles, the research methodology to study learners on the move, the PD model, our experience and understanding in fostering the technological and systemic conditions for the establishment of a sustainable 1:1 seamless learning environment, etc. Apart from the scaling up of the science curriculum, these research findings and deliverables are also serving as the basis for scaling the SLM to other school subjects. To date, two language subjects, namely, Chinese and English, have embarked on their respective journeys of transforming their curriculum into seamless learning environments, again with the eventual aims of cross-school scaling up.

“MyCLOUD” (My Chinese Language ubiquitOUs learning Days, *语飞行云*) (Wong, Chai, Chin, Hsieh, & Liu, 2012) is a levelling up effort of the completed study of “Move, Idioms!” (Wong, 2013a; Wong, Chin, Tan, & Liu, 2010). The DBR study aims to develop a holistic and scalable mobile- and cloud-assisted Chinese Language (CL) seamless learning environment that is informed by language learning/acquisition theories for P3-P5 students. The project involves a 2 ½-year school-based intervention (August 2011-November 2013). The intention is to enculturate students in carrying out CL vocabulary learning and communicative writing activities that encompass formal and informal learning settings. Under this approach, students are assigned mobile devices on 1:1, 24x7 basis in order to stimulate and support their language learning both within and beyond the classrooms. In addition, a cloud-based, device-independent MyCLOUD learning platform leveraging mobile and cloud computing technologies has been developed for the purpose. Another aim is to design new classroom practices that will be integrated into the existing formal curriculum as well as foster students’ competency to engender deep CL learning. The learning design principles developed by the researchers specifically foreground the bridging of language input (classroom, textbook-induced learning, and sharing of students’ linguistic artifacts) and output (students’ daily use of the target language through creations of linguistic artifacts such as photo taking / sentence making and social networking), and the bridging of learning contextualization and learning generalizations/reflections. On top of the more generic 10 features/dimensions of seamless learning as proposed by Wong and Looi (2011), the additional language learning-specific SLM dimensions rooted in various language learning theories may become the basis of the our future exposition of a Seamless Language Learning model. In addition, the research team is working towards spreading the curricular innovation to four more schools by 2014.
Aligned with the goal of implementing seamless learning and cultivating self-directed learners, smartphones were adopted in a trial design and implementation of the Primary 3 English language curriculum at Nan Chiau Primary School. This “WE Learn Project” is a scaling up of the seamless learning initiative in Primary 3 Science. By transforming the classroom from the traditional teacher-centered model to a learner-centered one, the project hopes to enhance the learning outcomes of students. In “mobilizing” the English curriculum and making it inquiry oriented, the teachers and curriculum developers drew on the two pedagogical strategies of P4C and Marzano’s P4C (Philosophy for Children), (Lipman, 1980) draws on the Socratic method of learning pioneered initially in Plato’s dialogues and focuses on learning how to ask a question and how to respond when asked a question. Marzano’s 6-steps to Better Vocabulary Instruction, (Marzano & Pickering, 2005) help children understand words by building relationships and links amongst the words, by using words in their proper contexts.

In science, in CL and the English language learning, while the particular styles of pedagogy have their differences, the pedagogies are, at their roots, inquiry-oriented, and learn-by-doing pedagogies that reinforce each other (Norris et al, submitted). In all these three subjects, the affordances of the mobile technology enable the students to use multiple modalities and multiple media, and to carry and use the smartphones anywhere and anytime as a learning hub.

At this point of writing (in November 2012), there are now approximately 350 P3 (3rd grade) students using mobile computing devices daily for science. Out of these 350, 120 students are also using the devices for English and Chinese language (MyCLOUD) learning. The school has plans to scale to approximately 700 students in P3 and P4 in science, to 350 students on P3 English, and 350 students in Chinese language learning.

**Conclusion**

Within the educational research community, many research projects focused on designing or establishing the efficacy of innovations that work well within specific contexts. They typically face the conundrum of narrowing the research-practice gap when it comes to changing or transforming practices in schools and other contexts for learning, and to scaling up to meet the needs of a broader audience. Such research projects are also not organized to address the challenge of long-term systemic improvement as research-practice partnerships often go in tandem with short-term funding of grants and program initiatives at foundations and government agencies. This paper describes a multi-year research programme for doing scaling and implementation research. The agenda is contextualized in the example of a mobile learning curricular innovation in a Singapore school as it goes to scale.

In this curricular innovation on mobile learning, we established that SLM can raise student achievement in the context of one class and one teacher. We have developed models for:

1. Transforming a curriculum to harness the affordances of mobile technologies
2. Technology infrastructure and support for a mobilized curriculum
3. Teachers’ professional development through working with 1 teacher for a grade level 3 (P3) class in 2009 and 2 teachers for two grade level 4 (P4) classes in 2010.

Because SLM demonstrated increased student achievement, the school decided to scale-up the roll-out of the transformed curriculum to all the eight P3 classes, by doing the planning in 2011 and doing the scaling in 2012. The next step in the research and implementation trajectory brings us to scaling research which seeks to make research count in practice. The goal is to study the adoption and the adaptation of the curricular innovation as it goes to scale to more classes, more levels and more subjects, and to documenting the benefits of balancing fidelity of implementation with adaptation to dynamic local contexts. The programme of research enables us to articulate the SLM curriculum, the supporting resources, the teacher learning and professional development models with analysis of their impact, efficacies as well as weaknesses. The scaling research will establish a model of scaling that recognizes the range and diversity of teachers’ local needs as well as the necessary adjustments they need to make in order for the innovation to be useable and effective, and how the school can support them to know how to adapt the innovation and yet retain the essence of its efficacy.

We hope that our narration of the ongoing research journey from innovation to practice and to scale can inspire other research initiatives that will address the multi-term, multi-pronged, multi-level and systemic aspects of school-based
innovations, and that yet at the same time, advance theory, frameworks, design principles, resources and strategies for effective and sustainable mobile learning.

Acknowledgements

The entire seamless learning program reported in this paper consists of several projects funded by Interactive Digital Media (IDM) Program (project number: NRF2007-IDM005-MOE-008 LCK); Office of Education Research, National Institute of Education, Nanyang Technological University (project numbers: OER 26/12 LCK, OER 61/12 WLH, OER 16/10 LCK, OER 17/10 LHW); and EduLab Program (project number: NRF2011-EDU002-EL005 (CI2)) respectively.

References


Ubiquitous Learning Project Using Life-logging Technology in Japan

Hiroaki Ogata*, Bin Hou2, Mengmeng Li2, Noriko Uosaki3, Kosuke Mouri2 and Songran Liu2

1Faculty of Arts and Sciences, Kyushu University, Japan // 2Faculty of Engineering, University of Tokushima, Japan // 3Osaka University, Japan // hiroaki.ogata@gmail.com

*Corresponding author

ABSTRACT

A Ubiquitous Learning Log (ULL) is defined as a digital record of what a learner has learned in daily life using ubiquitous computing technologies. In this paper, a project which developed a system called SCROLL (System for Capturing and Reusing Of Learning Log) is presented. The aim of developing SCROLL is to help learners record, organize, recall and evaluate ULLs. Using SCROLL, learners can not only receive personalized quizzes and answers to the questions, but also navigate and be aware of their past ULLs supported by augmented reality views. In particular, this paper introduces an approach that helps learners record their learning experiences in daily life from life-log photos with the help of SenseCam. To evaluate the effectiveness of this system, a case study of an undergraduate English course is presented to show how it can be used to facilitate seamless learning.

Keywords

Mobile learning, Ubiquitous learning, Learning log, Life logging

Introduction

CSUL (Computer Supported Ubiquitous Learning) or context-aware ubiquitous learning (u-Learning) is defined as a technology-enhanced learning environment supported by ubiquitous computing technologies such as mobile devices, RFID tags, and wireless sensor networks (Ogata & Yano, 2004; Wu, Hwang, & Chai, 2013). It is characterized by its augmented learning environment which presents information on personal mobile devices through the Internet based on the detection of physical objects in surrounding environment using sensing technologies (Hwang, Tsai, Chu, Kinshuk, & Chen, 2012).

The fundamental issues of CSUL are:

1. How to record and share learning experiences that happen anytime and in any place.
2. How to retrieve and reuse them for future learning.

To tackle these issues, LORAMS (Linking of RFID and Movie System) (Ogata, Matsuka, El-Bishouty, & Yano, 2009) was proposed. There are two types of learners in this system. One type consists of providers who record their experiences on video. The other type are learners who, when encountering some problems in their learning, may find the videos uploaded by the first groups of learners useful. The system automatically links between physical objects and the corresponding objects in a video and allows sharing among users. By scanning RFID tags, LORAMS shows users the video segments that include the scanned objects. Although this system is useful in certain environments, it is not currently easy to apply in practice. Therefore, we have begun a more practical research project called “ubiquitous learning log” (ULL) in Japan in order to store intentionally what learners have learned as ubiquitous learning log objects (ULLOs) and consequently reuse them. The project was conducted from October 2009 to March 2013 with the financial support by PRESTO of the Japan Science and Technology Agency (JST).

We define a ubiquitous learning log (ULL) as a digital record of what a learner has learned in daily life using ubiquitous technologies, and propose a model called LORE to show the learning processes from the perspective of the learner’s activity (Ogata et al., 2012). In this paper, we propose a system called SCROLL (System for Capturing and Reusing Of Learning Log) that helps learners log their learning experiences with photos, audios, videos, locations, QR-codes, RFID tags and sensor data, and share their ULLOs with others. Also, a learner can receive personalized quizzes and answers to their questions. This system is implemented both on the web and on the Android smartphone platform. With the help of built-in GPS and cameras in smartphones, learners can navigate and be aware of past ULLOs via the augmented reality view.

Originally, the term “learning log” was used for personalized learning resources for children. The logs were usually visually written notes of learning journals, which could become an integral part of the teaching and learning program
and which had a major impact on their drive to develop more independent learners. Research findings indicated that journals were likely to increase meta-cognition and reflective thinking skills through students becoming more aware of their own thought processes (Hung et al., 2014; Hwang, Wu, Zhuang, & Huang, 2013; Stockwell, 2007; Susan & White, 1994; Wood Daudelin, 1996). Our approach focuses on how to enrich learning logs and promote retention and meta-cognition by using mobile, ubiquitous and context-aware technologies.

Life-logs

Life-logs are a notion that can be traced back at least 60 years (Bush, 1945). The idea is to capture everything that ever happens to us, to record every event we experience and to save every bit of information we ever touch. For example, SenseCam (Hodges et al., 2006) is a sensor-augmented wearable still camera which is proposed to capture a log of the wearer’s day by recording a series of images and capturing a log of sensor data. This is a great tool for recording life logs, as it is a small digital camera that is combined with a number of sensors to help capture a series of images of the wearer’s whole daily life at the proper time, and can be worn around the neck (Figure 1). Originally this device was designed as a memory aid.

![Figure 1. SenseCam](image)

MyLifeBits (Gemmell, Bell, & Lueder, 2006) stores scanned material (e.g., articles, books) as well as digital data (e.g., emails, web pages, phone calls, and digital photos taken by SenseCam). The Ubiquitous Memory system (Kawamura, Fukuhara, Takeda, Kono, & Kidode, 2007) is a life-log system using a video and RFID tags. Another application, Evernote (www.evernote.com), is a tool to save ideas using mobile devices such as Android and iPhone. The common idea of these projects is to use life-log data as a memory aid. SCROLL, however, aims to utilize life-log data for the learning process.

**SCROLL**

*Design*

In this paper, a ubiquitous learning log (ULL) is defined as a record of what a learner has learned in daily life using ubiquitous technologies. ULL is considered as a set of ULLOs (Ubiquitous Learning Log Objects). The learning can also be considered as the extraction of meaningful knowledge from past ULLs that serves as a guide for future behavior (Wood Daudelin, 1996). Figure 2 shows the learning processes from the perspective of the learner’s activity model called LORE (Log-Organize-Recall-Evaluate). These four steps are explained as follows:

1. **Log what the learner has learned**: When learners face problems in daily life, they may learn some knowledge by themselves, or ask others for help. The system records what is learned during this process as a ULLO. We designed two modes to record learning contents – active and passive. Here is a typical scenario of the active mode – when a foreign student in Japan walks into a supermarket, there are many foods that he/she does not know how to say in Japanese. The student can take a photo of this food by Smartphone and ask someone how to say it in Japanese, then log the learning content as a ULLO including the photo, the name of the food in both Japanese and the learner’s mother language, location, time, etc. However, in the passive mode, a device such as a life-log camera can take photos of food and record the contextual information such as location and time automatically and wait for the learner to review the recorded contents before logging them as ULLOs.

2. **Organize ULL**: When a learner tries to add a ULLO, the system compares it with other ULLOs, categorizes it and shows similar ULLOs if they exist. There are many ways of categorizing ULLOs. For example, a foreign student
in Japan learned a new word “tofu” in a supermarket and logged this process as a ULLO. This ULLO can be categorized as “Japanese,” “Food,” “Supermarket,” etc. As such, it is difficult for the system to categorize it. Therefore, in the designed system, users can add their ULLOs into multiple categories and add several tags to each one. After that, they can review the learned contents by category/tag. Similar objects can be found by matching titles, content of photos, locations and categories, and then the knowledge structure can be regulated
and organized.

3. Recall ULL: Learners may forget what they have previously learned. Rehearsal and practice in the same or another context in idle moments can help to recall past ULLOs and to shift them from short-term to long-term memory. Therefore, the system provides some quizzes and reminds the learners of their past ULLOs.

4. Evaluate: It is important to recognize what and how learners have learned by analyzing their past ULLs, so that they can improve what and how to learn in the future. Therefore, the system refines and adapts the organization of the ULLOs based on the learners’ evaluation and reflection.

All of the above learning processes are supported by SCROLL.

We designed SCROLL to implement several types of learning, including self-directed and personalized learning, reflective learning, collaborative learning, situated learning, experiential learning, and seamless learning.

Self-directed and personalized learning

The first kind of learning is self-directed and personalized learning. We designed SCROLL based on the following two objectives that enable self-direction and personalization:

- The system can be aware of a learner’s current context. Currently, the context includes location and time. For the location information, the system can detect whether learners are near the place where they uploaded a learning log and whether there are location-based learning logs recorded by other learners nearby. If either requirement is met and if the availability of the device is high, the system will present a quiz based on the knowledge gained in that location or notify the user of the surrounding learning logs added by others.

- The system can record the context data when learners use the system as their context history and then detect their learning habits by analyzing their context history. If the system detects learning habits, and the circumstance meets these habits, it will issue a recommendation message to encourage the user to review what he/she has learned.

Reflective learning

An important goal of the SCROLL system is to help learners recall what they have learned after they have archived their learning logs. When a learner captures a learning log, in addition to its location-based property mentioned
above, a number of things are designed as retrieval cues for the learner. For instance, according to the picture superiority effect (Defeyter, Russo, & McPartlin, 2009; Shepard, 1967), learning logs with pictures are much more likely to be remembered than those without. In addition, according to the basic research on human learning and memory, practicing retrieval of information (by testing the information) has powerful effects on learning and long-term retention. Moreover, compared with repeated reading, repeated testing enhances learning even more.

For the above two reasons, taking advantage of photos, locations and so on, the quiz function is proposed. Three types of quizzes can be generated automatically by the system: yes/no, text multiple-choice and image multiple-choice.

Usually, learners can examine themselves by taking quizzes (Hwang, Tsai, & Yang, 2008), but two more ways that are instigated by the system are provided. One is that when a learner moves to the place where he/she captured the knowledge, the system can present quizzes about the learned knowledge. The other is that if learners have learning habits, the system will prompt them to review what they have learned using quizzes. These two methods are discussed in detail later in the paper.

Collaborative learning

We designed SCROLL to also encompass collaborative learning. Since learning logs are logs registered arbitrarily by each learner, collaborative learning in SCROLL adopts an asynchronous model. Any learner in this system is able to share ULLOs, and the system will show their shared ULLOs to others. Besides, they can also ask others questions about their shared ULLOs. In reflective learning, shared ULLOs can also be used to generate quizzes in order to help learners learn more objects.

Situated learning and experiential learning

According to Lave & Wenger (1991), situated learning is learning that takes place in the same context in which it is applied. Itin (1999) defined experiential learning as “learning from experience.” We introduce a concept called "task" in SCROLL to implement situated learning and experiential learning. Learning in the same context enhances the learning effect, and past experiences help learners learn effectively. Tasks refer to the activities through which they can acquire knowledge. Tasks are conducted in the circumstances where learning can happen such as in a school, hospital, post office and so on. For instance, if the system recommends the Japanese word “トマト (tomato)” to a learner in a supermarket, learners can talk with the staff in the supermarket using the word “トマト (tomato),” such as asking its price, location, related recipes and so on. It has been proved that by talking with a native Japanese speaker using the recommended word, learners can master the word well (Jonassen & Grabowski, 1993). The activity of asking about the information is a kind of “task.” Learners who save learning logs are responsible for providing what kind of knowledge can be gained by carrying out the task, and one learning log can be used in several tasks. Moreover, the system provides some predefined tasks in different contexts in order to reduce the learners’ burden when they save their learning logs. In addition, the tasks can be created by the learner and designated by the administrator of the system.

The system assigns an appropriate task for a learner according to the difficulty level of the task and the learner’s ability. For example, asking the price of the product is easy for learners, while asking about vegetable recipes is quite difficult for most learners. When learners receive the recommended learning log and the task, they are also asked to provide feedback for the system. For example, they are asked to take photos of the target object if the learning task is to find where the object is. Moreover, if the learning task is to learn about the place of the object, they need to collect and fill in the environmental information on the system. Only by providing feedback can the users prove that they have really gained the knowledge. Moreover, if the learners meet new problems when carrying out the tasks, they can record them in photos, videos, audios or texts and upload them to the system in order to ask for help. Such accumulated data is also meaningful for other learners.
Seamless learning

Recent progress in mobile and wireless technologies offers us a new learning environment, namely “seamless learning” (Wong & Looi, 2011). It allows learners to learn anytime, anywhere, and provides them with multiple ways of learning throughout the day. By seamless learning, we mean learning which occurs with seamless transitions between in-class and out-of-class learning (Hung et al., 2013). The American College Personnel Association (1994) has indicated the importance of linking students' in-class and out-of-class experiences via providing seamless learning environments to achieve academic success.

Based upon the above ideas, we designed the following Seamless Mobile-Assisted Language Learning Support System (hereafter called the SMALL System) (Uosaki et al., 2012) as a sub-project. The main objective of SMALL is to link learners’ out-of-class learning to their in-class learning. Once a learner uploads a newly learned word to SCROLL, our main system, SMALL, runs a search through the previously updated textbook data. If the new word is found in the textbook data, it jumps to the textbook page where this word is used. Another example is that when a user reads an uploaded textbook and clicks a word, then it jumps to the SCROLL system page to show how other learners have learned this word in different contexts in their out-of-class learning. In this way, users’ out-of-class and in-class learning can be intertwined. We learn words from contexts. In order to master words, it is important to come across them used in various situations.

System interface

We implemented SCROLL both on web and smartphone platforms. It consists of the following components:

ULL recorder

This component facilitates the way learners upload their ULLOs to the server whenever and wherever they learn. As shown in Figure 3 (1), in order to add a ULLO, a learner can take a photo, ask questions about it and attach different kinds of meta-data to it, such as its meanings in different languages.
ULL finder

If a learner registers a new ULL, the system checks whether the same object has already been stored by comparing the name fields of each object using a thesaurus dictionary. Also, a learner can search ULLs by name, location, text tag and time. Using this function, learners can understand what, where and when they learned before. Figure 3(2) and Figure 4 (left) show the list of the learner’s ULL, which helps him/her to recall all of the past ULLs. Besides, it allows the learner to be aware of others’ learning objects and to re-log them if deemed useful. This means that a learner can make a copy of another learner’s learning object into his/her own log. Therefore, learners can obtain a considerable amount of knowledge from others even though they have not experienced that knowledge themselves. By sharing ULLs with other learners and relogging the other learners’ ULLs, the acquisition of knowledge is enhanced.

ULL reminder

As shown in Figure 3(3) and Figure 4(right), the system generates simple multiple-choice quizzes based on the meta-data of the stored ULLs. For example, the idea of “quiz with image” is to ask a learner to choose an image that describes the word given by the system. The system immediately checks whether the answer is correct or not. These quizzes are generated according to the user’s profile, location, time and the results of past quizzes and helps learners to recall what they have learned (Li et al., in press). The quiz function is designed not only to help learners to reinforce what they have learned, but also to recommend what other learners have learned and to remind them of what they learned in the past according to their current location and their preferred time. In order to achieve these targets, they can take quizzes whenever they want. In addition, they can send their location information to the server continuously. Therefore, the server side can automatically assign quizzes for them based on their location and time information. It notifies them by showing an alert message and vibrating the mobile phone. Whenever they move around an area where they have encountered some objects, the system will send them quizzes regarding those objects. Furthermore, they can set a time schedule to receive the reminder quizzes.

ULL navigator

The ULL navigator provides mobile augmented reality that allows the learner to navigate through the ULLs. Like Wikitude (“Wikitude,” n.d.) and Sekai-Camera (“Sekai Camera Web,” n.d.), it provides a learner with a live direct view of the physical real-world environment augmented by a real time contextual awareness of the surrounding objects.
When learners are moving around with their mobile phones, the system sends alerts to the phone as soon as they enter the region of ULLOs according to the GPS data. This view is augmented, associated with a visual compass, and overlapped by the nearest objects in the four cardinal directions. Also, it provides the learners with a list of all surrounding objects. When the learner selects one or more of these objects, the Google map will be retrieved and marked with the learner’s current location and the selected objects. Moreover, the system shows a path (route) for the learner to reach the locations of the objects. This assists the learner in acquiring new knowledge by discovering existing ULLOs and recalling the learner’s own ULLOs.

An important component of the system is the ULL recorder. The current ULL recorder requires learners to capture learning contents (e.g., photos) manually. They might find this troublesome and it might disturb their learning activities. Thus, we have tried to find a better way to log learners’ learning content automatically and unconsciously. We found that SenseCam is able to do this. However, the second problem is: Among the very large amount of captured photos by SenseCam, what are the learning contents? After using SenseCam, we found that there are many kinds of context data such as temperature and brightness that can be used to help learners recall the captured objects. Besides, these data can also be used to help us improve our ULL reminders. This is an extra benefit of the passive capture of data.

**PACALL**

**Design**

Until now all the work that we have done has been using the active rather than the passive logging mode. This means that learners must record their learning experiences consciously. Compared to the passive mode, in the active mode we are more likely to miss learning chances since we are not necessarily able to record what we have learned, or sometimes we just forget to record it. Therefore, we introduced passive capture in our project with SenseCam and named the proposed system PACALL (PAssive C Apture for Learning Logs). In the real world, there are so many things that we have learned but we usually miss the chance to review them; that is, we do not know what we know. Similarity, it is certain that we are not able to learn what we have not noticed. Therefore, we considered this in the learning process.

Since this research is based on our previous work (Ogata et al., 2010) in which we used the active mode to register ULLOs, we need to make it clear how the passive mode differs from the active mode. We compare the features of both in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Passive</th>
<th>Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of photos taken</td>
<td>Many (~3000/day)</td>
<td>Few (&lt;1/day)</td>
</tr>
<tr>
<td>Data quality</td>
<td>Poor (by SenseCam, image only)</td>
<td>Good (by Camera /Smartphone/Tablet PC, image with GPS, Video etc.)</td>
</tr>
<tr>
<td>Recording time distribution</td>
<td>Continuous</td>
<td>Intermittent</td>
</tr>
<tr>
<td>Content completeness</td>
<td>High (user’s whole daily life)</td>
<td>Low (only specific scenes)</td>
</tr>
<tr>
<td>Consciousness</td>
<td>Unconscious</td>
<td>Conscious, Intentional</td>
</tr>
<tr>
<td>Reflection</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>Workload</td>
<td>Low (captured automatically)</td>
<td>Quite high (capture manually &amp; upload manually)</td>
</tr>
<tr>
<td></td>
<td>High (review &amp; upload a large number of photos)</td>
<td></td>
</tr>
</tbody>
</table>

In the first three rows of Table 1, the two modes are compared in terms of photos. When we use SenseCam, it takes photos automatically and continuously, while in the active mode, smartphone photos can only be taken at the time we intend to. As a result, more photos are taken in the passive mode than in the active mode. However, because of the storage problem and some other technology limitations, photos taken by SenseCam are lower in quality; they are however of sufficient quality to be used as ULLOs.

In the next two rows of Table 1, the comparison is made in terms of learning contents. When we use a camera or smartphone, many learning contents are logged in our spare time such as at lunchtime. However when we use...
SenseCam, because the recording is processed continuously, we can get photos whenever and wherever we are. As such, the photo contents cover our complete lives. The content type in the active mode, however, is richer than that in the passive mode because in this research the learning content captured by SenseCam consists only of photos.

In the last two rows, the comparison is related to the learners. In the passive mode photos are taken unconsciously, while in the active mode the learner must take photos consciously. When learners use SenseCam, they must review the whole learning process, and reflect on what they have seen and what they have learned and what they missed learning. This process will help them to remember the contents.

According to this comparison, we can see that the passive mode has many advantages over the active mode for language learning by photos. The quality of the photos is low, but it is still acceptable for our purpose. However the biggest disadvantage is the workload. SenseCam takes photos continuously. Consequently a huge number of photos are produced, and the more photos, the heavier the workload. If this workload is reduced, learners can learn language in the passive mode more easily. This is the key issue of using the passive mode in language learning. In this research, we focus on reducing the workload when reviewing the photos, and propose a system that can filter the photos to help learners review and upload ULLOs easily.

Figure 5 explains this process and shows how to support learning in the passive mode. We classify all the objects surrounding us into four groups – “(I) I know what I know,” “(II) I know what I don’t know,” “(III) I don’t know what I know,” and “(IV) I don’t know what I don’t know”. For example, for non-English speakers, when a learner walks outside and sees a fire hydrant, if he notices it and knows how to say it in English that is status (I). If he does not know how to say it in English, that is status (II). Since he has noticed it, and does not know how to say it, he can learn it from a dictionary or by asking someone else. Then the status will change from (II) to (I), which only happens consciously in active mode.

Another situation is the case when he did not notice it. If he has not noticed the fire hydrant how can he learn it? The answer is that he can’t without some form of assistance. Therefore, we would like to encourage learners to use passive mode to support their learning. There are also two kinds of “don’t know” situation. In one case, the learner already knows how to say it in English (status III). In this case, captured life-log photos can help him notice this fire hydrant and let him review it. In status IV, captured life-log photos let him know there is an object that he does not know, and then he can have a chance to learn this object (B to C). This is a good way to help learners become aware of what they do not know if they do not notice it.

This system is a sub-system of the Ubiquitous Learning Log, named PACALL. It stands for PAssive CApture for Learning Log. We set the whole capturing process in passive mode. After that, PACALL analyzes all of the captured photos and finds several important ones to help learners determine which are worth recording.
Figure 6 is the flow of PACALL when analyzing captured photos. It consists of 5 steps: Loading raw data, Filtering bad photos, Finding good photos, Photo recommendation, and Learning analytics. These steps are introduced in detail in the following.

**Loading raw data**

There are three types of raw data in PACALL: life-log photos, sensor data, and GPS data. Life-log photos are currently captured by SenseCam. In the future, we plan to apply this system to photos taken by smartphones or compact digital cameras which are far more commonly used. That will be more convenient and useful. Suppose a learner took a trip and took many photos. Then she can use this system to find photos which contain useful learning contents.

The sensor data are recorded by SenseCam, and the GPS data are created by the portable GPS unit.

**Filtering bad photos**

In this research, a bad photo is defined as a photo that is hardly recognizable or that is a duplicate of other photos. We define three types of bad photos:
- Dark: a dark photo taken with insufficient light.
- Duplicate: the photos are duplicated.
- Defocused: the photos are blurry and cannot be recognized well.

We use image processing to identify these bad photos. Currently, we are using OpenCV to detect dark photos, and LIRE (a plugin for Lucene) to detect duplicated photos.

**Finding good photos**

A good photo is defined as one that contains clear objects. We use OpenCV to find good photos mainly by feature detection.

After filtering bad photos and identifying the good photos, the rest are of mediocre quality. Those photos might contain learning contents, although they are not so clear. The top priority is given to good photos and then mediocre ones come next when shown to learners to choose.
**Photo recommendation**

Once photo sorting is finished, the next stage is learning assistance. Our challenge is to detect useful information from photos by machine, and recommend photos that contain information. We define four types of recommended photos:

- **Character photo**: a photo that contains characters. These characters are possibly used as learning contents. Here we are using text detection to find these photos.
- **Face photo**: a photo containing a face. Actually, these photos are usually not appropriate for learning content because of privacy issues. However, faces are also information from photos.
- **Taggable photo**: a photo that can be tagged by text. Tags are an important piece of information and can be used as a title of the photo.
- **ULLO-like photo**: If there is a similar photo that was already registered to the SCROLL as a ULLO, it can possibly be used as a ULLO as well.

**System interface**

In the previous section, we introduced our design of the flow of analyzing photos. In this section, we explain its functionalities (i.e., PACALL Uploader, PACALL Browser and PACALL Recaller) in detail. PACALL Uploader helps learners upload all the photos after capturing. We have made it easy to upload all the captured photos to the server. Because of the limitation of web technology, this process was not easy in the past. However with HTML5, it became possible. When learners want to upload a whole folder, they can select a photo folder and upload all the photos to the server. Also, the file of sensor data and GPS data will be uploaded.

After uploading the raw data (photos, sensor data and GPS data), the system will analyze all the data and show the results.

![Figure 7. Interface of PACALL Browser](image)

When all the photos are uploaded to the server, the learner can reflect on them with help from PACALL. The PACALL Browser has an interface for browsing all the photos, and it tags photos and provides some information to
help the learner find important photos (Figure 7). Currently, we provide three main functions in the browser – PACALL Filter, PACALL Searcher and PACALL Recognizer.

PACALL Filter classifies all the photos into categories such as Manual, Normal, Duplicate, Dark, Face and Recommendation. Here “Manual” means that a photo is taken manually by pressing the manual button of SenseCam. It usually happens when a learner finds something valuable to record. “Duplicate” and “Dark” mean bad photos. “Face” means the photos that contain faces, and “Recommendation” includes Manual, Faces and other good photos that contain information or similar photos that have been uploaded to SCROLL before. Such photos have tags under them such as 3d or 4d meaning that they were uploaded to the system three or four days ago.

When a learner clicks one photo in PACALL Browser, the PACALL Recaller will be opened. The photo and similar photos and sensor data will be shown on this page to help the learner recall the captured content. There is also an “Upload” button on this page. If the learner decides to upload this photo to SCROLL as a ULLO, he/she can click this button, and the photo will be uploaded to the SCROLL system directly and the page will jump to the learning log registration page (Figure 8). Figure 8 shows the interface of ULLO registration in the SCROLL system. On this page, a learner can see the location of the selected photo and other similar photos captured by SenseCam. If there are some similar photos that are already uploaded in SCROLL, they will also be shown on this page. Once “Upload Now” is clicked, the system will ask the user to answer a survey that lets the system know whether he/she knows it and whether he/she noticed this object when it was captured. The data can be used to evaluate our system and help the user analyze his/her learning situation. When an object is uploaded to the system, the SCROLL system will use the “organize,” “recall” and “evaluate” model to help users remember uploaded objects. For example, if a learner uploaded a photo and set the title as 消火栓 in Japanese, but does not know how to say it in Chinese, he/she can send a question along with the uploaded ULLO. SCROLL will send this question to all Chinese users. After receiving answers from them, the user can learn a new Chinese word. In the quiz module of SCROLL, a learner can answer quiz questions that are created automatically from uploaded ULLOs. By answering these questions, the learner’s knowledge will be enhanced.

![PACALL Recaller](image)

**Figure 8. PACALL Recaller**

**Evaluation**

We have conducted an evaluation experiment for PACALL. This section introduces the method and result of this experiment.
Method

Since this is an initial evaluation experiment, the study group consisted of 4 Japanese university students taking an undergraduate English course. In this experiment, they were asked to upload photos of learned objects along with titles both in Japanese and English. They used three methods of recording the photos. The entire evaluation experiment lasted for 3 weeks and consisted of 3 phases:

Phase 1: SenseCam

During this phase, students were asked to wear SenseCam every day for one week. Every evening, they needed to review all the life-log pictures and choose proper pictures to upload to SCROLL. They were requested to record the time spent.

Phase 2: Tablet PC

In our previous work (Ogata et al., 2012), we compared the learning effectiveness of Tablet PCs and a traditional learning method such as taking notes. It was found that using SCROLL on a Tablet PC was more effective than the traditional learning method. In this experiment, we compare SenseCam with the Tablet PC in terms of log methods, as logging with a Tablet PC is considered as active, while logging with SenseCam is considered as passive. During this phase, all the students were asked to record and upload the learning log objects every day using a Tablet PC for one week. We used a Samsung Galaxy Tab in this experiment. The operating system of this Tablet PC is Android, and we developed an Android client that can upload the photos to the system conveniently.

Phase 3: SenseCam+PACALL

During this phase, the PACALL system was introduced into the experiment. This phase was almost the same as Phase 1. All the students should wear SenseCam every day for one week. Every evening, they used the PACALL system to classify the life-log pictures and upload the pictures. They were requested to record their time spent. Besides, they were asked to count the number of classifications after all the pictures were uploaded.

Result

Learning chances - Differences in number of ULLOs uploaded per day

We examined the number of uploaded ULLOs among these three learning methods. Figure 9 shows the average number of uploaded ULLOs for each of the four participants.

![Figure 9. The average number of uploaded ULLOs](image)

In this chart, the horizontal axis shows the four subjects, and the vertical axis shows the average number of ULLOs uploaded by each subject. It shows that they uploaded more photos in the SenseCam and SenseCam + PACALL
modes than in the Tablet PC mode. As a result, the passive mode with SenseCam offered them more learning chances than the active mode. Moreover, it is found that PACALL increased the number of uploaded pictures in most cases (except S2). After this experiment, we examined why the number of uploaded pictures did not increase for S2, and then we interviewed him. We learned that at that time he was not so serious about the experiment and just managed to upload 3 pictures a day as a norm. However, the result of Figure 11 shows that it took him nearly half the time to review and upload photos in SenseCam + PACALL mode compared with the time spent in SenseCam mode, so it is expected that if he had been more involved and spent more time, the number of uploaded ULLOs would have increased. In normal circumstances in the subjects’ daily lives, the pictures captured by SenseCam were similar whether using PACALL or not, and the numbers of uploaded objects were almost the same. However, from the feedback, it was found that PACALL reduced the workload of reviewing life-log pictures. The learners could choose pictures and upload them more quickly. So the number of uploaded pictures using SenseCam + PACALL is larger than that using SenseCam only.

Learning quality - Can learners remember uploaded ULLOs that are taken unconsciously in passive mode?

What is the difference in the learning effect of the active and passive modes? This is the second question that we attempted to answer. Therefore, we gave all the students a test after each phase to see whether they had remembered the uploaded objects. We devised a test consisting of the uploaded pictures and asked them to write down the title of the pictures, then judged their memory, grading them from 0 to 5, where 5 means they remembered clearly and 0 means no recall at all. Figure 10 shows the results.

![Figure 10. Memory level of active and passive modes](image)

In Figure 10, we can see that the memory level of active mode is a little higher than that of passive mode. In active mode, the learner takes pictures consciously. Naturally the impression of the photos is deeper than that of those taken in passive mode. Besides, the number of uploaded photos in passive mode was larger than that of active mode which might be reflected in their memory level. Therefore this result is understandable. Even though their memory level was lower, when we consider the fact that they registered more photos as their learning logs, we interpret that our system gave them more chances to learn in passive mode. Therefore we believe that the system contributes to their learning.

Workload issue - How much value does the PACALL add to passive learning mode?

We examined how PACALL contributed to reducing the time spent on the whole procedure. We asked students to report the spent time, and Figure 11 shows the result.

This chart clearly shows that the developed system reduced the time spent reviewing the life-log pictures by nearly a half. Of course, the workload in passive mode is higher than that in active mode, but very few learning contents are missed. In the future, we will focus on how to reduce the workload and help learners to reflect on and find good photos more easily.
Feedback

We received some suggestions and feedback from the students which helped us to understand the usage of PACALL and to improve our system. Some typical feedback is listed as follows.

I think PACALL is easy to use. When I use the SenseCam without PACALL, I must find good photos in the folder from the browser. However when using PACALL, I just select them and click “upload.” The time is shown with the photo in PACALL, which is also helpful for selecting photos. Besides, inappropriate photos are already excluded by this system. It also helps.

It is better to use the Android Tablet PC in conjunction with the PACALL system.

In the passive mode, the learning contents are recorded even when I do not want to learn anything. On the other hand, in the active mode, photos can only be taken when I want to learn something.

I feel very embarrassed when using SenseCam.

The accuracy is not good enough for analyzing blurred photos.

The above comments show that this system is easy to use and the users seem moderately satisfied with the system. In the passive mode the learning contents could be recorded even if learners do not want to learn. In other words, the life-log pictures create more chances to learn vocabulary. However, the learners may feel embarrassed when wearing the SenseCam in public. Moreover, there is a privacy issue which is yet to be resolved. As for the problem of embarrassment, in the future, we believe that the SenseCam will get smaller and look better, and hopefully, learners will not feel so embarrassed wearing it. Besides, the algorithm for classification needs to be improved in the future.

Conclusion

This paper describes a ubiquitous learning log project in Japan called SCROLL. This project was partly supported by PRESTO of the Japan Science and Technology Agency (JST) from 2010 to 2013. It aimed to capture learning experiences in daily life and reuse them for learning and education. Especially, this paper focuses on capturing ubiquitous learning logs using life-log photos, which are automatically taken by SenseCam. We developed a system named PACALL to help learners find learning contents from life-log photos. Also, we took a further step in analyzing life-log photos for the main system. Therefore, PACALL is not only a learning content provider but also a learning content analyzer. Besides, the provided data of PACALL can also be used by the quiz module of SCROLL to determine the proper time for presenting learners with quizzes. We found that there are many useful pieces of information that can be mined from the life-log photos. In the future, we will improve the algorithm of image processing in this system and conduct an evaluation experiment.
Since SCROLL is intended to be used in general domains and for life-long learning, we will apply it to many application domains including foreign language, math, physics, and science education, and conduct a long-term evaluation with a larger sample of subjects. Another area of our future work is learning analytics. We plan to analyze the accumulated data in the learning logs to find learners’ learning patterns and learning habits in order to supply more appropriate learning materials in more appropriate places and at more appropriate times to improve the learning effects of the system.

Acknowledgements

This research work was supported by PRESTO from the Japan Science and Technology Agency, and the Grant-in-Aid for Scientific Research No. 21650225 from the Ministry of Education, Science, Sports, and Culture in Japan.

References


Context-Aware Mobile Role Playing Game for Learning – A Case of Canada and Taiwan

Chris Lu1, Maiga Chang1, Kinshuk1, Echo Huang2 and Ching-Wen Chen2

1School of Computing and Information Systems, Athabasca University, Canada // 2Department of Information Management, National Kaohsiung First University, Taiwan // chrischien630@gmail.com // maiga@ms2.hinet.net // kinshuk@athabascau.ca // echoh@nkfust.edu.tw // chingwen@nkfust.edu.tw

ABSTRACT

The research presented in this paper is part of a 5-year renewable national research program in Canada, namely the NSERC/iCORE/Xerox/Markin research chair program that aims to explore possibilities of adaptive mobile learning and to provide learners with a learning environment which facilitates personalized learning at any time and any place. One of the sub-projects of this 5-year national research program is to design and develop context-aware mobile learning services. The research team of the sub-project applied narrative theory to design a location based Context-Aware Mobile Role Playing Game (CAM-RPG) in order to give students feeling of living in the game world and role playing, exploring the game world, completing the quests, and learning things. A pilot study was then conducted to see how the two game features – context-awareness and story generation – influence students' attitude towards the use of the mobile educational game. The research findings suggest that the story generated in CAM-RPG positively influences users' attitude towards game use and increases users' perceived game usefulness. With the research findings, other components and outcomes of sub-projects, such as natural language processing, location-awareness, multiple input forms, social networking, and student modeling, can then be put together as one piece to provide students effective and efficient mobile learning experiences.

Keywords

Context-awareness, Location-based, Narrative theory, Educational game, Mobile game, Role-playing game, Technology acceptance model

Introduction

The exponential growth of wireless technology in recent years, increasing availability of high bandwidth network infrastructures (e.g., the SuperNet in Alberta), advances in mobile technologies and the popularity of handheld devices have opened up new accessibility opportunities for education. This has given rise to a five-year research program, funded by Canadian federal government and Alberta Provincial government in collaboration with various industry partners. The program aims to explore and develop different applications and content delivery systems, extending our understanding of mobile learning to provide rich learning experiences in order to not only improve the existing educational environment but also to widen access to education for the disadvantaged, particularly those living in remote and rural communities, who generally do not have access to learning opportunities and the disabled, who need specialized devices and applications for learning.

The learning environment that is being developed under this research program consists of different servers and databases, and provides several services for students. The location-awareness service is aimed to help mobile students forming face-to-face-learning groups. Moreover, innovative social networking functions are integrated in the learning environment. An adaptive mechanism is also developed that is responsible for providing learners with learning materials that fit their individual learning styles. The context-awareness service identifies the personalized context-aware knowledge structure in an ubiquitous/pervasive learning environment and is aimed to direct individual learners to learn and move in the real world using automatically generated guidance messages. Furthermore, learners are supported by an intelligent and multimodal asynchronous questions & answers (Q&A) knowledge sharing platform.

The program has three stages. The first stage consists of Canadian research team designing and developing the game and Taiwanese team designing and conducting the pilot study to verify the usability of the game. The second stage involves Canadian team improving the game according to the feedback received in the first stage and conducting a pilot study in Canada for both the iterative development process and cultural difference investigation. The last stage of the program involves application of the well-designed final product in a formal class and a comparative experiment involving both Canadian and Taiwanese students. The program is currently at the end of first stage.
In 2010, the research team of context-aware sub-project developed a Context-Aware Mobile Educational Game (CAMEG) (Lu, Chang, Kinshuk, Huang, & Chen, 2010a, 2010b). The game generates a series of learning activities (i.e., a learning activity chain) to enable students to interact with specific real objects (e.g., projector, rest room, pine tree, etc.) and virtual objects (payroll system, business policy, E-Commerce course, etc.) in authentic environments. The series of learning activities is automatically generated for individual students according to their learning history and surrounding context (i.e., learning objects associated with the chosen role that the student wants to play, the chosen learning theme, student’s location, etc.).

However, majority of the existing educational games, including mobile games, have not looked at such individual feelings. Focus has primarily been on how to teach specific discipline or curriculum in formal educational and on-the-job training settings (i.e., workplace, school campus, museum and historical site). These games become boring when students are simply asked to conduct certain activities one-by-one repeatedly. Few researchers have talked about how to design the contents of mobile educational games in order to make them attractive for the students. This paper focuses on this aspect with aim to improve effectiveness of the mobile educational games.

The rest of the paper is organized as follows. The next two sections introduce the research background by reviewing relevant literature on educational games. The research model and hypotheses used in this research are then described. This is followed by the description of the pilot design and the collected data. Statistical analysis methods are then used to find the answer to the research question. Finally, the implications of the findings are discussed and conclusions are drawn.

Background and motivation

In the last decade, many researchers have seen mobile learning (m-learning) as a further evolution of e-learning (Georgiev, Georgieva, & Smrikarov, 2004). Unlike computer-based learning (learning at a specific place with desktop computers), m-learning delivers education and training materials to a variety of lightweight devices such as personal digital assistants (PDAs), tablet PCs, smartphones, and mobile phones, which users can comfortably carry and use for learning anywhere, at anytime (Keegan, 2005). Beyond the learning devices, some researchers also think that the context of pedagogy differs between e-learning and m-learning. Especially for environmental sciences, m-learning brings potential benefits for learners' self-learning by realizing real-time and location-based learning materials (Jones, Scanlon, & Clough, 2013; Vogel, Spikol, Kurti, & Milrad, 2010).

Brown and colleagues argue that students can learn specific knowledge more efficiently by interacting with a situated environment (Brown, Collins, & Duguid, 1989). Learners can observe or touch the learning objects and can interact with the m-learning system immediately. Hwang, Yang, Tsai, and Yang (2009) also point out that context-aware learning is an innovative approach for detecting student situations and providing students personalized services and adaptive support. Wu and colleagues argue that context-aware ubiquitous learning enables students to interact with learning objects in the real world with the supports from the digital world (Wu, Hwang, & Tsai, 2013). It is important for a mobile learning system to be context-aware; hence, the research team decided to create an interesting context-aware mobile game for students learning domain knowledge.

Garris, Ahlers and Driskell (2002) applied an instructional model in games that uses game-feature-relevant instructional contents as inputs and makes the game-play a cycle. In this model, the repeatable judgment-behavior-feedback activity is a game cycle. These repeated activities can increase the student's motivation and enjoyment of playing the game, enable students to play the game continuously, and increase students' confidence in the gameplay (Garris, Ahlers, & Driskell, 2002).

Researchers have also identified the importance of story in the games (Connors, 2013; Simon, 2012; Sanders, 2011). Connors (2013) argues that story is fundamental for players remembering their gaming experiences and a game might be less impactful without the story. Simon (2012) argues that players may perceive two games to be exactly same if the games have no story, which would have negative effect on learners’ motivation to come back to play the games. Sanders (2011) argues that story can make players aware of the goal of the game and can keep them exploring the game.
In order to make the Context-Aware Mobile Role Playing Games (CAMEG) interesting for the users and to motivate them to play, narrative elements were taken into consideration in this research. Narrative theory covers the elements that a story needs (Conle, 2003); therefore, it was decided to design the story generation engine based on narrative theory. The narrative elements such as storyline, character, and interaction have been analyzed in the literature and used in the game-based learning system design (Ying, Wu, Chang, & Heh, 2009). Researchers have also integrated various narrative elements and designed different approaches for generating story (Akimoto & Ogata, 2011; Akimoto & Ogata, 2012). The research team therefore applied narrative theory to enhance CAMEG in order to give students feeling of living in the game world and role playing, exploring the game world, completing the quests, and learning things. At the end, the enhanced mobile educational game with stories - Context-Aware Mobile Role-Playing Game (CAMRPG) was developed in 2011 (Lu, Chang, Kinshuk, Huang, & Chen, 2011c).

The research team has subsequently been tackling the following research question: do the two game features – context-awareness and story generation – really influence students’ attitudes towards using such educational mobile role-playing games? A pilot study has been conducted, where a questionnaire (and associated statistical analysis) was employed to gather students’ attitudes toward the game.

**Story decorated context-aware mobile role-playing game**

To develop a lightweight, flexible, and scalable mobile educational game based on the research components designed by various sub-projects of the research chair program, a multi-agent architecture (MAA) has been used (Lu, et al., 2010b, 2011a). A multi-agent system is a software environment containing many agents who are responsible for their own tasks while collaborating with other agents whose responsibilities belong to the pre- and post-requisite tasks. Multi-agent architecture is particularly useful for developing mobile applications because it can divide a complex task into several smaller tasks and can assign these tasks to different agents. Moreover, these agents can work either within same device (e.g., a mobile phone) or on different machines/platforms as a distributed system (Balaji and Srinivasan, 2010).

Figure 1 shows the multi-agent architecture of the mobile educational game developed in this research. More details for the responsibilities of each agent and the collaboration among agents can be found in Lu, et al. (2011a)

![Multi-Agent Architecture Diagram](image)

**Figure 1.** Multi-Agent architecture of the proposed mobile educational game

Figure 2 shows the screenshots of the game-play of CAMRPG. During the game-play, the Player Agent is the only agent that interacts with the user and enables data exchange between the user and other agents.
As shown in Figure 2, in the game, roles and corresponding pre-defined themes are designed for students to have opportunity to choose what learning direction and discipline they really need. For instance, a student who takes Introduction to Management Information System course may want to know more about what enterprise support system is and what benefits a business can gain from it. In such circumstances, the student can choose a particular role and theme s/he wants to play, for instance, a chief information officer (i.e., step 2).

The game then generates learning activities for the individual students according to the chosen role and theme. Before the students are asked to do the learning activities, the game makes stories up automatically and uses the stories to populate the generated learning activities for the individual students (i.e., step 3). After the students finish reading the story, the game shows them the learning activities (i.e., step 4). The students can use the built-in camera to collect the required learning object(s) by taking pictures of the objects’ QR codes (i.e., step 5). Once the game has verified the correctness of the collected learning object(s), it delivers each student a piece of text-based learning material about the corresponding learning object (i.e., step 6). In addition to text-based learning contents, the learning contents can also be HTML-based, binary-based image, URL of webpage, media stream and Flash animation.

At the end, the game checks if the students have completed all generated learning activities for the chosen role and theme. If there are other activities left, the game takes the individual students back to step 3 to read another story and asks them to finish another learning activity (as flow B on Figure 2 shows). If no activity is left, the game takes the students back to theme selection screen (as flow A on Figure 2 shows). The students can then either choose another theme or can even take another role.

**Research model and hypotheses**

The research team decided to explore the connection between student's perceived usefulness of the game and the two features (i.e., context-awareness and story generation) step by step, with the following research question "do the two game features – context-awareness and story generation – really influence students' attitudes towards using educational mobile role-playing games?"
A number of models have been proposed in the literature for analyzing user perceptions and acceptance towards technological systems. A well-established and tested model in the literature is the Technology Acceptance Model (TAM), proposed by Fred D. Davis in 1986. This model has become one of the most common instruments used to explain the users' behavioral intention of using an innovative technology. Original TAM has four constructs: the perceived ease of use, the perceived usefulness, the attitude toward using the innovative technology, and the behavioral intention of using the innovative technology.

Some researchers have also examined the acceptance factors for educational games or entertainment games by adding their own variables to the original model to explore the influences of different external variables, for instance, gender, gaming experience, learning opportunities and the unified theory of acceptance and use of technology (UTAUT) (Bourgonjon, Valcke, Soetaert, & Schellens, 2010; Ibrahim, 2011). In the pilot study of this research, two external variables (i.e., the two game features, namely context-awareness and story generation) are proposed for inclusion in the original TAM.

The proposed research model is adopted from the research done by Ibrahim (2011) and Bourgonjon et al. (2010). Different from previous models, this research has four moderators, namely gender, gaming experience, smartphone experience, and context-awareness feature as variables. The reason for taking smartphone experience into consideration is to analyze whether or not the students who do not have experience in using smartphone encounter difficulty in using the game and perceive low ease of use than the students who have experience in using smartphone. Figures 3 and 4 show the macro view (i.e., all considered theories) and micro view (i.e., the detailed constructs) of the proposed research model respectively.

**Figure 3. Macro view of the proposed research model**

**Figure 4. Micro view of the proposed research model**
The hypotheses needed to be verified in the research model are listed in Table 1.

<table>
<thead>
<tr>
<th>Macro view</th>
<th>Micro view</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>H1: Attitude has a positive effect on behavioral intention.</td>
</tr>
<tr>
<td>H2</td>
<td>H2a: Perceived ease-of-use has a positive effect on attitude toward using CAMRPG.</td>
</tr>
<tr>
<td></td>
<td>H2b: Perceived usefulness has a positive effect on attitude toward using CAMRPG.</td>
</tr>
<tr>
<td></td>
<td>H2c: Perceived ease-of-use has a positive effect on perceived usefulness.</td>
</tr>
<tr>
<td>H3</td>
<td>H3a: Context-awareness feature has a positive effect on attitude toward using CAMRPG.</td>
</tr>
<tr>
<td></td>
<td>H3b: Story generation feature has a positive effect on attitude toward using CAMRPG.</td>
</tr>
<tr>
<td>H4</td>
<td>H4a: Context-awareness feature has a positive effect on perceived usefulness.</td>
</tr>
<tr>
<td></td>
<td>H4b: Story generation feature has positive effect on perceived usefulness.</td>
</tr>
</tbody>
</table>

**Pilot design and data collection**

The purpose of the pilot study was to analyze whether a mobile learning system with the two features improves learners’ willingness of using it. Initially, the researchers introduced the game and conducted a demonstration in a Management Information System (MIS) class at a national university in Taiwan. The researchers explicitly told the students that there was no compensation, reward, or recognition for anyone who participated in the study. It was also made clear that there were no consequences for not taking part in the study.

The experiment environment of the pilot study consisted of three laboratories located within one building of the university. Since all participants were taking undergraduate level MIS course at that moment (June, 2011), the MIS course contents and concepts were incorporated into the game and a virtual science park was built in that building. The park consisted of many famous IT businesses and companies that virtually resided in the park, and participants interacted with those organizations in the virtual park while playing the game.

The participants were asked to complete a demographic questionnaire before playing the game. All participants had 20 minutes to play the game in the authentic learning environment using the smartphones prepared by the researchers.

As the participants started to play the game, they received story-enhanced learning activities and looked for the required learning objects in the real world. The learning objects were associated with MIS topics/concepts and were presented in different formats, such as video clips, presentation slides, case studies, and real systems. In the gameplay, participants acted as information technology (IT) experts and received quests from their boss (i.e., a non-player-controlled character). The quests asked them to visit the science park (i.e., the authentic learning environment) and collect some important information for their company. While they were playing, they would learn about these learning objects actively through presentations and demonstrations instead of sitting passively in a classroom and receiving lectures from the course instructor.

After the game-play, they were asked to fill out the technology acceptance model questionnaire. The questionnaire had thirty one five-point Likert-scale items (ranging from 5 for "strongly agree" to 1 for "strongly disagree") to address four main constructs of Technology Acceptance Model (i.e., perceived ease of use, perceived usefulness, attitude toward using, and behavioral intention of using), and two examined constructs (i.e., context-awareness and story generation).

**Reliability analysis**

The questionnaire was adopted from previous research results, and its validity and reliability have been proven by Lu, Chang, Kinshuk, Huang, and Chen (2011b). The data collected in this research was analyzed before using it to examine/verify the hypotheses. Some participants did not show up at the scheduled time, hence, the corresponding
responses of the questionnaire were removed. In addition, responses of two more participants were removed because they had extreme values for all questions and had conflicting answers for the flip-flop items. The final valid sample therefore included 62 students, consisting of 34 male and 28 female students.

Table 2 lists the results of reliability analysis. The Cronbach's alpha for the overall questionnaire is 0.826, indicating that the questionnaire (and its items) can be seen as reliable because its internal consistency is good enough (i.e., exceeds 0.75) (Hair, Anderson, Tatham, & Black, 1998).

The results showed that all constructs, except the behavior construct, had good measure of reliability. The three items of the behavior construct were reviewed and it was concluded that these items might not explain the construct well in this research because of the different subjective situations this research has. The three items had no correlation with the other constructs either. Therefore, the behavior construct was removed.

| Table 2. Reliability analysis results of the technology acceptance model questionnaire |
|-----------------------------------|-----|----------------------------------|
| Construct                         | Item number | Overall Cronbach's alpha |
| Perceived ease of use (PEoU) | 5 | 0.743 (0.774 after PEoU03 and PEoU04’s removal) |
| Perceived usefulness (PU)          | 5 | 0.793 |
| Context-awareness feature (CA)     | 5 | 0.752 (0.807 after CA02’s removal) |
| Story generation feature (SL)     | 4 | 0.832 |
| Attitude toward using CAMRPG (ATT) | 4 | 0.807 |
| Intention of using CAMRPG (IT)     | 5 | 0.894 |

Note. Bold and underline = Cronbach's alpha value is lower than 0.75.

Validity analysis

Next, the internal commonality of items for each factor in the research model was examined using principal component analysis. Three items – PEoU03, PEoU04, and CA02 – were found to have factor loading less than 0.6 and therefore not good enough for presenting the construct. It was decided to remove these three items. The removal of PEoU03 and PEoU04 also improved the Cronbach's alpha value of "Perceived easy of use" construct by bringing it to 0.774, and the removal of CA02 improved the Cronbach's alpha value of "Context-awareness feature" construct to be 0.807. The remaining items could then be used to represent the factors respectively. Lower factor loading may have occurred due to unclear questions or misunderstanding. They need to be revised to fit the presented constructs in future studies and experiments. Table 3 lists results of all constructs in principle component analysis.

| Table 3. Validity analysis results of the technology acceptance model questionnaire |
|-----------------------------------|-----|-----------------|----|----|----|----|
| Factor                            | Item | 1 | 2 | 3 | 4 | 5 | 6 |
| Factor 1: Perceived ease of use (PEoU) α = 0.743 | I2: It is easy to learn how to play | .850 |    |    |    |    |    |
|                                  | I31: The system flow is clear and simple to me | .817 |    |    |    |    |    |
|                                  | I32: The terms and functions in the game are easy to understand | .799 |    |    |    |    |    |
|                                  | I35: User interface are easy to use | .759 |    |    |    |    |    |
|                                  | I36: I can get familiar with the learning objects quickly | .580 |    |    |    |    |    |
| Factor 2: Perceived usefulness (PU) α = 0.793 | I36: It provides me enough information for what I want to know | .762 | .758 | .741 | .733 | .703 |    |
|                                  | I37: I can get needed information quickly within the game | .758 | .741 | .733 | .703 |    |    |
|                                  | I44: The learning activities can save my time in learning | .741 | .733 | .703 |    |    |    |
|                                  | I5: This game makes me want to explore the game’s world | .733 | .703 |    |    |    |    |
|                                  | I52: This game provides me enough information for learning | .703 |    |    |    |    |    |
| Factor 3: Context-awareness feature (CA) α = 0.752 | I5: The learning objects are associated to my chosen theme | .844 |    |    |    |    |    |
|                                  | I54: If a quest required multiple learning objects, all of them can be found in the authentic learning environments | .761 |    |    |    |    |    |

107
Factor 4: Story generation feature (SL) $\alpha = 0.832$

$I_{14}$: I am engaged in the story and relevant quests $\quad 0.847$

$I_{15}$: The stories give me some ideas of what I should do $\quad 0.826$

$I_{16}$: The integration of storyline and quests is perfect $\quad 0.813$

$I_{17}$: The storyline makes the game more interesting $\quad 0.775$

Factor 5: Attitude toward using CAMRPG (ATT) $\alpha = 0.807$

$I_{18}$: I would like to use the game much more if I can team up with other players in the game $\quad 0.875$

$I_{19}$: I would like to try its upgraded version $\quad 0.805$

$I_{20}$: I would like to use the game much more if it has background music $\quad 0.795$

$I_{21}$: I hope the course's instructor to apply "CAM-RPG" into the course $\quad 0.710$

Factor 6: Intention of using CAMRPG (IT) $\alpha = 0.679$

$I_{22}$: I will play "CAM-RPG" continuously in the future $\quad 0.885$

$I_{23}$: I would like to use other similar systems in the future $\quad 0.885$

$I_{24}$: I will try to complete the tasks that the course's instructor asks me to do $\quad 0.863$

$I_{25}$: I will introduce "CAM-RPG" to other people in the future $\quad 0.819$

$I_{26}$: I would play "CAM-RPG" if many of my friends are playing $\quad 0.745$

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>10.048</th>
<th>2.526</th>
<th>2.222</th>
<th>1.586</th>
<th>1.292</th>
<th>1.242</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of variance</td>
<td>35.88</td>
<td>9.02</td>
<td>7.94</td>
<td>5.66</td>
<td>4.61</td>
<td>4.43</td>
</tr>
</tbody>
</table>

Note. Overall $\alpha = 0.826$, total variance explained is 67.54%

At the end, a valid and reliable technology acceptance model questionnaire for measuring participants' attitude towards mobile educational game with six constructs and twenty eight items was determined and confirmed. Quantitative statistical method was then used to get the answers for the research questions.

Data analysis and results

In order to answer the proposed research question, descriptive statistics was initially used to summarize the collected data and compare the constructs' mean and standard deviation values for different groups (e.g., gender smartphone use and player types). Independent t-test was then used to explore whether or not different groups of participants have different attitudes toward CAMRPG.

Descriptive statistics

The demographic questionnaire collected participant's gender information, experiences of playing games, time spent in playing games, and experiences with smartphones. Table 4 lists basic information for the final sample of 62 participants.

<table>
<thead>
<tr>
<th>Gender</th>
<th>N</th>
<th>Smartphone(s) using experience</th>
<th>Playing video games</th>
<th>Playing handheld video games</th>
<th>Playing computer games</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>34</td>
<td>10 (29.4%)</td>
<td>30 (88.2%)</td>
<td>29 (85.2%)</td>
<td>34 (100%)</td>
</tr>
<tr>
<td>Female</td>
<td>28</td>
<td>9 (32.1%)</td>
<td>21 (75%)</td>
<td>22 (78.5%)</td>
<td>27 (96.4%)</td>
</tr>
<tr>
<td>Total</td>
<td>62</td>
<td>19 (30.6%)</td>
<td>51 (82.2%)</td>
<td>51 (82.2%)</td>
<td>61 (98.3%)</td>
</tr>
</tbody>
</table>

Table 4. Demographic information of the participants
Table 4 shows that most participants had rich experiences of playing games, especially computer games. Video and computer games are both found to be major entertainment activities for them. In addition, only 30.6% of participants had experiences of using smartphones.

Quantitative analysis

Independent t-test was used to explore whether there were significant differences in technology acceptance between different groups of participants (e.g., gender, time spent playing computer games, and experiences of using smartphones). The statistical data analysis in Table 5 shows two meaningful results: (1) female participants have more positive feedback than male participants for all constructs; and, (2) there is no obvious difference between male and female participants in their responses for six constructs. The results are in line with the findings of previous researchers (Gwee, Chee, & Tan, 2010; Law, 2010; Papastergiou, 2009).

Table 5. Gender difference on six constructs of technology acceptance model.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Gender</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived ease of use</td>
<td>Female</td>
<td>28</td>
<td>4.3429</td>
<td>.40682</td>
<td>1.579</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>4.0765</td>
<td>.81205</td>
<td></td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>Female</td>
<td>28</td>
<td>4.3000</td>
<td>.37515</td>
<td>1.987</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>3.9882</td>
<td>.75629</td>
<td></td>
</tr>
<tr>
<td>Context-awareness feature</td>
<td>Female</td>
<td>28</td>
<td>4.0000</td>
<td>.29313</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>3.9882</td>
<td>.56126</td>
<td></td>
</tr>
<tr>
<td>Story generation feature</td>
<td>Female</td>
<td>28</td>
<td>4.0982</td>
<td>.51523</td>
<td>1.065</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>3.9412</td>
<td>.62480</td>
<td></td>
</tr>
<tr>
<td>Attitude toward using CAMRPG</td>
<td>Female</td>
<td>28</td>
<td>4.2589</td>
<td>.36945</td>
<td>1.519</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>4.0368</td>
<td>.74907</td>
<td></td>
</tr>
<tr>
<td>Intention of using CAMRPG</td>
<td>Female</td>
<td>28</td>
<td>3.9857</td>
<td>.60106</td>
<td>1.422</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>34</td>
<td>3.7471</td>
<td>.70032</td>
<td></td>
</tr>
</tbody>
</table>

From Table 5, it can be seen that although the mean values of both groups are quite high (positive) for all constructs, male participants have relatively higher standard deviation. This circumstance shows that male participants may have extreme high or low responses for these constructs. It is notable that the statistical analysis for experience of using smartphones shows no obvious difference between smartphone users and traditional mobile phone users.

Multiple regression

To explore the cause-effect relationships in the research model, a simple linear regression (i.e., use of attitude towards using CAMRPG to determine the intention of using CAMRPG) and several multiple linear regressions (e.g., use of perceived ease of use, perceived usefulness, context-awareness feature, and story generation feature to determine attitude towards using CAMRPG; and, use of perceived ease of use, context-awareness feature, and story generation to determine perceived usefulness) have been used. Such multiple regression analysis is typically used to examine and predicate the linear relationship between one dependent construct and one or more independent construct(s).

First, the independent factors were analyzed before entering the regression model in order to know whether there is a collinear problem in the statistics. A collinear problem is a statistical situation in which two or more predictors (independent constructs) in a multiple regressions are highly correlated. This situation causes an abnormally high R-square (i.e., explanatory power) in the regression model because the variances, standard error, and parameter estimates of predictors are probably inflated. It may also cause insignificant or incorrect coefficients (e.g., positive to negative) between predictors and affected variables.

The existence of linear dependence in the independent constructs can be determined by observing the collinearity statistic fields in Table 6. A collinearity statistic indicates that the construct may have a serious overlap (i.e., a collinearity problem, which means there is high correlation between the independent constructs) if the variance
inflation factor (VIF) is over 10 and tolerance tends to zero (Hair, Anderson, Tatham, & Black, 1998). The results show that there is no serious collinearity issue between the independent constructs in Table 6, in which tolerance > 0.1, VIF < 10, and no two variables' variances > 0.8 at the same line.

Table 6. Coefficients of perceived ease of use, perceived usefulness, context-awareness and story generation feature

<table>
<thead>
<tr>
<th></th>
<th>Unstandardized coefficients</th>
<th>Standardized coefficient</th>
<th>t</th>
<th>Significance</th>
<th>Collinearity statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>ß</td>
<td></td>
<td>Tolerance</td>
</tr>
<tr>
<td>(constant)</td>
<td>.441</td>
<td>.423</td>
<td>1.044</td>
<td>.301</td>
<td>.260</td>
</tr>
<tr>
<td>Perceived ease of use</td>
<td>.296</td>
<td>.137</td>
<td>.323</td>
<td>2.162</td>
<td>.035*</td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>.338</td>
<td>.146</td>
<td>.346</td>
<td>2.311</td>
<td>.024*</td>
</tr>
<tr>
<td>Context-awareness feature</td>
<td>.015</td>
<td>.156</td>
<td>.012</td>
<td>.999</td>
<td>.921</td>
</tr>
<tr>
<td>Story generation feature</td>
<td>.248</td>
<td>.111</td>
<td>.234</td>
<td>2.235</td>
<td>.029*</td>
</tr>
</tbody>
</table>

Note. Dependent variable: Attitude toward using CAMRPG *: p < 0.05

Table 7. Collinearity diagnostics

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Eigenvalue</th>
<th>Condition index</th>
<th>Variance proportions</th>
<th>(constant)</th>
<th>PEOU</th>
<th>PU</th>
<th>CA</th>
<th>SL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.968</td>
<td>1.000</td>
<td></td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>2</td>
<td>.015</td>
<td>18.274</td>
<td>.41</td>
<td>.12</td>
<td>.07</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>3</td>
<td>.009</td>
<td>23.118</td>
<td>.21</td>
<td>.06</td>
<td>.00</td>
<td>.00</td>
<td>.84</td>
<td>.84</td>
</tr>
<tr>
<td>4</td>
<td>.005</td>
<td>32.640</td>
<td>.30</td>
<td>.03</td>
<td>.32</td>
<td>.71</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.003</td>
<td>38.685</td>
<td>.09</td>
<td>.79</td>
<td>.61</td>
<td>.28</td>
<td>.13</td>
<td>.13</td>
</tr>
</tbody>
</table>

Note. Dependent variable: Attitude toward using CAMRPG. PEOU = Perceived ease of use; PU = Perceived usefulness; CA = Context-awareness feature; SL = Story generation feature.

Table 8 lists the coefficients of four independent constructs towards the dependent construct – attitude towards using CAMRPG. Three constructs (i.e., perceived ease of use, perceived usefulness, and story generation feature) present significant coefficient measures (ß = 0.323, 0.346, and 0.234, p < 0.05).

Path analysis for the multiple regressions model

Figure 5 shows the path diagram of the research model. The result of path analysis shows that the attitude towards using CAMRPG (ATT) has strong effects on the intention of using CAMRPG as 0.455 (p < 0.001) of path coefficient. The effects of perceived ease of use (PEoU), perceived usefulness (PU), context-awareness feature (CA) and story generation feature (SL) explain 74% of the attitude towards using CAMRPG, while perceived ease of use (PEoU), perceived usefulness, and story generation feature have significant effects (ß = 0.331, 0.437, and 0.234, p < 0.05) on the attitude towards using CAMRPG, but context-awareness feature does not (ß = 0.012). For the cause-and-effect relationship between the independent variables, the effects of perceived ease of use, context-awareness feature, and story generation feature explain 75% of perceived usefulness, while perceived ease of use (ß = 0.678, p < 0.001) and story generation feature (ß = 0.206, p < 0.05) have significant effects on perceived usefulness, but context-awareness feature does not (ß = 0.066).
Findings and discussions

Data analysis revealed several findings that can help in understanding users' attitudes towards and acceptance of the proposed mobile educational game as well as exploring the answer of the research question: do the two game features—context-awareness and story generation—really influence students' attitudes towards using educational mobile role-playing games?

These findings are categorized into three categories: common findings (i.e., those that have been proven in other research), important findings (i.e., those that are supported by this research), and unexpected findings (i.e., those that did not support our assumptions in this research).

Common findings

Findings suggested that the original technology acceptance model presents good results in cause-and-effect relationship of all factors (e.g., PEoU, PU, ATT, and IT). In particular, for the path coefficients found between perceived ease of use and perceived usefulness, the results indicate that ease of use is an important factor in context-aware mobile educational game design as well as other technology acceptance issues. Users appreciate a simple and easy-to-use interface, and a user friendly interface directly impacts perceived usefulness. In addition, attitude towards using CAMRPG and intention of using CAMRPG also present strong significant coefficients in our research model. These findings have been proven in many studies that have focused on the acceptance towards information systems.

Important findings

First of all, the descriptive statistical data (i.e., Table 6) shows that the responses from both males and females were positive in terms of appreciation of the proposed CAMRPG. In addition, responses of female participants to all factors were relatively higher than those of male participants in the pilot.

The result did not show any significant differences between participants who have experience using smartphones and those who only have experience using traditional mobile phones. The reason may be that the participants in this pilot were undergraduate students and they were all familiar with mobile phones and games. Therefore, experience of
using smartphones did not affect acceptance of innovative technology. On the other hand, from the perspective of national research program, this result suggested that there is no need to worry about whether or not a user has used a smartphone while deploying such context-aware mobile role-playing game for learning.

**Unexpected findings**

From the path analysis results, it was found that most of the proposed factors qualified to explain the dependent variables, except the context-awareness feature factor. One reason perhaps is that the context-awareness feature is transparent to its users. For instance, a participant will receive from the game only those learning activities that involve learning objects in a library if the game detects that the participant is in library at that moment. So the participant would not feel what exactly the feature does for him or her. Another reason could be that the experiment environment in the pilot might not have represented the concept of context-awareness well enough to make participants aware of this game feature. For instance, the pilot was conducted using a virtual science park that was built in a university building used by the participants regularly for attending classes, which made it difficult for participants to have immersive feelings that they were in San Francisco or Helsinki. Finally, this pilot did not cover different buildings and did not continue over a longer time period. In such case, the participants could not experience scenarios like signing on to get quests at different places. These shortcomings might have caused the context-awareness feature factor to present relatively lower measures and an insufficient cause-and-effect relationship on the path coefficient.

**Conclusions**

This paper presented the outcome of the first stage of the context-aware sub-project, under the auspices of the 5-year national research chair program, namely the context-aware mobile role playing game, in which its kernel – learning activity generation engine and story generation engine – can automatically generate a series of story-based learning activities. This game can help users in learning by role-playing in authentic learning environments. The story makes up the learning activity chain resulting in more interesting and immersive learning process. Integrating story into a mobile educational game increases the perceived effectiveness and satisfaction toward the game, especially for the male students. On the other hand, the story reduces the perceived efficiency of using the system.

The findings indicate that participants in the pilot found the context-awareness feature of the game to be less important for the game-play and this factor did not affect their attitude towards using the game. The findings also identified the importance of authentic environment in mobile learning. The pilot study designed in this research clearly demonstrated that the context-awareness ability of the system was not even noticed by students, since they were asked to imagine a floor of a teaching building in school campus as a country in the world, hence there was a lack of an authentic environment. Such mismatch between the virtual and the real world has potential to reduce the perceived usefulness of the context-awareness functionality.

To make users aware of the advantages of a context-aware mobile educational game, subject selection (e.g., learning environment, selected learning topic, and learning materials) would be an important issue. The current game seems to work well for outdoor teaching/learning as well as learning based on treasure hunting paradigm at particular sites (e.g., museums, botanical gardens, and historical sites). It is also suitable for replacing orientation/training courses for freshmen and new students of the graduate programs. However, such game might not be suitable in environments in which learning objects have no strong connections to either the learning topic or the environment (e.g., trying to learn a business intelligent system from a desktop computer in a laboratory).

The pilot study encompassed only a short-term intervention whose effect may not be carried for long run. The research results also provided a clear picture of what learning topics may be more appropriate for applying context-aware mobile role-playing games, what authentic environment and learning objects are the best for deploying context-aware mobile learning systems, and what features are important to students who use mobile role-playing games.

The next stage research will focus on continuing the architecture design and proof of concept of the services developed under the national research program. This includes the incremental improvement of various modules
based on proof of concept evaluations in Canada as well as continuing to integrate the developments within the overall system. As the research results show that the context-awareness feature is difficult to be noticed when the learning environment is a mix of mismatched virtual and real worlds, augmented reality may help in enhancing the perceived usefulness of context-awareness feature of the game. Also, in order to increase the effectiveness of the game, ordinary learning activities, such as field trips and remedial learning can be integrated into the game. Future plans of the context-aware sub-project include: (1) study and application of augmented reality concept within the interactive mobile learning systems in order to provide students the benefits of context-awareness; (2) develop Android version of CAMRPG and deploy it in rural areas for K-12 education and field trips; and, (3) integrate the outdoor remedial instructions and worksheet idea together to provide students an even more personalized ubiquitous learning experience according to their academic performances and the context surrounding them.

References


Potentials of Mobile Technology for K-12 Education: An Investigation of iPod touch Use for English Language Learners in the United States

Min Liu*, Cesar C. Navarrete and Jennifer Wivagg

Department of Curriculum & Instruction, The University of Texas at Austin, 1912 Speedway Stop D5700, Austin TX 78712, US // mliu@austin.utexas.edu // ccnavarrete@utexas.edu // jwivagg@keystoneschool.org

*Corresponding author

ABSTRACT
This case study investigated a m-learning initiative by a large school district in the United States to provide iPod touch devices 24/7 to teachers and students of English Language Learners. We described the initiative and presented the research findings of its implementation for two years at elementary and middle school levels. The results revealed the iPod touch was used to support language and content learning, provide differentiated instructional support, and extend learning time from classroom to home. However, several challenges were identified such as significant time demand on the teachers, technical issues, the need for professional training and dedicated support staff. Implications for teachers, instructional technologists, school administrators, and researchers were discussed.

Keywords
iPod touch, Mobile device, M-learning, English language learners, K-12 education

Introduction
Emerging mobile technology is having a transformative impact on how people live, learn, work, and play. According to 2012 Horizon Report for K-12 Edition (Johnson, Adams, & Cummins, 2012), “mobile devices & apps and tablet computing as technologies expected to enter mainstream use in the first horizon of one year or less.” Incorporating mobile devices in K-12 education has experienced increased interests and schools in the United States (US) are looking for new opportunities and possibilities introduced by the mobile technology. We describe an initiative by a large school district in the southwest region of US to provide iPod touches 24/7 to teachers and students of English Language Learners (ELL). We present research findings of its implementation for two years from 2010 to 2012 at elementary and middle school levels. ELL students are those who speak diverse languages such as Spanish, French, Portuguese, Chinese, and Japanese. With different levels of English proficiency (some have been in US schools for several years, while others are new to English language instruction), they represent distinct academic challenges in language acquisition. Especially in the region, Spanish-speaking students are a rapidly growing population. The goal of investigating iPod touch use over two implementation cycles is to gain an in-depth understanding of the potential benefits of such devices as a teaching and learning tool for the ELL population. The findings of this research have implications for teachers, instructional technologists, school administrators, and researchers as they explore and consider mobile technologies for K-12 education.

Relevant research
Mobile learning has garnered considerable interests from the educational community (Koole, 2009; Traxler, 2011) and literature on mobile learning has identified such affordances as flexibility, accessibility, interactivity, and motivation and engagement. Implementing flash card learning, for example, Kiger, Herro, and Prunty (2012) found third-graders improved their multiplication skills using iPod devices. Studies have shown Internet-enabled mobile devices can support cognitive learning (Peng & Chou, 2007); mathematics learning (Kalloo & Mohan, 2011); language and literacy learning (e.g., Coe & Oakhill, 2011; Kemp & Bushnell, 2011); and game-based learning (e.g., Liao, Chen, Cheng, Chen, & Cha, 2011).

For ELL students, ready access to information technology is considered a critical factor (Cummins, 2000). Mobile Internet-enabled devices can provide ELL students with access to learning resources for “comprehensible input” in language acquisition necessary for academic success (Cummins, 2000, p. 541). Using the mobile devices for audio with trade books provided by the school teacher-librarian in collaboration with teachers, Patten and Craig (2007) found iPod shuffle devices to support reading and writing with ELL students. By adding audio support to their text
reading and journal writing, the use of the iPod devices were also found to significantly increase student engagement and allow the students greater connection to the “U. S. popular culture” (Craig & Paraiso, 2007, p. 1840).

Despite the perceived affordances of m-learning, there is a paucity of research on using mobile devices with ELL students, especially how ELL teachers and students use such devices in classrooms for teaching and learning. This study investigates ELL teachers’ and students’ use of iPod touches in elementary and middle school classrooms as well as their perceptions of using it as a teaching and learning tool. Our research question is: In what ways can iPod touch devices serve as a teaching and learning tool for English Language Learners?

Research project

Research background

The mobile initiative, for which this research was conducted, took place in a large school district in the southwest region of US. The district has a 2011 enrollment of about 18,000 students in grades K-12 and covering an area of about 600 square miles. Of that population, about 4.5% are students with limited English proficiency. Spanish is the primary language spoken at home for about 90% of ELL students in the district, while the remaining ELL students speak a variety of other languages at home. Because of different levels of English proficiency, teaching ELL students is a unique academic challenge. In an attempt to address a substantial gap between the test scores of ELL students versus other students in the district, in September 2009 the school district purchased the latest mobile technology with Internet at the time - iPod touch - for every ELL student and teacher at the middle school level. By allowing ELL students, 24/7 utilization of the devices for accessing additional educational resources, the district is hoping to boost the ELL students’ English language proficiency. In this district, ELL students in grades K-5 are placed in bilingual instruction classrooms that consist of ELL students only. In sixth grade, ELL students transition away from bilingual classroom into mainstream regular education classroom where ELL students are mixed with non-ELL students with support from English as Second Language (ESL) teachers. In 2011-2012, this mobile initiative expanded from middle school level (grade 6-8) to 4-5 grades at elementary level for ELL teachers and students only. In this paper, we report the research conducted in two cycles: Year One implementation from 2010-2011 and Year Two implementation from 2011-2012.

Participants

The participants in this research were two ELL middle school teachers and their students for the first cycle (during 2010-2011 school) and two ELL elementary school teachers and their students for the second cycle (during 2011-2012 school year). Two sixth-grade teachers were selected because they were among the first to implement the initiative. Two teachers, fourth and fifth grade, were selected because they were part of the initiative during the second cycle and were teaching only ELL students. Teachers’ demographics are provided Table 1. In addition to iPod touch, all these teachers had access to other technologies such as an interactive whiteboard, a desktop computer, a Mac laptop, and a document camera in their classrooms.

<table>
<thead>
<tr>
<th>Data collection period</th>
<th>Level</th>
<th>Teacher</th>
<th>Years of teaching</th>
<th>Grade levels taught</th>
<th>Total students taught/managed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle one: 2010-2011</td>
<td>Middle school</td>
<td>Virginia</td>
<td>26</td>
<td>6-8</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Claire</td>
<td>7</td>
<td>6-8</td>
<td>31</td>
</tr>
<tr>
<td>Cycle two: 2011-2012</td>
<td>Elementary school</td>
<td>Clara</td>
<td>14</td>
<td>4-5</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lydia</td>
<td>23</td>
<td>4-5</td>
<td>42</td>
</tr>
</tbody>
</table>
**Data sources and analysis**

The primary data sources were (1) interviews with the teachers, (2) classroom observations, and (3) surveys with the students.

Interviews were conducted with each teacher at three different times: the beginning of their iPod touch implementation, the mid-school year and toward the end of the school year. The goal of these interviews was to understand the teachers’ and students’ experience and perceptions toward iPod touch use. Questions varied because of the different focus for each interview during the year. Sample questions are:

At Beginning-of-Year:
- How you are using the iPods?
- How are the students using the iPods?

During Mid-year:
- How are you using the iPod touch?
- Please describe how does the iPod touch help student learning? How does it constrain student learning? Provide examples.
- What, if any, challenges are there in developing iPod touch activities? Examples?
- How is iPod touch used outside of the school day?

At End-of-Year:
- How did you use the iPod touch in your teaching? What has worked and what has not?
- In your opinion, does the iPod touch make a difference in the students’ learning? Provide examples?
- In what way does iPod touch support your students’ learning? How does that compare to student learning without the iPods?
- Do you see any challenges in using the iPod? Explain.

All interviews were conducted face-to-face or through conferencing software, audio-recorded, and transcribed. Researchers also observed each teacher’s use of iPod touch in their classrooms. Field-notes were taken and details of the activities were written afterwards. The guidelines of Strauss and Corbin (1990) were followed in the data analysis: First, each interview was independently coded by one researcher to generate a list of initial codes, then another researcher reviewed and verified the codes. Necessary coding modification, realignment, and refinement were made during the process. This data analysis went through several iterative cycles until codes were categorized and themes emerged. Any disagreement was discussed until 100% inter-rater reliability was reached. Interview quotes were presented below as unedited. Observation notes were analyzed in a similar way and were used to supplement the interview data.

Three surveys were given to the students: at the beginning, middle, and end of the school year. The goal of the surveys was to seek students’ self-reported usage (as there was no tracking software available for iPods) and find out their perception toward having iPod touch as a learning tool. Sample survey questions are “How often do you use the following applications/features at school?” “How often do you use the following applications/features at home?” “How helpful or fun do you think the following applications or features are?” and demographic questions. Survey responses were analyzed descriptively and frequency data were tallied. Since not all students responded to all survey items, the results provided the percentages of those who responded to the questions. Throughout the data collection process, the research team met regularly to discuss what they observed in different classrooms, shared insights, and performed peer debriefing.

**Findings**

**Year-one implementation**

Virginia taught three classes per day to ELL students and two of these were dedicated to beginner ELL students while Claire taught four periods of ELL classes with three of these dedicated to beginners. In their second year of
iPod use in their classrooms, Virginia and Claire enthusiastically embraced the opportunity the iPod touch initiative presented and were the first group of teachers who implemented it in their classrooms.

Interviews with Virginia and Claire and classroom observations revealed both teachers had very positive perceptions toward using iPod touch and found ways to incorporate the device in their daily teaching activities as well as home assignments. Data analysis highlighted a few affordances of mobile technology in their case: Using iPod touch to (1) support language and content learning with Internet-based multimedia resources, (2) provide differentiated instructional support to accommodate students’ individual needs and create collaborative learning opportunities, and (3) extend learning time from classroom to home use and establish a better home-school connection.

Support language and content learning

The student iPod touches were preloaded and synced by the teachers with native apps along with an Internet browser as well as a variety of learning games, videos, audio books, multimedia, and textbooks. Virginia and Claire had students access a wide range of multimedia resources to support English language learning in classes as well as assigned homework.

During the classroom instruction, the students were able to access resources such as translation dictionaries that provided audio pronunciations along with images to support the vocabulary acquisition and audio textbooks. Claire described an English learning activity, “They go find a picture and the name of the picture, we use [the app] ComicLife for that, so they have a picture and vocabulary word and at the beginning they add a Spanish word to it.” Students used the devices for just-in-time translations as well as multimedia materials to support content learning in math, science, and social studies. For example, Virginia described, “I used a lot of the science videos and I try to keep it on the topic that the science teachers are teaching.” The iPod touch, as a small and portable computer with Internet access, was available to students 24/7 and provided them with a plethora of multimodal resources for English language and content learning.

Provide differentiated instructional support

In addressing various English language learning levels of these students, the teachers used varied apps and resources to allow students for access to appropriate levels of language and academic instruction. As an example, Claire spoke of how the iPod “allows more development, more customized learning” and said that with the iPod, “they have their private tutor” with non-judgmental support in their activities. Virginia also spoke of how the iPod use helped the students, “It’s private, it’s between them and, so they don’t feel stupid.”

The teachers assigned the activities and games appropriate for students’ specific language levels and were able to scaffold students more easily with iPod touch from basics in phonetics and sight words to more fluency and comprehension in advanced topic and subjects. Virginia described her differentiated instruction: “I give instructions: 7th grade, you go to this video, 8th grade you go to this video.” Collaborative learning was also afforded in allowing the students to share information gathered individually with other students in classroom discussion. Claire illustrated, “They feel free to say whatever they think and they don’t feel that pressure [to participate].” Besides supporting the students in shared learning activities, Virginia had her students access current events and collaborated in identifying key information and answered questions on the topic.

Extend learning time from classroom to home

The students were assigned homework activities using iPod touch for extended learning at home. Virginia elaborated on her literacy learning focus with iPods for “helping them read better and being oral speakers” with audio recordings. The ability to take the iPods home allowed the student further academic learning away from the classroom. Emily, a seventh grade student, summarized her iPod use for learning, “I practice more in the iPod.” With the multifarious learning games and apps, the teachers believed their students were more engaged in further practices than traditional homework materials and activities offered.
Claire considered the importance of allowing student access to the iPod away from school and said, “The iPod brings more a family into the issues of learning.” Some students’ parents were able to access the students’ grades at school as well as being able to monitor their children’s actual work on the iPod. Additionally, the teachers were able to assign the student homework that included the siblings and parents. For example, Claire had the students conduct a video interview of their parents talking of their educational experiences as children. Virginia included homework that had the students read to their siblings, “… take one of my baby books and read it to your little bothers and sisters.” Parents were able to access the Internet via iPod touch for their own use as well. Virginia identified an incident of finding “job application” and discovered that the student’s parent was seeking employment information via the device.

### Middle school students’ usage and perception

The 6th, 7th and 8th graders were asked to report how often they used the available features on their iPod touch on a weekly average both at school and at home. The features available to the students can be grouped into three categories: Resource tools (e.g., calculator, calendar, accessing Internet, maps, listening to music and podcasts, and checking weather), media creation tools (voice recorder, notes, still and video cameras) and other applications (often in the form of games). Table 2 showed students’ responses for school and home use. Comparing mid-of-school usage to end-of-school usage, the data showed that a considerable percentage of students used resource and media creation tools for two or more hours per week during the mid-school year; however, at the end-school year, almost all of the students reported using these applications 0-1 hour per week. The results indicated that the highest use of the iPod touch occurred during mid-school year when teachers and students were engaged in using the resource and media creation tools as part of weekly learning activities. The noticeable decrease of usage pattern toward the end of the school year could be explained in part by the fact that much of the school time during the end-school year was spent on preparing for various state mandated tests and end-of-year curriculum exams and testing itself. Comparing school and home usages, the data indicated approximately 63% of the students spent 2 or more hours using the Internet at home (with 25.9% spent more than five hours at home while 16.7% spent more than five hours at school). This is also true for listening to music with 40.7% spent 2-5 hours at home while 11.5% at school and 25.9% spent more than 5 hours at home while 15.4% at school. That is, students spent more time using the Internet (both school or non-school related) and listening to music (non-school related) at home than at school.

We asked the students to name three features they used most often and found most helpful Accessing Internet, listening to music, and using the calculator were the three features the students used most often both during mid-school year and at the end-school year for resource tools. Voice recorder, video camera, and notes were used most often for media creation tools. Students found Internet and calculator most helpful for the resources tools and voice recorder and notes for media creation tools.

### Table 2. How often students used iPod applications/features on average each week

<table>
<thead>
<tr>
<th>Feature</th>
<th>School</th>
<th>Home</th>
<th>School</th>
<th>Home</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>At school</strong></td>
<td><strong>At home</strong></td>
<td><strong>Mid-year %</strong></td>
<td><strong>End-year %</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Duration in hours</strong></td>
<td></td>
<td><strong>Mid-year %</strong></td>
<td><strong>End-year %</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Mid-year %</strong></td>
<td><strong>End-year %</strong></td>
<td><strong>Mid-year %</strong></td>
<td><strong>End-year %</strong></td>
</tr>
<tr>
<td>Calculator</td>
<td><strong>0-1</strong></td>
<td>48.1 (n = 27)</td>
<td>94.2 (n = 17)</td>
<td>66.7 (n = 27)</td>
</tr>
<tr>
<td></td>
<td><strong>2-5</strong></td>
<td>40.7</td>
<td>5.9</td>
<td>33.3 (n = 17)</td>
</tr>
<tr>
<td></td>
<td><strong>More than 5</strong></td>
<td>11.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Calendar</td>
<td><strong>0-1</strong></td>
<td>77.8 (n = 27)</td>
<td>100 (n = 17)</td>
<td>85.2 (n = 27)</td>
</tr>
<tr>
<td></td>
<td><strong>2-5</strong></td>
<td>11.1</td>
<td>0</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td><strong>More than 5</strong></td>
<td>11.1</td>
<td>0</td>
<td>7.4</td>
</tr>
<tr>
<td>Internet</td>
<td><strong>0-1</strong></td>
<td>45.8 (n = 24)</td>
<td>94.1 (n = 17)</td>
<td>37.0 (n = 27)</td>
</tr>
<tr>
<td></td>
<td><strong>2-5</strong></td>
<td>37.5</td>
<td>5.9</td>
<td>37.0 (n = 27)</td>
</tr>
<tr>
<td></td>
<td><strong>More than 5</strong></td>
<td>16.7</td>
<td>0</td>
<td>25.9 (n = 17)</td>
</tr>
<tr>
<td>Maps</td>
<td><strong>0-1</strong></td>
<td>66.7 (n = 24)</td>
<td>100 (n = 16)</td>
<td>64.0 (n = 27)</td>
</tr>
<tr>
<td></td>
<td><strong>Mid-year %</strong></td>
<td><strong>End-year %</strong></td>
<td><strong>Mid-year %</strong></td>
<td><strong>End-year %</strong></td>
</tr>
</tbody>
</table>

119
<table>
<thead>
<tr>
<th></th>
<th>0-1</th>
<th>2-5</th>
<th>More than 5</th>
<th>0-1</th>
<th>2-5</th>
<th>More than 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>73.1 (n = 26)</td>
<td>11.5 (n = 27)</td>
<td>15.4 (n = 17)</td>
<td>33.3 (n = 27)</td>
<td>40.7 (n = 17)</td>
<td>25.9 (n = 0)</td>
</tr>
<tr>
<td>2-5</td>
<td>29.2 (n = 25)</td>
<td>23.5 (n = 16)</td>
<td>87.5 (n = 17)</td>
<td>76.4 (n = 17)</td>
<td>25.0 (n = 11.8)</td>
<td>20.8 (n = 0)</td>
</tr>
<tr>
<td>More than 5</td>
<td>4.1 (n = 27)</td>
<td>0 (n = 17)</td>
<td>0 (n = 0)</td>
<td>20.0 (n = 0)</td>
<td>0 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>Weather</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>56.0 (n = 25)</td>
<td>36.0 (n = 24)</td>
<td>8.0 (n = 17)</td>
<td>54.2 (n = 24)</td>
<td>25.0 (n = 17)</td>
<td>20.8 (n = 0)</td>
</tr>
<tr>
<td>2-5</td>
<td>64.0 (n = 25)</td>
<td>12.5 (n = 24)</td>
<td>0 (n = 0)</td>
<td>76.4 (n = 17)</td>
<td>11.8 (n = 0)</td>
<td>11.8 (n = 0)</td>
</tr>
<tr>
<td>More than 5</td>
<td>15.4 (n = 17)</td>
<td>0 (n = 0)</td>
<td>0 (n = 0)</td>
<td>25.9 (n = 0)</td>
<td>0 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>Podcasts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>58.3 (n = 24)</td>
<td>29.2 (n = 24)</td>
<td>12.5 (n = 27)</td>
<td>51.9 (n = 27)</td>
<td>25.9 (n = 17)</td>
<td>22.2 (n = 0)</td>
</tr>
<tr>
<td>2-5</td>
<td>64.0 (n = 25)</td>
<td>5.9 (n = 17)</td>
<td>0 (n = 0)</td>
<td>94.1 (n = 17)</td>
<td>5.9 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>26.9 (n = 27)</td>
<td>0 (n = 17)</td>
<td>0 (n = 0)</td>
<td>50.0 (n = 26)</td>
<td>0 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>Media creation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice recorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>57.7 (n = 26)</td>
<td>26.9 (n = 26)</td>
<td>15.4 (n = 27)</td>
<td>50.0 (n = 26)</td>
<td>38.5 (n = 17)</td>
<td>11.5 (n = 0)</td>
</tr>
<tr>
<td>2-5</td>
<td>64.0 (n = 14)</td>
<td>5.9 (n = 17)</td>
<td>0 (n = 0)</td>
<td>94.1 (n = 17)</td>
<td>5.9 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>19.0 (n = 13)</td>
<td>0 (n = 0)</td>
<td>0 (n = 0)</td>
<td>25.9 (n = 0)</td>
<td>0 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>76.2 (n = 21)</td>
<td>19.0 (n = 14)</td>
<td>4.8 (n = 13)</td>
<td>71.4 (n = 21)</td>
<td>19.1 (n = 13)</td>
<td>9.5 (n = 0)</td>
</tr>
<tr>
<td>2-5</td>
<td>70.7 (n = 20)</td>
<td>0 (n = 15)</td>
<td>0 (n = 0)</td>
<td>100 (n = 20)</td>
<td>0 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>15.0 (n = 22)</td>
<td>0 (n = 13)</td>
<td>0 (n = 0)</td>
<td>100 (n = 22)</td>
<td>0 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>Still camera</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video camera</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>50.0 (n = 20)</td>
<td>35.0 (n = 20)</td>
<td>15.0 (n = 17)</td>
<td>50.0 (n = 20)</td>
<td>40.0 (n = 13)</td>
<td>10.0 (n = 0)</td>
</tr>
<tr>
<td>2-5</td>
<td>63.1 (n = 19)</td>
<td>18.8 (n = 16)</td>
<td>0 (n = 0)</td>
<td>76.4 (n = 17)</td>
<td>23.6 (n = 0)</td>
<td></td>
</tr>
<tr>
<td>More than 5</td>
<td>15.8 (n = 22)</td>
<td>0 (n = 17)</td>
<td>0 (n = 0)</td>
<td>22.7 (n = 17)</td>
<td>0 (n = 0)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Not all students responded to each question. N indicated the number of the students who completed the question. This is also the case for Table 3. Other applications mostly included games the teachers downloaded for the students.

Challenges revealed

Several challenges were observed during this first cycle: Significant amount of time required of the teacher to learn to use iPod touch, finding appropriate apps, developing lessons that would integrate the mobile devices, managing the devices by charging, synching, downloading, and updating, and dealing with technical problems. These tasks presented challenges for the teachers who already had a demanding workload with teaching ELL students. In addition, the teachers had to attend necessary technology training required by the school district.

Another challenge was encountering technology issues such as loss of Wi-Fi capacity for use of multiple devices at once at school as well as teachers helping individual students with login problems. The devices were also potentially prone to breakage or loss as the students took them home and other settings away from the classroom. For example, during the school year, several devices were broken or lost. A few thefts occurred throughout the year and one theft of 15 iPod touches towards the end-school year with only one recovered. Other examples of damage included accidental mishandling and washing with clothes.
Similar data collection procedure was followed for both cycles. However, as researchers, we realized asking middle school students to self-report their weekly iPod usage can be a challenge as this age group often had a hard time remembering an activity within a week’s duration and had trouble understanding hours vs. minutes. Therefore, in surveying elementary students for the second cycle, we asked the students to report daily usage and made the response time duration from hours to minutes.

**Year-two implementation**

Clara and Lydia were partner teachers for 4th and 5th grade ELLs and both held a bilingual teaching certification. Clara taught math and science and Lydia taught English Language Arts and Social Studies. As a team, they collaborated on group activities, student discipline, and iPod integration. Clara also served as the manager of the devices, updating and syncing the devices for all 42 students. Compared to Virginia and Claire, Clara and Lydia, in their initial year of iPod integration, were less enthusiastic about this mobile initiative and they implemented the project because they were encouraged by the school district to participate in the initiative. The analysis highlighted a few affordances of mobile technology in their case: Using iPod touches to (1) support language and content learning with Internet-based multimedia resources and (2) provide differentiated instructional support to increase student engagement and collaboration in the classroom.

**Support language and content learning**

Clara and Lydia shared insights on their experience, explaining their efforts in incorporating iPod touches into their teaching. Similar to the middle school teachers, their experiences showed they used the tool for engaging students in language as well as content learning. They described how the iPod touch provided ELL students access to multiple resources that facilitated their learning. Clara detailed the process for multimedia use for vocabulary development: “They go online and find a picture that is associated with the word and download it or save it to their iPod and transfer it to their StoryKit. Then, they write the definition that goes with it. Then, they go into recording themselves reading their definition.” As another example, Lydia spoke of how she used the iPod for language and reading skills development, “We are using them in terms of recording their fluency or recording to listen to their reading.” Clara also stated, “All their science homework is uploaded into their iTouches” for easy sharing and grading, not possible with traditional instruction.

**Provide differentiated instructional support**

Clara described the differentiated learning opportunities made possible by being able to access the resources at students’ own levels as “none of my kids have ever been on the same level.” The students were engaged in using such resources as dictionaries and playing educational games at a level appropriate to their learning and therefore reinforcing skills such as reading and multiplication. Clara described her typical use in this way, “There is no more downtime in my classroom because any down time we had, even when I’m passing out papers, their iTouches are always on their desk and they know that: ‘Okay, get it out and start working on your times tables’ or ‘get it out and work on divisibility.’” Moreover, the students enjoyed using the iPod touch for learning. Lydia spoke of a writing activity in which the students went on a field trip and created a story using StoryKit with images, “They loved it. They had enjoyed the field trip anyway and were able to go in to create their own story pretty much just with peer assistance.” Additionally, Internet access allowed students to pursue topics of their interest as Lydia described a student’s use, “He will look up his favorite baseball players, what their stats are, and he will have a new line of interests depending on what he what he’s reading in his novels sometimes.” The students found the writing activities compelling when using images from the field experience. Clara added, “It’s fun for them. It’s not paper and pencil.”

**Elementary students’ usage and perception**

The 4th and 5th graders were asked to report how often they used the available features on their iPod touch on a daily average both at school and at home. The features available to the students were also grouped into three categories as
with the middle school students: Resource tools, media creation tools, and other applications. Table 3 shows elementary students’ responses for school and home use.

The overall usage patterns indicated an increase in use with the resource tools at school but a decrease in home use; an increase in use for the media creation applications at school, except for the camera, while the usage remained the same or decreased at home. In addition, home use of the video camera appeared to increase at the end-school year. For calculator, calendar, music and reminders, more than 50% of the students indicated not using them both at school and at home. These elementary students typically spent 30 minutes or less in using applications/features daily for approximately 2 to 2.5 hours each week. While relatively few apps/features were used for over an hour daily, of particular interest was that audio book use exceeded an hour.

For resources tools, when comparing the mid- and end-of-year usage, it showed a decrease in calculators use both at school and at home. Conversely, the use of calendar, Internet, and maps increased at school and at home. Internet use at school increased by 32% at the 31-60 minute duration while Internet use at home remained the same at the mid-year and end-school. For the media creation tools, the data indicated the voice recorder, notes, and video camera showed a modest increase in usage at the 1-30 minute time frame at school, while at home the voice recorder use dropped and video camera use increased. These usage patterns corroborated with teachers’ interviews indicating the types of school activities the students were required to do. For the other applications category, the trend indicated that both school and home use increased at the 31-60 minute duration at school and home.

When the students were asked to indicate how helpful or fun the applications/features are, more than half of the students found all of the applications/features helpful with the exception of music. Reporting the percentages of the features that exceed 50% of student responses, all resource tools, except for music, were found helpful and all media creation tools were found helpful. The features with the highest percentage of the students finding it helpful was audio books of the resource tools (helpful with 92.3% at mid-year and 92.1% at end-year), and the voice recorder of the media creation tools (82.0% at mid-year and 81.6% at end-year). The high usage of audio books and voice recorder was consistent with the teachers’ description of assigned learning tasks with the iPod touch.

The responses of “how fun” features that exceed 50% of students’ responses showed: calculator was identified as Not fun (55% at end-year), Internet was indicated as Very fun (55% at mid-year and 63.1% at year-end) and other application category (predominantly as games) was indicated Very Fun (68.7% at mid-year). Responses of the other features were spread out and offered mixed results.

<table>
<thead>
<tr>
<th>Resource tools</th>
<th>Duration in minutes</th>
<th>Mid-year %</th>
<th>End-year %</th>
<th>Mid-year %</th>
<th>End-year %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculator</td>
<td>0</td>
<td>66.7</td>
<td>78.9</td>
<td>51.3</td>
<td>76.3</td>
</tr>
<tr>
<td></td>
<td>(n = 39)</td>
<td></td>
<td></td>
<td>(n = 39)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>21.1</td>
<td>0</td>
<td>23.7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(n = 38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31-60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(n = 39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>61-90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Calendar</td>
<td>0</td>
<td>71.2</td>
<td>52.6</td>
<td>69.2</td>
<td>57.9</td>
</tr>
<tr>
<td></td>
<td>(n = 39)</td>
<td></td>
<td></td>
<td>(n = 39)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>47.4</td>
<td>30.7</td>
<td>42.1</td>
<td>30.7</td>
</tr>
<tr>
<td></td>
<td>(n = 38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31-60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(n = 39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>61-90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Internet</td>
<td>0</td>
<td>22.5</td>
<td>10.5</td>
<td>48.7</td>
<td>44.7</td>
</tr>
<tr>
<td></td>
<td>(n = 40)</td>
<td></td>
<td></td>
<td>(n = 39)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>67.5</td>
<td>44.7</td>
<td>33.3</td>
<td>36.8</td>
</tr>
<tr>
<td></td>
<td>(n = 38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31-60</td>
<td>7.5</td>
<td>39.5</td>
<td>15.4</td>
<td>18.4</td>
</tr>
<tr>
<td></td>
<td>(n = 38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>61-90</td>
<td>2.5</td>
<td>5.3</td>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td>Maps</td>
<td>0</td>
<td>52.6</td>
<td>36.8</td>
<td>55.3</td>
<td>26.3</td>
</tr>
<tr>
<td></td>
<td>(n = 38)</td>
<td></td>
<td></td>
<td>(n = 38)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-30</td>
<td>44.7</td>
<td>60.5</td>
<td>36.8</td>
<td>65.8</td>
</tr>
<tr>
<td></td>
<td>(n = 38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31-60</td>
<td>2.6</td>
<td>23.7</td>
<td>5.2</td>
<td>7.9</td>
</tr>
<tr>
<td>Time</td>
<td>Music</td>
<td>Weather</td>
<td>Reminders</td>
<td>Clock</td>
<td>Audio book</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>1-30</td>
<td>5.1</td>
<td>41.0</td>
<td>13.2</td>
<td>50</td>
<td>15.0</td>
</tr>
<tr>
<td>31-60</td>
<td>2.6</td>
<td>2.6</td>
<td>2.6</td>
<td>2.6</td>
<td>2.5</td>
</tr>
<tr>
<td>61-90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>92.3</td>
<td>56.4</td>
<td>13.2</td>
<td>50</td>
<td>47.3</td>
</tr>
<tr>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 40)</td>
</tr>
<tr>
<td>0</td>
<td>92.1</td>
<td>47.3</td>
<td>2.6</td>
<td>2.6</td>
<td>5.3</td>
</tr>
<tr>
<td>1-30</td>
<td>17.9</td>
<td>60.0</td>
<td>28.9</td>
<td>7.7</td>
<td>2.5</td>
</tr>
<tr>
<td>31-60</td>
<td>2.5</td>
<td>5.1</td>
<td>2.6</td>
<td>2.6</td>
<td>5.3</td>
</tr>
<tr>
<td>61-90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>79.5</td>
<td>35.9</td>
<td>28.9</td>
<td>7.7</td>
<td>5.4</td>
</tr>
<tr>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 37)</td>
</tr>
<tr>
<td>1-30</td>
<td>17.9</td>
<td>60.0</td>
<td>28.9</td>
<td>7.7</td>
<td>5.4</td>
</tr>
<tr>
<td>31-60</td>
<td>2.5</td>
<td>5.1</td>
<td>2.6</td>
<td>2.6</td>
<td>5.3</td>
</tr>
<tr>
<td>61-90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>13.5</td>
<td>41.0</td>
<td>5.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 39)</td>
<td>(n = 37)</td>
</tr>
<tr>
<td>1-30</td>
<td>64.8</td>
<td>54.0</td>
<td>40.5</td>
<td>44.7</td>
<td>37.5</td>
</tr>
<tr>
<td>31-60</td>
<td>27.0</td>
<td>50.0</td>
<td>45.4</td>
<td>54.7</td>
<td>34.2</td>
</tr>
<tr>
<td>61-90</td>
<td>10.5</td>
<td>50.0</td>
<td>27.0</td>
<td>55.3</td>
<td>34.2</td>
</tr>
<tr>
<td>0</td>
<td>10.5</td>
<td>50.0</td>
<td>27.0</td>
<td>55.3</td>
<td>34.2</td>
</tr>
</tbody>
</table>
Challenges revealed

During the second year of implementation, the two elementary teachers communicated similar challenges as those by the two middle school teachers. For these two technology novices, they especially indicated significant amount of additional time required of them to carry out the project. In addition, these teachers pointed out they had to attend additional technology training during the holidays as well as in the classroom. Although the teachers indicated the benefits of the training, they also pointed out this was in addition to their regular workload. Apart from the technical issues, appropriate use of the devices was identified as an issue in the elementary setting. The teachers described individual students who misused the camera function as well as cases of attempts to bypass teacher-managed settings. Monitoring appropriate student use while supporting mobile device use in K-12 setting poses a specific challenge. Moreover, the teachers indicated a need for change in teaching approaches. While the purpose of this mobile initiative was to improve ELL students’ success on state assessments, the various affordances provided by mobile devices were found to encourage and support teachers to adopt more independent, learner-centered pedagogical approaches.

Discussion and implications

Discussion

The goal of this research is to investigate how iPod touch devices can serve as a teaching and learning tool for English Language Learners. Encompassing a two-year period with two implementation cycles, the findings revealed that iPod touch use can offer ELL students critical support for language and content learning (Banister, 2010; Craig et al., 2007; Kiger et al., 2012; Patten & Craig, 2007), as well as engaging students in enjoyable learning (Ullman, 2010). Three affordances of mobile technology are highlighted with the results: Using iPod touch to support language and content learning, provide differentiated instructional support, and extend learning time from classroom to home (Chan et al., 2006; Looi, Seow et al., 2010; Looi, Zhang et al., 2010). For these ELL students, several features are especially important: audio books, Internet access, and media creation tools. The iPod touch provides them with another learning platform in a just-in-time manner to facilitate their language acquisition and content learning. Teachers also have easier access to resources and tools to adapt and adjust their instruction to meet students’ needs (Lacina, 2008; Ullman, 2010). The noticeable changes in the teacher’s curriculum practices is to allow students to have more control over their learning. Because of the instant access to useful resources (e.g., translator, dictionary, voice recorder, and other applications) at anytime and anywhere, the students can use tools they previously did not have access to learn content materials. This is especially important for ELL students as they can practice language skills in different ways using multiple modes at school and at home. The teachers in this study noted a need to shift their pedagogy toward a more student-centered approach with affordances offered by mobile devices. Kukulsk-hulme and Traxler (2005) described the “discursive/didactic dichotomy” within m-learning as a complexity in teaching approaches in traditional environments (p. 27). Additionally, consistent with previous research suggesting that engaging games can provide enjoyable content learning (Huizenga et al., 2009), the results showed that the students often accessed the game apps both at school and home and considered such apps as fun. Educators should take advantage of the new opportunities and possibilities provided by mobile devices to “engage, motivate, support, and interest students” (Liao et al., 2011, p. 86) to positively influence learning.

Although there are clear benefits for using iPod touch in teaching and learning as this research indicated, the results also showed teachers will need significant added support for learning how to use and manage the devices. This is especially so for teachers who themselves are not enthusiastic technology users or technology novices. Using mobile technology requires teachers to allocate significant time and effort to learn to use the device as well as learn to integrate effectively in their teaching (Kukulsk-Hulme & Traxler, 2005; Traxler, 2011). The management of the devices for classroom wide use, in terms of charging, synching and dealing with multiple technical demands can also add a substantial load on the teachers (Franklin & Peng, 2008; Vogel, Kennedy, & Kwok, 2009).

Implications

Given the preliminary encouraging evidences of m-learning for ELL students and teachers, what we have learned from this research project for the past two years not only has implications for this district, but also should offer insights for teaching and learning practices for other schools that are considering similar initiatives.
For teachers

Teachers who plan to incorporate mobile devices should be prepared to invest considerable amount of time in learning the functionalities of the devices as well as in designing, planning, and developing effective learning activities. Preparing to use the device for instruction will require prior preparation during the summer, before the school year commences. Moreover, teachers should be flexible in their teaching practice, reconsidering curriculum as well as teaching and learning approaches in order to effectively leverage the full range of affordances that mobile devices can offer. ELL teachers should take advantage of the new possibilities to offer differentiated instruction in meeting the specific language and academic needs of ELL students. To accommodate students’ learning needs and encourage student-directed learning, teachers can consider management plans to monitor student learning and create activities that utilize a wide range of apps/features (e.g., audio recordings, video productions, digital stories) available on mobile devices.

For instructional technologies

School district’s instructional technology department played a key role in this implementation. Prior to and during the implementation, the instructional technologists should provide training for the teachers, focusing on the basic capabilities and operations as well as identifying appropriate apps and suggesting ideas on how to integrate the iPod touch into the overall curriculum. Additionally, instructional technologists must anticipate, identify and quickly resolve technical challenges as well as help teachers in making sure that devices are synced with the appropriate apps, media, software, and security settings. They need to make sure students and teachers know how to log in to their devices and access wireless Internet, while coordinating with the network infrastructure staff in ensuring for sufficient and reliable network resources availability. Technology staff also needs to remain available for troubleshooting technical problems so that these types of issues do not lead to frustration and disuse.

For school administrators and policy makers

The results of this research suggest that realizing the benefits of the iPod touch program requires initial and ongoing support from the school administration. Support should include dedicated instructional technology staff to provide training and on-going support as well as provide funding to purchase and maintain the iPod touches. It is suggested that policies and processes for fixing or replacing broken, lost, and stolen devices should be in place. Administrators should also consider how to get buy-in from teachers when rolling out a technology initiative such as this one and set and enforce guidelines for safe and appropriate use of the devices and their apps. Equipping entire classes with mobile devices can create a significant demand for wireless bandwidth and administrators need to make sure the infrastructure can meet the demand. Without the support to ensure a fully functional device most of the time, student and teacher frustration may lead to declining use and thus undermine or negate the affordances of these devices and their potential to enhance learning.

For researchers

Conducting classroom-based, mobile learning research is a challenge, especially in ELL classrooms. Given the nature of ELL instruction in US public schools, there are many classroom transitions in middle schools on any given day. ELL students can be in different content classes and capturing classroom iPod interactions consistently poses a challenge. Because there was no tracking software on iPods, usage has to rely on students’ self-reporting. Students’ responses often are short and less specific. Follow-up interviews could be considered to further examine students’ perception.

Conclusion

The experience of the two-year implementation of this mobile initiative and research findings suggest that mobile devices such as iPod touch can provide ELL students significant support for language and content learning and extend learning time from classroom to home. Audio books, Internet access, and media creation tools are found to be
especially important for these ELL students. While several affordances are revealed both at elementary and middle school ELL settings, substantial support in the form of professional training and administration support and encouragement is also highlighted.

References


The Impact of a Principle-based Pedagogical Design on Inquiry-based Learning in a Seamless Learning Environment in Hong Kong

Siu Cheung Kong and Yanjie Song
Department of Mathematics and Information Technology, The Hong Kong Institute of Education, Hong Kong // sckong@ied.edu.hk // ysong@ied.edu.hk

ABSTRACT
An inquiry-based learning pedagogy coupled with a seamless learning environment is a potential way to realise the educational goal of learner-centred learning in digital classrooms in the 21st century. An overarching research framework is proposed for preparing teachers to effectively develop pedagogical designs that are premised on theoretical principles and facilitate inquiry-based learning in a seamless learning environment. We carried out an initial study using the overarching framework. Three questions are addressed: how a principle-based pedagogical design was developed and implemented, the effect that the principle-based pedagogical design had on students’ domain knowledge gains and inquiry skills and how students advanced their domain knowledge and developed their inquiry skills. One teacher and 27 students from a local primary school were involved in the study. Both qualitative and quantitative data were collected and analysed over two weeks. Six inquiry-based learning lessons focusing on a scientific ‘rustproofing’ learning unit were conducted in a seamless learning environment, initiated in a digital classroom and extended to online discussions on a social network platform. The results reveal innovative ways of developing and implementing the pedagogical design in the rustproofing learning unit and demonstrate the pedagogical design’s positive effect on students’ domain knowledge gains and inquiry skills. In addition, how the students advanced their domain knowledge and inquiry skills were also explored and discussed.

Keywords
Inquiry-based learning, Primary school education, Principle-based pedagogical design, Seamless learning environment

Introduction
Educational reform calls for a paradigm shift to learner-centred domain knowledge learning. It is well recognised that the inquiry-based learning approach is a useful pedagogy for realising learner-centred learning (Marshall, Smart, & Horton, 2010). The inquiry-based learning process helps learners to develop inquiry skills, which are an important type of 21st century skill. The development of inquiry skills takes root during a child’s senior primary school years (Lakkala, Lallimo, & Hakkarainen, 2005). Digital classrooms are on the rise; students are connected and learn in a ‘one learner to one computer’ setting, and teachers are expected to be prepared to lead students to learner-centred learning in such classrooms as early as the primary school stage. The use of online learning platforms inside and outside digital classrooms supports resource access and peer interaction to develop students’ domain knowledge and inquiry skills (Kong & So, 2008; Lakkala et al., 2005). Incorporating the inquiry-based learning pedagogy into a seamless learning environment may thus be a potential method for realising learner-centred educational goals and driving teachers to apply and reflect on pedagogical designs. This study presents a design-based research framework for principle-based pedagogical design for inquiry-based learning in seamless learning environments. It details and reports the results of an initial study conducted using this framework in a Hong Kong primary school.

Research framework
With the goal of finding a meaningful and sustainable method of developing the teacher competence necessary to facilitate inquiry-based learning in a seamless learning environment, this research study seeks to address two issues. First, the method for developing teacher competence should be in line with the method for developing learner competence in inquiry-based learning. Second, the method for developing teacher competence should present evidence of the development of teacher competence and student learning improvement. This study adopts a principle-based approach to developing and implementing pedagogical designs for inquiry-based learning. This approach, which differs from the conventional approach that emphasises “best practices” with prescribed procedures, provides more flexible scaffolding under guiding principles. It attempts to build up teachers’ capacity to promote inquiry-based learning in a manner aligned with learners’ inquiry-based learning practice. This study also adopts a
design-based research approach to developing and refining pedagogical designs for inquiry-based learning guided by instructional principles. It exposes teachers to the process of progressive refinement in pedagogical designs driven by principles and supported by empirical evidence. In light of these issues, we propose an overarching research framework on the use of principle-based pedagogical designs for inquiry-based learning in a seamless learning environment, as shown in Figure 1.

Figure 1. Research framework for principle-based pedagogical designs for inquiry-based learning in a seamless learning environment

A design-based research method is adopted to prepare teachers to effectively develop pedagogical designs that are premised on instructional principles. Learners learn in a seamless learning environment to develop the necessary skills to practice inquiry-based learning.

Inquiry-based learning includes three approaches: structured, guided and open inquiry, listed in ascending order of the learner’s autonomy over setting investigation problems and planning problem-solving procedures (Colburn, 2000). The literature suggests that the guided inquiry approach is especially suitable for young learners, as teachers only select core issues that are worthy of a learner’s inquiry (Hakkarainen, 2003; Marshall et al., 2010; Song & Looi, 2012). According to Wong and Looi (2011), seamless learning environments provide learners with opportunities to make use of diverse resources and tools in digital formats for learning and communication, which is initiated in digital classrooms and extended to online interactions. The technological support of a seamless learning environment allows learners to conveniently share and store multimedia resources, and to easily exchange and track discussion ideas with peers during the inquiry process. During class time, learners in digital classrooms are connected, and use digital technologies in a ‘one learner to one computer’ setting (Chan, 2010; Kong, 2011). Beyond the limited class time, learners typically use learning platforms to communicate with peers online, mostly to extend discussions or to engage in deeper discussions after class.

Pedagogical design refers to the organisation plan for learning activities and the actual implementation of the plan in a learning unit (Lakkala et al., 2005). Researchers have reported that principle-based pedagogical designs are more adaptable and conducive to transforming inquiry-based learning practices (Schwarz, 2009; Song & Looi, 2012; Zhang, Hong, Morley, Scardamalia, & Teo, 2011). The principle-based approach to pedagogical design defines the core principles of learning and teaching. According to Schwarz (2009), Zhang (2010) and Zhang et al. (2011), principle-based pedagogical designs focus on guiding principles and customizable practices. Teachers are afforded the flexibility to reflectively judge and adapt classroom decisions to accommodate different learning and teaching possibilities. Based on knowledge-building and social-constructivism theories (e.g., Scardamalia, 2002; Vygotsky, 1978), a set of theoretical principles premised on 12 knowledge-building principles and progressive inquiry principles (Lakkala, Muukkonen, Paavola, & Hakkarainen, 2008; Scardamalia, 2002; Song & Looi, 2012; Zhang et al., 2011) is considered suitable for pedagogical designs for inquiry-based learning.
Teachers need support from evidence-based research to make continuous pedagogical reflections. As such, the design-based research approach is suitable for gaining new insights. Design-based research attempts to combine theory-driven design with empirical analyses of practices in real settings. It creates a path to connect interventions to outcomes through an iterative mechanism of design, evaluation and refinement (Bell, 2004; Hoadley, 2004). Teachers are provided with iterative opportunities (as shown in Figure 1) to use principle-based pedagogical designs to enhance their competence in leading inquiry-based learning in a seamless learning environment. This study details and reports the results of an initial study on design-based research that explored the effect of using the principle-based approach to pedagogical designs for science inquiry in the seamless learning environment of a Hong Kong primary school.

The initial study was conducted using the overarching framework in a learning unit on ‘rustproofing’ conducted in the school’s Primary 4 class.

This study

Research context

The study took place in the initial cycle of design-based research on principle-based pedagogical designs for inquiry-based learning that aimed to develop teacher competence and learners’ science domain knowledge and inquiry skills in a seamless learning environment at the primary level in Hong Kong. According to recent territory-wide surveys on the development of technology-enhanced education in Hong Kong (Li & Kong, 2011), local primary school teachers are typically capable of integrating technology into their daily teaching methods. Further, local primary school learners are ready to use technology for inquiry-based learning, as they demonstrate a basic information literacy competency that is important in the inquiry process. These surveys reveal that primary schools in Hong Kong have built a foundation for introducing educational innovations that integrate pedagogical designs for inquiry-based learning into technology-supported learning environments.

The study purposefully sampled a primary school in Hong Kong as its partner school. One experienced science teacher and one Primary 4 class with 27 students (15 female and 12 male) were invited from the partner school to participate. The science inquiry focused on a learning unit on rustproofing conducted in six lessons over 2 weeks for senior primary school learners. The following research questions were addressed.

1. How did the teacher develop and enact the principle-based pedagogical design?
2. What effect did the principle-based pedagogical design have in helping students to gain domain knowledge and inquiry skills in the seamless learning environment?
3. How did the students advance their domain knowledge and develop their inquiry skills?

To address these questions, our pedagogical design involved the adoption of the 5E inquiry-based learning model to guide the students’ science inquiry, and five instructional principles for pedagogical practice in a seamless learning environment supported by a social network (i.e., Edmodo). These principles are elaborated in the remainder of this section.

5E inquiry-based learning model

According to EDB (2008), the focus of science education is to promote students’ scientific thinking through inquiry-based learning approaches. Although open inquiry provides optimal opportunities for students’ cognitive development and scientific reasoning, teacher-guided inquiry may provide better opportunities for students to focus on the development of particular science concepts (Song & Looi, 2012). To balance the two inquiry approaches, we developed a 5E inquiry-based pedagogical model as follows: (a) “engage” in inquiry topics and questions, (b) ‘explore’ the inquiry methods and processes, (c) “explain” the inquiry analyses and outcomes, (d) “evaluate” the inquiry processes and outcomes and (e) “extend” the inquiry topics and questions. The process is cyclic and progressive but not linear, and may not involve all of the components in each learning cycle.
Five instructional principles

To explicate the processes and dynamics of science inquiry for knowledge advancement using the inquiry-based pedagogical approach, we adapted five core instructional principles from a set of progressive inquiry principles (Song & Looi, 2012) and other related research (Scardamalia, 2002; Yeo & Tan, 2010). These principles are premised on social constructivist principles, and include (a) working on real problems, (b) encouraging diverse ideas, (c) providing collaborative opportunities, (d) using authoritative sources constructively and (e) performing a formative assessment (see Table 1).

<table>
<thead>
<tr>
<th>Principles</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working on real problems</td>
<td>Setting up real-life problems rather than abstract concepts (Scardamalia, 2002; Song &amp; Looi, 2012).</td>
</tr>
<tr>
<td>Encouraging diverse ideas</td>
<td>Encouraging students to express their ideas voluntarily. There is no right or wrong answer. Every idea is valued and unique (Song &amp; Looi, 2012).</td>
</tr>
<tr>
<td>Providing collaborative opportunities</td>
<td>Emphasis on the importance of collective effort and responsibility in the learning process (Scardamalia, 2002; Song &amp; Looi, 2012).</td>
</tr>
<tr>
<td>Using authoritative sources constructively</td>
<td>The meaningful use of authoritative sources for continual science meaning-making (Yeo &amp; Tan, 2010).</td>
</tr>
<tr>
<td>Performing a formative assessment</td>
<td>Provision of peer assessment and teacher feedback concurrently in the collaborative process (Scardamalia, 2002; Song &amp; Looi, 2012).</td>
</tr>
</tbody>
</table>

Teacher principle-based understanding

Before the beginning of the study, the teacher received two 1.5-hour training sessions from two researchers. In the first session, the researchers prompted and discussed the inquiry-based learning model and five instructional principles with the teacher using PowerPoint slides. The researchers then asked the teacher to reflect and pose questions. In the second session, the teacher chose one of his lessons to elaborate how he would conduct it using instructional principles, and discussed his pedagogical design with the researchers.

Seamless learning environment supported by a social network—Edmodo

This study took place in a seamless learning environment. The teacher provided pedagogical support in implementing the 5E model and five instructional principles for inquiry-based learning. The seamless learning environment comprised digital technologies that allowed students to access learning resources and interact with peers in inquiry-based learning. Inside the digital classroom, each student was given a mobile computing device comprising a tablet PC with Internet connectivity and an embedded camera. The social network Edmodo (see Figure 2), was used as a learning communication platform to support the students’ learning at the individual, group and whole-class levels both inside and outside the classroom. Edmodo is a secure microblogging medium conducive to collaborative knowledge construction (Ma, Ko, Chu, & Song, 2012). It can be used across formal and informal learning settings, allowing students to collaborate, communicate, submit assignments and upload and download files, and teachers to share lecture notes with students, connect to useful websites, upload and download learning references for students, create online quizzes and release news and events. The platform can run on different operating systems (e.g., iOS or Android).

Figure 2. Interface of Edmodo for the Primary 4 science class
The learning activities organisation plan

The principle-based pedagogical design included an organisation plan of the learning activities (see Table 2) and the actual enactment of the learning unit plan on rustproofing, which comprised six lessons and activities that Primary 4 students carried out at home or between lessons over two weeks. According to the plan, the students formed six groups of four or five members each and were expected to collaboratively lead their own experimental inquiries into an “expert rustproofing design” project. Two prompts were provided on Edmodo during the experimentation process. First, the ‘Forms for Experimental Rustproofing Designs’ prompt asked students to record their inquiry plans. The students were required to fill out three design methods with hypotheses (Appendix I), and each student was required to take on a responsibility in the design. Second, the ‘Observational Forms for Rustproofing Experiments’ prompt helped students to monitor their experimental process and scaffold their reflections. The students were required to document the rustproofing process over a week and to include the observers’ names (Appendix II). Both of the Edmodo forms were linked to GoogleDocs, which allowed the students to fill them in directly.

<table>
<thead>
<tr>
<th>No.</th>
<th>Aim</th>
<th>Activity</th>
<th>Means of interaction</th>
<th>Teaching and learning resources</th>
</tr>
</thead>
</table>
| Lesson 1 | To engage students on the topic of rustproofing | - Storytelling was used to make students understand why they needed to rustproof and to arouse their curiosity on how to do so.  
- Individual students were required to discover rustproofing methods and prepare to share their discoveries with group members in the next lesson. | F2F + online | LCD projector, Tablet PCs, Internet, Social network: Edmodo |
| At home | To discover rustproofing methods individually | - Individual students continued researching rustproofing methods and uploaded their findings to Edmodo to share with their peers.  
- The students posted questions to Edmodo and commented on or responded to other students’ posts. | Online learning | Desktop, laptop, iPad, iPhone, etc. Edmodo |
| Lessons 2 and 3 | To determine the three best experimental rustproofing designs in groups | - Each group member shared and explained his or her findings on the experimental rustproofing designs.  
- The three best experimental rustproofing designs were discussed and worked out by combining every member’s ideas.  
- The three best experimental designs were filled in on the ‘Forms for Experimental Rustproofing Designs’, prepared by the teacher on Edmodo to share with peers. | F2F + online | LCD projector, Tablet PC Edmodo |
| At home | To plan and prepare the material for the rustproofing experiment in the next session | - Students could post their questions to Edmodo. They could also comment on or respond to other students’ posts. | Online | Desktop, laptop, iPad, iPhone, etc. Edmodo |
| Lessons 4 and 5 | To conduct the experiment based on the proposed experimental designs in groups | - Each group conducted three rustproofing experiments by placing three iron clips into three plastic cups full of water and certain other materials.  
- The cups containing the clips and materials were placed on windowsills. | F2F + online | Three iron clips provided by the teacher for the students to conduct the experiments. Students brought the rustproofing materials from home, including the plastic cups. |
**Breaks between lessons**

- Group members observed and took pictures each day to document the changes in the clips in the different cups and filled out the ‘Observational Forms for Rustproofing Experiments’ on Edmodo.

**Lesson 6**

- The groups presented and explained their experimental results to the other groups. Each group showed three experimental designs for protecting the iron clips from rust.
- Each group commented on the other groups’ work and chose the group that achieved the best rustproofing results.
- A badge was awarded to the best group.

**At home**

- Students were required to complete a worksheet on the topic of rustproofing on Edmodo.
- Students posted questions to Edmodo.

### Data collection and analysis

To understand the effects of the principle-based pedagogical design on students’ domain knowledge of rustproofing and inquiry-based learning skills, and on how the students advanced their domain knowledge and inquiry strategies, we collected the following data.

- Data on the development and implementation of the pedagogical design, including posts on Edmodo, the learning activities organisation plan, lesson videos, group experimental design forms, group experimental observational forms, the assignment and teacher interviews and reflections.
- Data on the effects of principle-based pedagogical design on domain knowledge gains, including pre- and post-domain tests and assignments. The pre- and post-domain tests were identical and consisted of 20 multiple-choice questions and five open-ended questions on rustproofing. The assignments took the form of worksheets on rustproofing knowledge (Appendix III). The students submitted their assignment directly to Edmodo, and the teacher commented on their work directly to provide immediate feedback. The data on the effects of the principle-based pedagogical design on student inquiry skills included pre- and post-questionnaire surveys on perception changes towards inquiry learning skills before and after the inquiry-based learning approach. The questionnaire focused on students’ perceptions of inquiry learning skills, and comprised 12 items rated on a five-point Likert scale (with 5 indicating strong agreement and 1 indicating strong disagreement). The 12 items were designed to address the five inquiry skills under the 5E inquiry-based learning model, with items 1-3 addressing the “questioning” skill; items 4-5 addressing the “exploring” skill; items 8-11 addressing the “explaining” skill; items 6-7 addressing the “evaluating” skill and item 12 addressing the “extending” skill. The Cronbach’s alpha reliability scores were 0.849 and 0.902 for the pre-test and post-test, respectively, implying that the questionnaire was reliable.
- Data on how the students advanced their domain knowledge, including group experimental designs (experimental design forms on Edmodo), group experimental results/products (observational forms on Edmodo) and group artefacts (photos documenting the experimental process), and data on how the students advanced their inquiry skills, including posts on Edmodo, lesson videos, teacher and student interviews and field notes.

Table 3 shows the data sources used to investigate the three research questions.

<table>
<thead>
<tr>
<th>Data</th>
<th>Research questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre- and post-domain tests</td>
<td>*Q1</td>
</tr>
<tr>
<td>Pre- and post-questionnaire surveys</td>
<td>*Q2(a) *Q2(b) *Q3(a) *Q3(b)</td>
</tr>
</tbody>
</table>

Table 3. Data sources for analysis
Results

Development and implementation of the principle-based pedagogical design

The data analysis results on the development of the principle-based design show that the learning activities organisation plan included the five elements of the inquiry-based learning model in a seamless environment. The inquiry learning activities went from engaging students on the topic of rustproofing, to exploring rustproofing methods and making hypotheses for the experimental designs, to evaluating the experimental rustproofing designs through active experimentation and explaining and sharing the designs and finally to consolidating and extending the rustproofing knowledge to help the students to become rustproofing ‘experts’. The entire inquiry process was carried out seamlessly between classes and the students’ homes with the support of the social network platform (Edmodo).

The data analysis results on the implementation of the principle-based pedagogical design indicate that the teacher premised the rustproofing learning activities organisation plan on the five instructional principles (see Table 1). The demonstration of the five inquiry-based elements and five instructional principles in a seamless learning environment during the enactment of the rustproofing learning unit is illustrated in Table 4.
Table 4. Principle-based pedagogical implementation of inquiry-based learning in a seamless learning environment

<table>
<thead>
<tr>
<th>Implementation of the organisation plan</th>
<th>Inquiry elements</th>
<th>Instructional principles</th>
<th>Seamless learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesson 1</td>
<td>Engage: Engaging in the topic of rustproofing</td>
<td>P1</td>
<td>Class</td>
</tr>
<tr>
<td>At home</td>
<td>Explore: Exploring rustproofing methods</td>
<td>P4</td>
<td>Home</td>
</tr>
<tr>
<td>Lessons 2 and 3</td>
<td>Explain: Sharing and planning rustproofing experiments and making hypotheses in groups</td>
<td>P2 and P3</td>
<td>Class</td>
</tr>
<tr>
<td>At home</td>
<td>Engage: Preparing and coordinating the rustproofing experiments</td>
<td>P1 and P3</td>
<td>Home</td>
</tr>
<tr>
<td>Lessons 4 and 5</td>
<td>Evaluate: Evaluating the hypotheses through conducting experiments</td>
<td>P1, P3 and P5</td>
<td>Class</td>
</tr>
<tr>
<td>Breaks between lessons</td>
<td>Evaluate: Observing and documenting the rustproofing process</td>
<td>P1, P3 and P5</td>
<td>Breaks</td>
</tr>
<tr>
<td>Lesson 6</td>
<td>Explain: Explaining and sharing group work</td>
<td>P2, P3 and P5</td>
<td>Class</td>
</tr>
<tr>
<td>At home</td>
<td>Extend: Consolidating and extending knowledge related to rustproofing</td>
<td>P5</td>
<td>Home</td>
</tr>
</tbody>
</table>

Note. *P1 = working on real problems; P2 = encouraging diverse ideas; P3 = providing collaborative opportunities; P4 = using authoritative sources constructively; P5 = performing a formative assessment.

Table 4 shows that the teacher flexibly adopted different principles at different stages of the science inquiry. We also asked the teacher to reflect on his pedagogical plan and enactment process and outcomes based on guided questions on the inquiry-based learning approach and instructional principles, and arranged a time to interview him to hear his reflections. Some of the questions (Q) and excerpts from his reflections (R) are presented as follows:

**Q1:** When teaching the rustproofing learning unit, do you think it is important for the students to conduct hands-on experiments in an authentic environment? Why?

**R1:** If I told the students the reasons for rusting and how to prevent rusting directly, it would take 3 minutes. However, by providing opportunities for the students to lead their own science inquiry, they not only tested their own hypotheses in the experiments, but also underwent a process of discovery and collaboration: to identify problems in their everyday lives, raise questions to explore resources and conduct experiments to solve the problems.

**Q2:** Did you encourage the students to express their diverse ideas in their inquiry? How?

**R2:** I valued each student’s questions. How? Digital technology and the Internet extend our learning spaces. I seldom asked students to ask questions face to face in class. They could post their questions anytime, anywhere to the Edmodo social network platform, both inside and outside the classroom. They can get quick feedback from their peers or from me. If I found that many students were concerned about a problem, I would discuss the problem in class. Using Edmodo, all students’ questions are treated equally and their learning is extended beyond the classroom. This can be called seamless learning.

It is worth noting that the teacher’s good understanding of inquiry-based pedagogies and the principles of working on real problems and encouraging diverse ideas allowed him to apply the instructional principles in his pedagogical practices in multiple contexts, with the support of the social network platform.

**Effect on students’ domain knowledge gains and inquiry skills development**

We investigated the effect of principle-based pedagogical design on students’ domain knowledge gains through pre- and post-domain tests on rustproofing. Table 5 shows the results of the tests. Significant differences were found between the pre- and post-domain test results (pre-average score = 11.64; post-average score = 22.50, \( p < 0.05 \)). We can thus conclude that the students made significant advancements in their rustproofing knowledge after their inquiry.

Table 5. Pre- and post-domain test results on the rustproofing learning unit

<table>
<thead>
<tr>
<th>Class</th>
<th>No. of students</th>
<th>Pre-test M</th>
<th>S.D.</th>
<th>Post-test M</th>
<th>S.D.</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4C</td>
<td>27</td>
<td>11.64</td>
<td>2.652</td>
<td>22.50</td>
<td>3.837</td>
<td>14.854*</td>
</tr>
</tbody>
</table>

*Note: Total test score = 40

\( *p < .05 \)
In terms of the rustproofing learning unit assignments, the average scores for the 27 students were 79% (7.11 out of 9 questions in total). Although some of the students' scores were not high, their worksheets revealed that in many cases marks were deducted due to incorrect rendering of Chinese characters rather than their content knowledge. Figures 3(a) and 3(b) show screen captures of two students' worksheets marked by the teacher. Figure 3(a) shows that the student used the incorrect Chinese word “份 (part)” rather than “分 (component)” – the two words are the same in Pinyin, and their characters are also similar. Figure 3(b) shows that one student did not know how to write the Chinese word “滅 (extinguish).”

The effect of the principle-based pedagogical design on students’ inquiry skills was examined through pre- and post-questionnaire surveys. The results are shown in Table 6. They reveal that only the pre- and post-questionnaire results for item 1 (I know how to start thinking about how to solve a scientific problem) (mean = 4.35) and item 10 (I know how to explain my ideas to my peers when learning science) (mean = 4.16) showed significant differences. Items 1 and 2 relate to the “questioning” and “explaining” inquiry skills, respectively, and indicated an improvement in the students’ skills in raising questions and explaining ideas and concepts to peers.

Table 6. Pre- and post-questionnaire results on the rustproofing learning unit (translated version)

<table>
<thead>
<tr>
<th>Question</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>1. I know how to start thinking about how to solve a scientific problem</td>
<td>3.77</td>
<td>0.652</td>
<td>4.35</td>
</tr>
<tr>
<td>2. I know how to solve a scientific problem step by step</td>
<td>3.88</td>
<td>0.816</td>
<td>4.31</td>
</tr>
<tr>
<td>3. I know how to find scientific problems that I am interested in solving</td>
<td>4.08</td>
<td>0.935</td>
<td>4.15</td>
</tr>
<tr>
<td>4. I know where to find the information to solve a scientific problem</td>
<td>4.00</td>
<td>0.980</td>
<td>4.27</td>
</tr>
<tr>
<td>5. I know how to explore information/resources on my own when solving a</td>
<td>3.96</td>
<td>0.824</td>
<td>4.08</td>
</tr>
<tr>
<td>scientific problem</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I know how to improve the ways to solve a scientific problem</td>
<td>3.88</td>
<td>0.864</td>
<td>4.27</td>
</tr>
<tr>
<td>7. I know how to try different ways of solving a scientific problem</td>
<td>4.13</td>
<td>0.850</td>
<td>4.13</td>
</tr>
<tr>
<td>8. I know when to ask help from my peers when learning science</td>
<td>4.04</td>
<td>1.076</td>
<td>4.08</td>
</tr>
<tr>
<td>9. I know when to ask help from teachers when learning science</td>
<td>3.88</td>
<td>0.900</td>
<td>3.88</td>
</tr>
<tr>
<td>10. I know how to explain my ideas to my peers when learning science</td>
<td>3.72</td>
<td>0.891</td>
<td>4.16</td>
</tr>
<tr>
<td>11. I know how to explain my ideas to my teacher when learning science</td>
<td>4.24</td>
<td>0.970</td>
<td>4.20</td>
</tr>
<tr>
<td>12. I know how to work together with my peers to solve a scientific</td>
<td>3.88</td>
<td>1.035</td>
<td>4.29</td>
</tr>
<tr>
<td>problem</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3.87</td>
<td>0.900</td>
<td>4.15</td>
</tr>
</tbody>
</table>

* p < .05

Ways that the students advanced their domain knowledge and developed their inquiry skills

We traced the artefact development of each student group in the rustproofing learning unit. Each group designed three methods. Among the 18 experimental methods proposed by the groups, only one method (i.e., using paint to coat the clip) was the same between two groups. All of the other methods were different from each other (see Table 7). In addition, after tracking the experiment outcomes, the degree of rusting in each method was evaluated and scored from 0 (no rusting) to 10 (most severe rusting), which is also shown in Table 7.
Table 7. Experimental designs and rustproofing outcomes by group

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of design</th>
<th>Hypothesis of rustproofing theories using the methods. (All of the following methods were hypothesised to prevent air from entering the cup or to prevent water from having direct contact with the clip.)</th>
<th>Degree of rusting</th>
</tr>
</thead>
<tbody>
<tr>
<td>G 1</td>
<td>(1)</td>
<td>Use Vaseline and paint to coat the clip and put oil into the cup along with the water.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use Vaseline to coat the clip.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Use paint to coat the clip.</td>
<td>7</td>
</tr>
<tr>
<td>G 2</td>
<td>(1)</td>
<td>Use Vaseline to coat the clip and then wrap the clip with plastic wrap. Put oil and w-4 into the cup.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use Vaseline to coat the clip and then wrap the clip with plastic wrap. Put oil into the cup.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Use Vaseline to coat the clip and then wrap the clip with plastic wrap.</td>
<td>3</td>
</tr>
<tr>
<td>G 3</td>
<td>(1)</td>
<td>Put the clip into a storage bag.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use Vaseline to coat the clip, and then use the dryer to dry the clip before putting it into the water.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Use a candle to dry the clip after removing it from the water.</td>
<td>3</td>
</tr>
<tr>
<td>G 4</td>
<td>(1)</td>
<td>Wrap the clip with glue paper.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Wrap the clip with plastic wrap.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Put oil into the cup.</td>
<td>2</td>
</tr>
<tr>
<td>G 5</td>
<td>(1)</td>
<td>Use Vaseline to coat the clip and then put oil and soya oil into the cup.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use Vaseline to coat the clip and then put oil into the cup.</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Wrap the clip with glue paper and then put the oil into the cup.</td>
<td>7</td>
</tr>
<tr>
<td>G 6</td>
<td>(1)</td>
<td>Use paint to coat the clip, seal the clip inside a bottle and then put the bottle into the water.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>Use paint to coat the clip, use a dryer to dry the paint, seal the clip inside a bottle and then put the bottle into the water.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>Use paint to coat the clip.</td>
<td>7</td>
</tr>
</tbody>
</table>

We also examined the students’ captured photos daily to keep an observational record of the three experiments in each group. We obtained some of the artefacts created by Group 2 as an example. Figure 4(a) shows a screenshot of the three experimental results on the first day using the three methods (see Table 7). Figure 4(b) shows a screenshot of the experimental results on the last day (1 week after). Figure 4(c) shows the degrees of rusting (0-10) in the group’s three methods, evaluated by the group itself, peers in other groups and the teacher (3, 0 and 0, respectively). These artefacts documented the students’ deepened understanding of rustproofing from experimental design through to observation, presentation and evaluation.

![Figure 4](image.png)

*Figure 4. Photo of the first day of the experiment; photo of the second day of the experiment; photo of the experimental results evaluation (a – left; b – middle; c – right)*

Group 2 was chosen as having the best experimental rustproofing design, and was awarded a badge declaring its members rustproofing “experts,” which encouraged the students to make further science inquiry. It is worth noting
that although the students captured photos of the experimental results each day, they rarely entered observational records on the degree of rusting into the “Observational Forms for Rustproofing Experiments.”

To examine how the students advanced their inquiry skills, we investigated students’ inquiry processes supported by Edmodo, where seamless learning across physical spaces (in the classroom and between lessons using tablet PCs, and at home using various devices) and individual and social spaces (online learning and class interactions) was documented. Figure 5 indicates the limited numbers of students’ posts on Edmodo relating to their science inquiry skills (questioning = 1.7%, exploring = 2.2%, explaining = 5.1%, evaluating = 3.9% and extending = 0%). However, other posting categories accounted for the majority of the contributions: coordinating the experimental designs and observations (47.9%), social interactions (23%), greetings (14%) and news sharing (8%). This indicates that the social network platform played an important role in students’ project work orchestration and establishment of intimate relationships.

<table>
<thead>
<tr>
<th>Inquiry skills</th>
<th>No. (%) of postings on Edmodo</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questioning</td>
<td>7 (1.7%)</td>
<td>What shall I use in order not to make the oil not spill over?</td>
</tr>
<tr>
<td>Exploring</td>
<td>9 (2.2%)</td>
<td>How many ways are there to prevent rusting (with Hyperlink)?</td>
</tr>
<tr>
<td>Explaining</td>
<td>21 (5.1%)</td>
<td>This is the method of using oil to prevent rusting (Hyperlink: iron rusting- [Wiki – Wikipedia – the free encyclopedia])</td>
</tr>
<tr>
<td>Evaluating</td>
<td>16 (3.9%)</td>
<td>Good!</td>
</tr>
<tr>
<td>Extending</td>
<td>1 (0)</td>
<td>Gilded iron or metal can prevent rusting</td>
</tr>
<tr>
<td>Others</td>
<td>355 (86.7%)</td>
<td></td>
</tr>
<tr>
<td>Coordinating</td>
<td>196 (47.9%)</td>
<td>Chu L, don’t forget to bring the hair dryer.</td>
</tr>
<tr>
<td>Interaction</td>
<td>94 (23%)</td>
<td>You get my phone, Ho Y?</td>
</tr>
<tr>
<td>Greetings</td>
<td>61 (14.9%)</td>
<td>Good morning, everyone!</td>
</tr>
<tr>
<td>News sharing</td>
<td>8 (2%)</td>
<td>The second group got the honor of “Rustproof Expert”</td>
</tr>
<tr>
<td>Total postings</td>
<td>409 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5. Categories of students’ posts on Edmodo related to rustproofing and other inquiry skills

We also interviewed a group of students using the questions guided by the five elements in the inquiry-based learning model to understand their rustproofing inquiry processes. The students expressed great passion for leading their own rustproofing research in real life (supported by Edmodo), where their inquiry experience bridged seamlessly across different spaces. One student made the following statement.

In the past, in science class, we only learned from textbooks. It was boring. Now, we can explore and discuss the best ways for our experimental designs on rustproofing. It is interesting that we can do hands-on experiments ourselves, observe the process of the rustproofing process and take pictures [of the clip rusting results] day by day . . . If I learned [rustproofing] from textbooks, it would be easy for me to forget what I have learned. But doing hands-on experiments makes it difficult to forget the methods and principles [of rustproofing]. I understand the rustproofing theories better.

The students’ interview results indicate that they acquired solid domain knowledge and developed inquiry skills from the experience of hands-on experimental design and practice.
Discussion

Principle-based pedagogical design—gains

This study attempted to develop and implement a principle-based pedagogical design for inquiry-based learning in the seamless learning environment of a Hong Kong primary school. The results show that a teacher with training experience and a good understanding of the principles and inquiry-based learning model was better able to plan and enact student-centred inquiry activities in which students had control over their own learning. This echoes the findings in previous studies that suggest that enhancing teachers’ inquiry-based pedagogical competence in a technology-supported learning environment requires teachers to understand the basic theoretical principles behind inquiry-based pedagogies and how to apply the principles to pedagogical practices (Scardamalia, 2002; Song & Looi, 2012, Zhang et al., 2011). In our study, the students showed improvements in their domain knowledge learning and inquiry skills, especially in terms of “questioning” and “explanation,” which are considered essential elements of inquiry-based learning (Hakkarainen, 2003).

However, students cannot develop a meta-cognitive awareness of inquiry strategies without adequate scaffolding (Lakkala et al., 2005). Our study adopted a guided inquiry-based learning model by providing prompts (“Forms for Experimental Rustproofing Designs” and “Observational Forms for Rustproofing Experiments”) for the students’ experimental designs and observations. Although each group of students generated different experimental designs and hypotheses, they were all on the right track in the inquiry. According to Lin and Lehman (1999), metacognitive skill development is typically fostered by providing students with opportunities to reflect on and monitor their learning performance and revise their investigatory strategies. In this regulative process, students are reflective inquirers looking to accomplish projects and gain a deeper understanding of domain knowledge and inquiry skills (Loh et al., 2001). Further, in the “Forms for Experimental Rustproofing Designs,” the students were required to take on different responsibilities in completing the experiment, which increased their awareness of taking collective responsibility for advancing the group’s knowledge (Scardamalia, 2002; Zhang et al., 2011).

In this study, the Edmodo social network platform provided a seamless learning environment for the students to coordinate the inquiry projects and establish a rapport in groups and with peers, which played an important role in advancing their rustproofing knowledge and developing their inquiry strategies. In addition, the students could share their groups’ products with their peers at any time and anywhere on the platform, which allowed them to evaluate other groups’ work and construct knowledge collaboratively. The embedded peer group assessment in the pedagogical design meant that the assessment responsibility was turned over to the students, helping them to develop increased agency when evaluating their own learning progress (Zhang et al., 2011). The students could also submit assignments to the platform and obtain the teacher’s feedback in a timely manner, which encouraged them to learn. This study contributes to the literature on the use of principle-based pedagogical design for guided inquiry-based learning in science in a seamless learning environment at primary level. Nevertheless, it also has some limitations.

Principle-based pedagogical design—losses

We identified several issues and limitations to be addressed in future work. First, in terms of the five instructional principles adopted in the research, the teacher was not able to grasp the gist of the principle of using authoritative information constructively in his reflections. Some students copied and pasted information from the Internet directly without acknowledging the source. However, the teacher believed that as long as he asked the students to explore learning resources on the Internet, he was adhering to the principle. He did not further scaffold the students on making constructive use of sources. According to Yeo and Tan (2010), the constructive use of authoritative sources involves the interpretation of meaning in context and plays an important role in deepening and expanding students’ science domain knowledge. Hence, in our next cycle of research, we must elaborate this principle to the teacher. Second, the teacher designed the ‘Observational Forms for Rustproofing Experiments’ (see Appendix II) for the students to document the daily degree of rust over five consecutive weekdays. The findings show that none of the six groups completed this task. The ‘Degree of rusting of the iron clip: 0/10’ item might have been too abstract for Primary 4 students to estimate and record. Nevertheless, the students took some pictures to document the daily rusting process during their break time. In providing scaffolds such as prompts, we suggest that the teacher must...
consider the students’ level and cater to their needs. Finally, the results of the research cannot be generalised due to the short time span. Further interactive studies are required to investigate whether the five instructional principles suffice for developing teachers’ principle-based understanding of inquiry-based learning, whether the inquiry-based learning model must be refined and how to make better use of social network technology to support seamless learning environments.

Conclusions and future work

This study explores the use of a principle-based pedagogical design for inquiry-based learning in a seamless learning environment, with resource access and peer interactions initiated in classrooms and extended to online interactions. The inquiry-based learning approach was integrated into the domain knowledge learning process to promote students’ development of inquiry skills. The results demonstrate the effective development and implementation of the pedagogical design in a rustproofing learning unit and the positive effect of the design on students’ domain knowledge gains and inquiry skills. The results reveal the need for further research to accumulate experience and scale-up pedagogical interventions within and across schools. The future scale-up of research efforts should be planned under the “design-based implementation” and “designing for diffusion” approaches (Dearing & Kreuter, 2010; Penuel, Fishman, & Cheng, 2011) to build capacity among teachers in the within- and cross-school settings. Further research will instil in target teachers a positive perception of the research efforts that address persistent practice problems from multiple perspectives during the capacity scale-up across the Hong Kong primary school sector.

References


Appendix I
Forms for experimental rustproofing designs

<table>
<thead>
<tr>
<th>Group:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group members:</td>
</tr>
<tr>
<td>1.</td>
</tr>
<tr>
<td>2.</td>
</tr>
<tr>
<td>…</td>
</tr>
</tbody>
</table>

*Experimental design 1 (Experimental designs 2 and 3 used the same forms as experimental design 1.)*

<table>
<thead>
<tr>
<th>Material</th>
<th>Student in charge</th>
<th>Hypotheses of rustproofing theories</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>…</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Design diagrams (in order)

Appendix II
Observational forms for rustproofing experiments

<table>
<thead>
<tr>
<th>Observation date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation time</td>
</tr>
<tr>
<td>Observer(s)</td>
</tr>
</tbody>
</table>

*Experimental design 1 (Also experimental designs 2 and 3)*

Degree of rusting of the iron clip: 0/10
Spatial Visualization Learning in Engineering: Traditional Methods vs. a Web-Based Tool

Carlos Melgosa Pedrosa, Basilio Ramos Barbero* and Arturo Román Miguel

Graphic Expression Department, Higher Polytechnic School, University of Burgos, Avda. Cantabria s/n. 09006 Burgos, Spain // cmelgosa@ubu.es // bramos@ubu.es // aroman@ubu.es

*Corresponding author

(Submitted June 21, 2012; Revised February 1, 2013; Accepted April 12, 2013)

ABSTRACT

This study compares an interactive learning manager for graphic engineering to develop spatial vision (ILMAGE_SV) to traditional methods. ILMAGE_SV is an asynchronous web-based learning tool that allows the manipulation of objects with a 3D viewer, self-evaluation, and continuous assessment. In addition, student learning may be monitored, which saves a significant amount of time for the teacher, as both correction and grading are automatically performed. Our aim in this study is to establish whether the application is an effective tool for learning spatial visualization. Students of engineering graphics following an industrial engineering degree course at the University of Burgos (Spain) participated in a pilot project over two academic years. The students were separated into two groups: an experimental group that studied with the ILMAGE_SV web application and a control group taught with traditional methods. Our study demonstrates that the results of both methods, with regard to the development of spatial vision, are in general equal. Nevertheless, ILMAGE_SV is more efficient for students who experience greater difficulties with spatial vision and for students with no prior knowledge of technical drawing.

Keywords

E-Learning, Web3D, Virtual interactive learning, Interactive graphics, Spatial ability

Introduction

In a previous study (Melgosa, Ramos, & Baños, in press), an interactive tutorial (IT) was presented, which functions as a web-based spatial visualization ability (SVA) learning support tool for students of engineering graphics and as an administrative tool for teachers to track student learning. It is an open-access Internet application that students can use to complete exercises and exams taken at random from its database or as directed by the teacher. IT_SVA is designed to communicate knowledge of SVA more evenly and to address weaknesses in SVA among students following undergraduate engineering degrees. Continuous assessment and follow-up of student learning achievements on engineering courses are requirements of the guidelines for the European Higher Education Area.

The IT_SVA serves three types of users (student, teacher, and administrator) and has four main parts: (1) a content management system (CMS); (2) an intelligent tutorial system (ITS); (3) a web-based tool for exercise management and correction; (4) a database. ILMAGE_SV (Interactive Learning Manager for Graphic_Engineering: Spatial Vision) is available at the following URL http://www2.ubu.es/calcubu/ and is a type of IT-SVA. The main difference of this tool compared to other tutorial learning systems is its interactivity due to the use of 3D viewers, allowing computerized simulations of the mental process of rotating 3D models.

The use of ILMAGE_SV, associated with the methodology proposed by Pérez and Serrano (1998) (summarised in section Justification), helps train students to develop spatial visualization skills and brings with it four important features:

1. A 3D model that may be manipulated on-screen as a tool to assist with mental visualization and rotation exercises of particular difficulty to students.
2. Preliminary tests that may be taken to identify the level at which the student experiences spatial visualization difficulties and that recommend a starting level in the ILMAGE_SV application to the student.
3. Instant access to self-evaluation records for the student.
4. Automated tracking of student learning for the teacher.

The main differences between our application and some (but not all) of the other web applications are that it has manipulable 3D objects, calculates and displays the results and grades of the finished exercises, tracks learning
achievements on a database, and has been used and/or validated in comparative studies. We highlight that the possibility of manipulating interactive 3D models through a 3D viewer was the feature that our students rated most highly. We therefore consider that these Web3D tools should incorporate a 3D viewer.

It is evident that engineering students should possess very well developed SVA, so that they can effectively communicate and progress as professional engineers (Ferguson, Ball, McDaniel, & Anderson, 2008; Brus, Zhoa, & Jessop, 2004; Sorby, 2001), as one of their core competences is the drafting and execution of projects. A definition given by McGee of spatial visualization is: “The ability to mentally rotate, twist, or invert pictorially presented visual stimuli” (McGee, 1979). Various experiments to improve this ability have been performed, some in a significant way, such as those in which students have real models in their hands to understand what they have to visualize (Ferguson et al., 2008; Sorby, 2001). We proposed the use of virtual on-screen models that may be manipulated by means of 3D viewers. Their manipulation helps students with the process of mental rotation of objects. This idea is also shared by Piburn, Reynolds, McAuliffe, Leedy, and Birk (2005), who argued that the manipulation of 3D computer objects in a virtual terrain could significantly improve students’ spatial abilities.

ILMAGE_SV has been validated with engineering graphics students at the University of Burgos (Spain) by means of a user-satisfaction survey over two academic years. We may highlight the survey results, which suggest that more than 89% of our students would use this application again to improve their spatial visualization skills and that 98% would recommend this application to other people with an interest in SVA.

However, we also think we should confirm whether ILMAGE_SV is an effective learning tool for graphic engineering students, which is precisely our aim in this study.

Figure 1 shows one of the various types of exercises stored on the database, in which the student can manipulate the 3D model with the eDrawings viewer that appears on the left. The set problem is described in the middle and, in the column on the right, the student types in the answers in the form of letters that correspond to the numbered surfaces shown in the problem statement. In the example, the two first correct answers have been introduced, the third is incorrectly answered, and no further answers have been given. The screen shown in Figure 1 appears when the button is activated to view the results, giving the number of correct and incorrect responses and unanswered questions, as well as the grade obtained in the exercise. At the same time, the database stores the results so that the teacher may use the ILMAGE_SV-based student evaluation.

We previously demonstrated that the design, structure, resources, and ease of use of the ILMAGE_SV web application are of sufficient quality to be used as a teaching tool. They have also confirmed that the content of the ILMAGE_SV application in terms of its capacity to motivate, its usefulness, and its appropriateness are suitable to
help students develop spatial visualization skills (Melgosa et al., in press). However, the effectiveness of ILMAGE_SV as a tool that improves the learning of spatial visualization was not evaluated. Thus, we wish to confirm in this article whether ILMAGE_SV is an effective tool for learning spatial visualization, if its use improves learning in comparison with traditional methods, whether it is better for students who experience greater difficulties with spatial visualization, and whether it is better for those students with no previous knowledge of technical drawing.

**Justification**

At present, universities in Europe are currently in the process of introducing university studies according to the criteria of the Bologna plan, which focus on student learning rather than on a teacher-led approach. This situation, together with the emergence of CAD, has led many universities to consider either total or partial elimination of descriptive geometry from their engineering course curricula. There are more and more researchers who acknowledge the need to search for alternatives to descriptive geometry, in order to develop spatial visualization skills on engineering courses.

Various researchers (Hake, 2002; Sorby, 2000; Miller & Bertoline, 1991; Scribner & Anderson, 2005) believe that the most critical component of graphic representational skills is the ability to apply spatial visualization to objects. Contero, Company, Naya, and Saorín (2006) and Connolly (2009) described the continued importance of spatial reasoning in the engineering curricula and stressed that it should be taught using sketching as well as modern technology. From their perspective, emphasis should be placed on orthographic projection skills, mental imagery of 3D objects, and the use of web-based drills, interactive multimedia, and tutorials.

Web3D has great potential for a wide number of educational applications that require visual understanding (Strong & Smith, 2001; Web3D Consortium, 2010), although research into educational techniques and systems associated with its use is very limited. The 3D models (which can be animated) allow students to understand aspects of the taught subject that are not clearly seen in an image, when they are hidden inside the models. It should be noted that more emphasis has been placed on the visualization of 3D objects, because 3D immediately enhances the learning process (Liarokapis et al., 2004). Students can explore 3D visualization through interactive Web3D content of the teaching material, which helps them to understand it more effectively.

According to Ault and Samuel (2010), visualization skills are a strong indicator of success in engineering, science, and a variety of other careers. Studies have shown that training can enhance visualization skills in a relatively short time. Researchers generally agree that spatial visualization skills are enhanced by 3D drawing and by manipulating physical 3D objects. Traditionally, engineering graphics courses have included a strong component of descriptive geometry and sketching. Since the advent of computer-aided design systems in the early 1980s, nearly all US engineering schools have ended their courses in descriptive geometry and most schools have also dropped manual drafting and sketching from their introductory graphics courses. Universities around the world have followed suit. As a result, there has been a noticeable decline in the visualization skills of engineering students.

There are now numerous studies that apply CAD, the manipulation of 3D models in virtual environments, or the completion of different types of tests such as those used by Sorby, Wysocki, and Baartmans (2003), in order to improve spatial visualization. For example:

- Tsutsumi, Schröcker, Stachel, and Weiss (2005), using the mental cutting test (MCT), confirmed in a significant way that students following descriptive geometry courses obtained better visualization results than those taking design courses.
- Sorby and Baartmans (1996), in the conclusions to a course on 3D visualization skills, stated that the completion of computer-based exercises (in this case, with I-DEAS software) improved learning among students with greater learning difficulties.
- Rafi, Anuar, Samad, Hayati, and Mazlan, (2005) also confirmed with significant results that the use of Web3D applications as pedagogic tools with 3D models in VRML format improved the development of spatial skills.
- Ferguson et al. (2008) found that the PSVT:ROT (Purdue Spatial Visualization Test: Rotations) test led to significant improvements in the spatial visualization of students that had real models in their hands to perform the different practical exercises, as opposed to students who did not physically have these models in front of them.
Even so, other investigations suggest otherwise:

- Leopold (2005), Sorby and Gorkska (1998), Sorby (2000), Yue (2001), and Godfrey (1999) all suggested that 3D CAD experience in itself does not appear to enhance visualization skills.
- Koch (2006) found no significant differences between drawing and solid modelling design methods used for solving technical problems ($p = .752$).
- Konukseven (2010) compared traditional teaching methods on an engineering graphics course with a 3D-based teaching method in VRML format, which engineering students could visualize on the Web. His results suggested both methods were of the same quality and showed that an effective way to keep students active is to use creative visualization and to offer them opportunities to interact with the courseware.

In our opinion, the use of 3D CAD software and the 3D viewers, which assist with mental manipulation of the models and thinking in 3D, produces significant results. Nevertheless, there is no significant improvement in visualization skills when CAD software is used for 3D modelling.

We have to advance in the use of self-correction test-type exercises in the educational virtual environments (EVEs), in order to try to improve the development of visualization skills. These EVEs should have tools that measure student learning and Web3D models as a resource to support learning and to improve spatial visualization skills, formerly acquired in descriptive geometry classes. The most recent literature suggests that, among those methods that use test exercises to improve visualization ability, the enhancing visualization skills-improving options and success (EnViSIONS) project is the most extensively used.

Teachers on the EnViSIONS project (Veurink et al., 2009) developed spatial visualization ability on an introductory course in 10 study modules: isometric sketching; orthographic projection (normal surfaces, orthographic projection); inclined and single curved surfaces (not included in workbook/software [Sorby et al., 2003]); flat patterns; rotation of objects about a single axis; rotation of objects about two or more axes; object reflections and symmetry; cutting planes and cross sections; surfaces and solids of revolution; and combining solids. Following their use at six universities, these materials have been shown to improve visualization significantly. Three minimum study modules are necessary, according to Veurink, Hamlin, and Sorby (2008), to obtain significant results: isometrics; orthographics; and rotations about a single axis.

We applied the methodology suggested by Pérez and Serrano (1998) in ILMAGE_SV, to develop spatial visualization skills that are divided into six levels: (1) identification and recognition, (2) understanding, (3) application, (4) analysis, (5) synthesis, and (6) evaluation. We included the first four levels, cited earlier, and held the fifth and sixth levels outside ILMAGE_SV in the classroom. Pérez and Serrano (1998) demonstrated, in a significant way, that the performance of 72.6% of students, who began with this methodology, was above a pre-established average level, as against the 47.4% of students who performed above that same level before the introduction of this training. This methodology, in relation to the assessment of spatial visualization parameters, is at an intermediary position between the three components proposed by Veurink and the ten modules in the EnViSIONS project.

Since 2003, a group of teachers at the University of Burgos have been firmly convinced that the spatial visualization skills of students enrolling on engineering courses have been weaker than in previous years, an opinion that is also expressed by other authors (Veurink et al., 2009; Sorby, 2009; Duff & Kellis, 2009; Brus & Boyle, 2009; Knott & Kampe, 2009). One of the causes is that technical drawing is an optional material in studies leading up to pre-university exams (the Bachillerato, in Spain). This reason coupled with the shorter length of study plans, the emergence of CAD, and an increasing number of exercises done on 3D all encourage us to look for alternatives to traditional teaching to improve spatial visualization.

**Design of the investigation**

The graphic expression study module at our university is divided into four parts: technical drawing, geometry, descriptive geometry and CAD. On this study module, although spatial visualization improves over the academic year, the fundamental concepts—types of projections, principal views, and minimum necessary views—are learnt in the first part (technical drawing) and at the start of the course. This investigation focuses on those concepts and related practical sessions, which amount to three classroom hours of taught classes, to which we should add the private study time of each student.
The design of the investigation applied two different methods to test the efficiency of the two methods and to confirm which was best for the development of spatial visualization. One group of students used a traditional method (T), in which the teacher explained the concepts in class and the students completed three practical drawing exercises. The other group used the experimental method (E), in which the students had access, in a computer room, to the ILMAGE_SV software package described above. This application has three videos with the same theoretical concepts as in group T, in which the teacher sets up three practical exercises with 52 test-type questions, as shown in Figure 1. The students worked their way through the graded application levels exercises progressively: identification of surfaces, main views, developments, contact surfaces between blocks, and minimum necessary views.

We used the following indicators in the design phase of the pilot project, in order to form homogeneous groups of students, to ascertain which students held previous knowledge and to compare the visualization results between groups:

- A test of spatial ability, at the start of the study module to classify students by levels of spatial visualization.
- The previous studies completed by students, to ascertain whether the newly enrolled students have previous knowledge of technical drawing.
- An exam at the end of the pilot project, divided into the variable “vision” (spatial visualization ability) and the variable “sketch” (sketching skills).
- An exam upon completion of the course with a final visualization test, in order to confirm the influence of the study module’s other components on the development of spatial visualization in groups E and T.

Initially, the pilot project was solely designed for the 2009–10 academic year, but the results were not significant because some of the samples were small. Another pilot project was repeated with the same variables in the 2010–2011 academic year. After the pilot project, in the first academic year, we agreed to introduce “increase in knowledge” as a new variable (INCR_MRT), which required a new test to assess the improvements in SVA. This test was performed both before and after the pilot project in the second academic year.

**Measurement instruments**

At the start of the course the differential aptitude tests-space relations (DAT-SR) test enabled us to classify students by levels of aptitude for spatial visualization. This test, with a total of 60 questions, requires mental manipulation of objects in a tri-dimensional space. It involves identifying one object among the four options, which, when unfolded, corresponds to the surface development of a die (Figure 2).

![Figure 2. Example of the DAT-SR test](image)

Before and after the experiment, in order to establish the increase in spatial visualization knowledge, we used the mental rotations test (MRT), which measures mental rotation ability. The test consists of 20 questions. Each item involves identifying two figures, from among four possibilities, that correspond to the drawing on the left, but have been rotated by several degrees (Figure 3).

![Figure 3. Example of MRT test](image)
At the end of the pilot project, in order to ascertain the extent of student learning achievements with the two methods (E and T) and which method was the most successful, they take an exam in two parts to test the vision variable and the sketch variable. One example of this exam, in Figure 4, consists of three exercises. The first involves hand sketching a view of a part seen from point A, in which the part must be mentally rotated on its axes; the second consists of sketching the way the part develops; and the third, of sketching the minimum necessary views of the part. The intention is to assess spatial visualization skills (mental rotations and spatial visualization).

Lastly, a new test, the “final vision test” variable, was introduced at the end of the course, which was intended to track student progress throughout the course in groups E and T, in the development of spatial visualization. This test was obtained from the selection of 24 test questions taken from the DAT-SR (Bennett, Seashore, & Wesman, 1997), MRT (Vanderberg & Kuse, 1978), PSVT-ROT (Guay, 1977), and the Lappan (1981) tests.

The Lappan test consists of 32 items, each of which shows the projection of a construction formed by cubes and only one out of five options has to be selected, which corresponds to another of its projections. The PSVT-ROT test comprises 30 items, where one of five possible responses corresponds to a piece that has been rotated 90 or 180 degrees about one of its axes.

**Pilot project**

In the first week of the course, all students on the University of Burgos industrial engineering course, in each of the two academic years 2009–2010 and 2010–2011, took the DAT-SR test to ascertain their previous knowledge of spatial visualization. The percentile of the DAT-SR test was used with a coupling or pairing technique to distribute students between groups E and T. In this distribution of students between the two groups, the different control variables (gender, first-time enrolment, previous studies—pre-university or vocational training—and previous knowledge of technical drawing) were taken into account, beginning with gender, so that these variables would be balanced in the two groups. The pilot project over the two academic years involved 112 students in group T and 132 students in group E. The t-test statistically confirmed that these groups were equivalent with regard to spatial visualization, measured by the DAT-SR test.

During this first week of the 2010–2011 academic year, students from both groups also took the MRT test and repeated it after the pilot project, so as to validate any increase in their knowledge of spatial visualization.

The pilot project lasted for three hours over a period of three weeks, at the start of the course when students were learning about the basic concepts of spatial visualization.

Group T conducted three practices with manual exercises: identification of surfaces in views, sketching the way a part develops, principal views, minimum necessary views, and views when rotating a part around given axes (example Figure 5).

Students from group E conducted three practical exercises with ILMAGE_SV. They completed 52 test-type exercises of progressive difficulty, involving the same concepts as those for group T, but without sketching (Figure 6).
Following the pilot project, in the fifth week, a visualization exam was given to all students without the use of help tools and completed by hand sketching. In this exam, the variables “vision” and “sketch” were independently assessed (Section “Measurement instruments” and Figure 4).

The end of the pilot project coincided with the final study module exam, in which a test with the points described at the end of section 3.1 was added to the end-of-year exam.

Results of the pilot project and their analysis

In the second academic year of the pilot project, students completed the MRT pretest (40 point test) and repeated a MRT post-test after the pilot project ended. Eighty-five students participated in the pretest and 106 students in the post-test. After
the pilot project, it was confirmed that 76 students had responded to both tests, of whom, 45 belonged to the experimental group and 31 to the traditional group, all of which initially confirmed that both methods (T and E) were appropriate to improve spatial visualization after the experiment.

We present the analysis of five studies from among those performed after the pilot project had ended.

We studied whether the average of the post-test minus the pretest differences was significantly greater than zero, in groups T and E, to validate improvements with the two methods. To do so, the t-student statistics test was used. In this test, the MRT variable should follow a normal distribution in both groups, which is in fact the case, as we can confirm from Table 1.

<table>
<thead>
<tr>
<th>Group</th>
<th>Kolmogorov-Smirnov² Statistic</th>
<th>df</th>
<th>Sig.</th>
<th>Shapiro-Wilk Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>.128</td>
<td>31</td>
<td>.200</td>
<td>.956</td>
<td>31</td>
<td>.234</td>
</tr>
<tr>
<td>E</td>
<td>.127</td>
<td>28</td>
<td>.200</td>
<td>.972</td>
<td>28</td>
<td>.643</td>
</tr>
</tbody>
</table>

Note. a. Lilliefors Significance Correction. This is a lower bound of the true significance.

The “One-Sample T Test” run on the SPSS programme (Table 2) confirmed that the variable difference of post-test minus pretest (MRT_INCR) in the two groups was significantly greater than zero. Because \( \alpha < 0.05 \), the null hypothesis of equality to zero of improvement in both methods is rejected and we can accept that both methods have been shown to improve the learning of spatial visualization significantly.

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean of MRT_INCR</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>95% Confidence Interval of the Difference</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>3.806</td>
<td>4.875</td>
<td>.875</td>
<td>4.348</td>
<td>30</td>
<td>.000</td>
<td>2.018</td>
<td>1.741</td>
<td>1.277</td>
<td>-.816</td>
</tr>
<tr>
<td>E</td>
<td>5.321</td>
<td>4.869</td>
<td>.920</td>
<td>5.783</td>
<td>27</td>
<td>.000</td>
<td>3.433</td>
<td>5.321</td>
<td>1.741</td>
<td>-.815</td>
</tr>
</tbody>
</table>

In Table 2, we see that the improvement in the traditional group, at a confidence level of 95%, was at an average value of between 2 and 5.6 points, while the improvement in the experimental group stood at an average value of between 3.4 and 7.2 points. It may also be seen that the average increase is 3.8 points in group T and 5.3 points in group E.

A second analysis of these two pilot projects consisted of testing whether the improvement was greater in group E than in group T. The t-student statistical test was therefore completed for independent samples that compared the average of the variable increased visualization MRT_INCR for both groups E and T. The level of significance was 0.178 in the t-test for equality of measurements of the variable MRT_INCR (Table 3) and because \( \alpha > 0.05 \), we cannot reject the null hypothesis and therefore have to accept that the averages are equal, despite the average increase being 1.7 points greater than in group E with respect to group T.

<table>
<thead>
<tr>
<th>Group</th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRT_INCR</td>
<td>Equal variances assumed</td>
<td>Sig.</td>
<td>t</td>
</tr>
<tr>
<td>T</td>
<td>.189</td>
<td>.665</td>
<td>1.363</td>
</tr>
<tr>
<td>E</td>
<td>1.364</td>
<td>56.519</td>
<td>.178</td>
</tr>
<tr>
<td>MRT_INCR</td>
<td>Equal variances not assumed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

149
We can make a global comparison of the variable “exam” and individual comparisons of the variables “vision” and “sketch” in groups E and T. To do so, the averages were once again compared for the two groups using the SPSS software package t-test for independent samples. We also confirmed the equality of the variable visualization with the Mann-Whitney U non-parametric statistical test, which is used for abnormal variables.

**Table 4. Independent samples test**

<table>
<thead>
<tr>
<th></th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Exam</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>.043</td>
<td>.837</td>
<td>- .624</td>
<td>242</td>
<td>.533</td>
<td>-.116</td>
<td>.187</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>- .627</td>
<td>238.59</td>
<td>.532</td>
<td>-.116</td>
<td>.187</td>
<td>-.482</td>
<td>.250</td>
</tr>
<tr>
<td><em>Vision</em></td>
<td>.513</td>
<td>.475</td>
<td>- .382</td>
<td>242</td>
<td>.703</td>
<td>-.055</td>
<td>.144</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>- .383</td>
<td>238.65</td>
<td>.702</td>
<td>-.055</td>
<td>.143</td>
<td>-.337</td>
<td>.227</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Sketch</em></td>
<td>2.751</td>
<td>098</td>
<td>-1.060</td>
<td>242</td>
<td>.290</td>
<td>-.061</td>
<td>.058</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>-1.072</td>
<td>241.91</td>
<td>.285</td>
<td>-.061</td>
<td>.057</td>
<td>-.175</td>
<td>.051</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The t-test in Table 4, regardless of whether equality of variances is assumed, gives us a signification value that is greater than 0.05. The difference in means is not significant in the three tests (“exam,” “vision,” and “sketch”), so we must assume the equality of means. In this Table, we can see that the difference of means assumes negative and positive values in the confidence interval.

In Table 5, the level of significance of the variable vision with the Mann-Whitney U non-parametric test was 0.8 > 0.05. The difference of means for the variable visualization is not significant, so we assume equality.

**Table 5. Test statistics**

<table>
<thead>
<tr>
<th>Test Statistics</th>
<th>Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>7258,500</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>16036,500</td>
</tr>
<tr>
<td>Z</td>
<td>- .243</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.808</td>
</tr>
</tbody>
</table>

*Note. a. Grouping Variable: Num_group*

We may also affirm that the introductory component in ILMAGE_SV neither significantly improves nor worsens learning of the variable “vision.” Furthermore, the two experimental groups obtained similar results for sketching, which is confirmed in Table 4, despite the various exercises completed in ILMAGE_SV in group E, in comparison with manual completion by hand sketching of the exercises in group T. Finally, Table 4 confirms that the score for sketching is slightly higher in group T compared to group E, but not significantly. We may therefore say that students who used ILMAGE_SV in the pilot project (Group E) did not obtain worse results than students from the traditional group (Group T) in a significant way in the assessment of sketching, despite having dedicated fewer hours to drawing during the pilot project.

At the end of the academic year, all students completed a new visualization test called the “final vision test” as part of the final exam in the study module. The results of the differences in means, when comparing the two groups, E and T, were not significant, so we may only assume that the two methods were equally valid.
In a third study, we identified the students for whom ILMAGE_SV was more appropriate. We divided the students who had used ILMAGE_SV into three categories of spatial ability: those of greater ability (1), average ability (2) and lower ability (3). The division into each category was done by 35/65 percentile values of students who had completed the MRT pretest, and the study variable was MRT_INCR.

Figure 7 shows that the average increase in knowledge in the MRT test is greater for students with weaker abilities than in the other two categories. All three categories were compared to see whether significant differences existed between them. The result of the ANOVA statistical test was $F = 3.9$, with a significance level of $0.028 < 0.05$, which indicates that at least one of the three categories has a different average, but does not indicate which one.

Table 6 shows the tests chosen from among various alternatives to ascertain the categories with significant differences: Tukey’s honestly significant difference (Tukey HSD) test and the Scheffé test. The Scheffé test is the most conservative. Both tests coincide in as much as a significant difference only exists between the averages of categories 1 (greater spatial ability) and 3 (lower spatial ability).

Table 6. Multiple comparisons

<table>
<thead>
<tr>
<th>Test</th>
<th>(I) cap_esp 35</th>
<th>(J) cap_esp 35</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cap_esp 65</td>
<td>cap_esp 65</td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>Tukey HSD</td>
<td>1</td>
<td>2</td>
<td>−1.530</td>
<td>1.606</td>
<td>.610</td>
<td>−5.428</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>−4.289*</td>
<td>1.542</td>
<td>.021</td>
<td>−8.033</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1.530</td>
<td>1.606</td>
<td>.610</td>
<td>−2.368</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>−2.759</td>
<td>1.672</td>
<td>.236</td>
<td>−6.816</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>4.289*</td>
<td>1.542</td>
<td>.021</td>
<td>0.545</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>2.759</td>
<td>1.671</td>
<td>.236</td>
<td>−1.299</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>4.289*</td>
<td>1.542</td>
<td>.029</td>
<td>−8.199</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>−1.530</td>
<td>1.606</td>
<td>.638</td>
<td>−5.601</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>−2.759</td>
<td>1.672</td>
<td>.267</td>
<td>−6.997</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>4.289*</td>
<td>1.542</td>
<td>.029</td>
<td>0.378</td>
</tr>
<tr>
<td>Scheffe</td>
<td>2</td>
<td>3</td>
<td>2.759</td>
<td>1.672</td>
<td>.267</td>
<td>−1.479</td>
</tr>
</tbody>
</table>

Note. The mean difference is significant at the 0.05 level.

This study was repeated, dividing the categories by the 25/75 percentiles, to obtain at least this significant difference between categories 1 and 3. And finally, if divided into two categories of spatial ability by the 50 percentile, significant differences between both groups may also be found.
It may also be seen from Table 6 that the difference of means between categories 1 and 3 always has negative values in the confidence interval. Because the three divisions by categories (percentages 25/75, 35/65, and 50/50) had no influence on the significance of the results, in our opinion, students with poorer capabilities improve more if they use ILMAGE_SV, since this tool helps to perform mental rotations of 3D models by simulation on the computer. Students with better capabilities have further developed this capability and need less help or none at all.

Thus, we can affirm that students with lower spatial ability improved their visualization ability more than students with greater spatial abilities in a significant way, if their learning achievements were supported by studying with ILMAGE_SV.

In a fourth type of data analysis, the variable “vision” from the exam at the end of the pilot project was compared two by two for all students from the two academic years and between the two groups (E and T). These students initially had the same spatial ability in the DAT-SR test, taking into account that each group had been divided up into three categories of spatial visualization ability: greater ability (1), average ability (2) and lower ability (3) between the 35/65 percentiles. The procedure is similar to the earlier case that also applied the Tukey test (Table 7), which performs multiple comparisons between categories.

Table 7. Multiple comparisons

<table>
<thead>
<tr>
<th></th>
<th>DAT_SR_35_65</th>
<th>DAT_SR_35_65</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(I)</td>
<td>(J)</td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>Tukey</td>
<td>E3</td>
<td>T3</td>
<td>-0.485</td>
<td>0.232</td>
<td>0.297</td>
<td>-1.152</td>
</tr>
<tr>
<td>HSD</td>
<td>E2</td>
<td>T2</td>
<td>-0.0398</td>
<td>0.254</td>
<td>1.000</td>
<td>-0.768</td>
</tr>
<tr>
<td></td>
<td>E1</td>
<td>T1</td>
<td>0.369</td>
<td>0.222</td>
<td>0.558</td>
<td>-0.269</td>
</tr>
</tbody>
</table>

The results in Table 7 indicate that neither the differences between students of greater spatial visualization ability in groups E and T nor between those of lower ability are significant. But it may be said that the results are significantly equal among students of average ability between groups E and T.

Although the results of greater and lesser SVA are not significant, the average values for the vision variable are slightly higher among students with greater abilities in group E as opposed to those in group T. Contrary results are found among students of lower ability; the average values of the variable visualization are greater in group T than in group E (Figure 8). But, as mentioned earlier, the results are not significant.

If, instead of the vision variable, we study the variable “final vision test,” obtained from the results of the student’s end-of-year exam, then we should examine the results in Figure 9 and Table 8. Table 8 confirms that the three comparisons have a level of significance of \( \alpha > 0.95 \), which allows us to affirm that the results are equal for methods T and E among students of greater, equal, and lower spatial visualization ability.
In Spain, students can access university engineering courses with or without previous knowledge of technical drawing, as it is an optional material in pre-university teaching. Thus, in a fifth analysis of the data, we studied whether significant relations existed for the vision variable between students from groups E and T with and without previous knowledge of technical drawing. Variance in the vision variable is not equal among the students of the four subgroups, which makes it impossible to apply the comparison using the Tukey HSD statistic test.

When the variances are unequal, the Welch test is used to compare the differences between the averages. Its results were $F = 2.27$ and $\alpha = 0.085$, which were therefore not significant at a significance level of 0.05. However, when we studied the results at the significance level of 0.1, we found significant differences between the subgroups. Moreover, when the variances are not equal, the Tukey test can be replaced by the Games Howels test. In Table 9, the differences in the vision variable in group T between the students with and without previous knowledge are significant at a level of $\alpha = 0.093$ and at a confidence level of 90%. We can also see that the differences for the vision variable in group E, between students who have and those who have no previous knowledge, stands at $\alpha = 0.941$, so

---

**Table 8. Multiple comparisons**

<table>
<thead>
<tr>
<th>(I)</th>
<th>DAT_SR_35_65</th>
<th>(J)</th>
<th>DAT_SR_35_65</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tukey</td>
<td>E3</td>
<td>T3</td>
<td></td>
<td>.096</td>
<td>.390</td>
<td>1.000</td>
<td>−1.029</td>
<td>1.220</td>
<td></td>
</tr>
<tr>
<td>HSD</td>
<td>E2</td>
<td>T2</td>
<td></td>
<td>−.022</td>
<td>.429</td>
<td>1.000</td>
<td>−1.258</td>
<td>1.214</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E1</td>
<td>T1</td>
<td></td>
<td>.269</td>
<td>.357</td>
<td>.975</td>
<td>−.759</td>
<td>1.298</td>
<td></td>
</tr>
</tbody>
</table>

**Table 9. Multiple comparisons**

<table>
<thead>
<tr>
<th>(I)</th>
<th>ET_DTNT</th>
<th>(J)</th>
<th>ET_DTNT</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>90% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Games-</td>
<td>E_DT</td>
<td>E_NDT</td>
<td></td>
<td>.126</td>
<td>.222</td>
<td>.941</td>
<td>−.391</td>
<td>.642</td>
<td></td>
</tr>
<tr>
<td>Howell</td>
<td>T_DT</td>
<td>E_DT</td>
<td>−.216</td>
<td>.158</td>
<td>.521</td>
<td>.521</td>
<td>−.582</td>
<td>.149</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T_NDT</td>
<td>T_DT</td>
<td>−.323</td>
<td>.232</td>
<td>.509</td>
<td>.219</td>
<td>−.642</td>
<td>.391</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E_NDT</td>
<td>T_DT</td>
<td>−.342</td>
<td>.216</td>
<td>.396</td>
<td>.396</td>
<td>−.847</td>
<td>.163</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T_DT</td>
<td>T_NDT</td>
<td>.198</td>
<td>.275</td>
<td>.890</td>
<td>.434</td>
<td>−.443</td>
<td>.839</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E_DT</td>
<td>T_NDT</td>
<td>−.323</td>
<td>.233</td>
<td>.509</td>
<td>.509</td>
<td>−.866</td>
<td>.219</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E_NDT</td>
<td>T_DT</td>
<td>−.198</td>
<td>.275</td>
<td>.890</td>
<td>.890</td>
<td>−.839</td>
<td>.443</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T_DT</td>
<td>T_NDT</td>
<td>−.540</td>
<td>.228</td>
<td>.093</td>
<td>.093</td>
<td>−1.072</td>
<td>−.008</td>
<td></td>
</tr>
</tbody>
</table>

*Note. The mean difference is significant at a level of 0.1.*
we may say, at a confidence level of 90%, that there are no significant differences between students from group E who have previous knowledge and those who have none, after the pilot project.

Figure 10 and the results in Table 9, indicate that the differences of means in group E, between students who have previous knowledge of technical drawing and those who do not, is much less than the differences in group T. One interpretation of these results is that the use of ILMAGE_SV is of greater help in the learning of spatial visualization for those students who arrive at university with no previous knowledge of technical drawing.

![Figure 10. Average for vision in the two groups E and T for students with and without previous knowledge](image)

**Conclusions**

Our main conclusions are as follows:

- We can affirm that students who used the traditional method as much as those who use the experimental method with the ILMAGE_SV application increased their spatial vision capacity after the pilot project in a significant way.

- We sought to demonstrate that learning among the Group E students improved more than among the Group T students through the variable increase of the MRT test indicated in a significant way that the average increase was equal in both groups, despite the average increase being 1.7 points greater in group E than in group T.

We can also affirm that the results of the evaluation of the vision variable at the end of the pilot project were equal in both groups and coincided with the results of Konukseven (2010). The fact that the general results were equal for both groups, that the use of ILMAGE_SV is asynchronous and is adapted to the student’s speed of learning, and that its 3D viewer adapts itself to the mental rotation of the models means we can recommend the use of this tool for the development of SVA.

- This tool has been demonstrated to be more effective for students whose SVA is lower, because it aims to develop the preliminary concepts of SVA. It is therefore recommended for students with a lower SVA. We should propose exercises of greater difficulty, in order to make it more effective for students with greater levels of SVA.

- If we compare the results of the exam variable of vision, after the pilot project, which students of greater, average, and lower ability obtained in groups E and T, respectively, the results are equal in a significant way. If the results of the “final vision test” variable are compared at the study module, we may affirm that the students of greater, equal, and lower spatial ability from groups E and T obtained very similar results. We have therefore maintained equality between the SVA categories over time.

- The results for sketching among students who used ILMAGE_SV (group E) were no worse than those in group T.
• The use of ILMAGE_SV is suitable for students at the start of their engineering studies at university who have no previous knowledge of technical drawing. Thus, this tool (ILMAGE_SV) is recommended for training prior to the start of the course for those students who come to engineering studies with no previous knowledge of technical drawing.

We should not forget that this tool covers only the fundamental aspects of SVA, without including geometry, sketching, knowledge of technical drawing standards, and CAD, where SVA may also be developed. This may be one of the reasons why the results of different researchers are contradictory.

The use of ILMAGE_SV is viable despite the initial time spent on its development, because: 1) it is used in different groups and over various years; 2) it saves evaluation time because it is automatic; 3) its flexibility means new graphic engineering modules may easily be added. A future line of research could be to determine the number of years before the application is in fact viable.

At present, one line of research is to use ILMAGE_SV in study modules on section views and auxiliary views in which 3D viewers may be used to improve student learning. Another line of research is the adaption of the ILMAGE_SV structure to other knowledge areas where SVA is necessary.

In conclusion, we can state that a recurrent problem is that approximately 20% of engineering students have experienced difficulties with the development of their spatial visualization abilities (Veurink et al., 2009; Sorby, 2009; Duff & Kellis, 2009; Brus & Boyle, 2009; Knott & Kampe, 2009), and that one method by which students can address this problem is to complete preliminary and complementary courses to develop these spatial abilities (Veurink et al., 2008; Sorby, 2009).

References


156


Developing Digital Courseware for a Virtual Nano-Biotechnology Laboratory: A Design-based Research Approach

Hsiu-Ping Yueh¹*, Tzy-Ling Chen², Weijane Lin³ and Horn-Jiunn Sheen⁴

¹Department of Bio-Industry Communication and Development, National Taiwan University, Taiwan // ²Graduate Institute of Bio-Industry Management, National Chung Hsing University, Taiwan // ³Department of Library and Information Science, National Taiwan University, Taiwan // ⁴Graduate Institute of Applied Mechanics, National Taiwan University, Taiwan // yueh@ntu.edu.tw // tlchen@dragon.nchu.edu.tw // vjlin@ntu.edu.tw // sheenh@ntu.edu.tw

*Corresponding author

ABSTRACT

This paper first reviews applications of multimedia in engineering education, especially in laboratory learning. It then illustrates a model and accreditation criteria adopted for developing a specific set of nanotechnology laboratory courseware and reports the design-based research approach used in designing and developing the e-learning material. According to findings of the present study, the courseware developed satisfies the “e-Learning Courseware Quality Checklist version 3.0” in most dimensions. This paper concludes by presenting the researchers’ findings and problems encountered in this ongoing study, and it describes a future study plan that will include the student input regarding these laboratories. This study should contribute to the innovative use of technology in facilitating engineering laboratory learning and in instructional design practice and research.

Keywords

E-learning, e-Learning Courseware Quality Checklist, Instructional design, Virtual nano-biotechnology laboratory, Nanotechnology

Introduction

Nanotechnology, which is the science, engineering and technology conducted at the nanoscale (United States National Nanotechnology Initiative, 2000), is a cutting-edge technology with great potential, and it is widely applied in many fields with inter- and cross-disciplinary approaches (Guggisberg, Formaro, Gyalog, & Burkhart, 2003). Since the invention of the scanning tunneling microscope (Binnig, Rohrer, Gerber, & Weibel, 1982) opened the era of nucleation tools and applications studies, applications and research on the nanometer scale have prospered. However, nanotechnology laboratories are often expensive and of limited usage, so teaching nano-science with such equipment is rarely possible in practice. Considering the needs of today’s nanotechnology engineers to be well educated in all aspects of nanotechnology design and implementation, efforts have been made to facilitate the investigation with the scientific process, a core competency in engineering education usually demonstrated through laboratory skills and performance. The current rapid growth in research to explore sound pedagogical methods in teaching interdisciplinary subjects of nanotechnology aims to improve student learning of relevant laboratory skills (Adams, Rogers, & Leifer, 2004; Sullivan et al., 2008; Swiss Virtual Campus, 2000; Tahan et al., 2005; Yueh & Sheen, 2009).

Along with the development of information and communication technology (ICT), which includes audio/ video instruction, computer-assisted instruction, intelligent tutoring systems, internet/ web-based instruction, and current Web 2.0 technologies, a substantial body of research has focused on the search for effective applications of technology in support of teaching and learning in various fields. The use of multimedia resources allows an integrative presentation of messages. In recent practice in engineering education, multimedia resources have become essential in teaching complicated engineering concepts and in demonstrating skills more vividly (Fernandez et al., 2011; Walker, Stremler, & Brophy, 2008). As Mayer (2001) argued, multimedia, when used in support of learning, can help learners not only to recognize and retain the presented materials but also to construct a coherent mental representation of what they have learned.

The virtual laboratory has been widely applied in science and engineering education in different formats. Through the use of different tools and platforms, laboratory exercises and experiences can be conveyed by wide-area network bandwidth on the Internet and satellites that allow remote access (Forbus et al., 1999; Wyatt, Arduino, & Macari,
learning courseware, was originally developed as part of a national project funded by the government of Taiwan. It consists of five dimensions (Table 1) that are intended to encompass the range of functions involved in supporting a student-centered learning experiences develop upon is considered the most important of all in the selection of the framework for quality assurance (Inglis, 2008). The eLCQC 3.0, which is designed for assessing the key quality factors of all aspects of e-learning courseware, it is recognized as a fit and adequate framework for this study. The original eLCQC was first designed web-based modules that simulate laboratory assignments, and students interact with these learning modules virtually (Koretsky, Amatore, Barnes, & Kimura, 2008; Watkins, Hall, Chandrashekharra, & Baker, 2004). Other studies have combined remote and local laboratories so that students can manipulate real laboratories devices by performing simulated activities with web interfaces (Ogot, Elliot, & Gulmac, 2003; Sivakumar, Robertson, Artimy, & Aslam, 2005; Spanis & Atti, 2005). Still others have advocated the importance of providing students with flexible access to information on the underlying concepts behind experiments and laboratory work, and intelligent tutoring systems have been developed to facilitate learning in the virtual laboratory environment (Yueh & Sheen, 2009).

With virtual laboratories being applied so widely in engineering education, the effectiveness of these practices has been the focus of many studies (Hurley & Lee, 2005; Koretsky, Amatore, Barnes & Kimura, 2008; Ogot, Elliot, & Gulmac, 2003; Sivakumar, Robertson, Artimy, & Aslam, 2005; Spanis & Atti, 2005). As Reilly (2008) noted, virtual laboratories, whether accessed locally or remotely, have the potential to provide greatly enhanced and more effective deep-learning experiences. Although such innovations look promising and enhance student learning, many educators and researchers have encountered problems with judging the quality of multimedia design and its effect on learning improvement. It is also our belief that multimedia virtual laboratories, which are particularly suitable for learning about nanotechnology, require a quality assurance framework for their design, with emphasis placed on student-centered learning experiences. Thus, this paper reports on one cycle of designed-based research (DBR) in which e-learning courseware for a nanotechnology virtual laboratory was developed and the e-Learning Courseware Quality Checklist version 3.0 (eLCQC 3.0) was used for quality assurance. The description of the emergence of a new virtual nano-bio engineering laboratory and the preliminary iterative analysis of the quality of its design could be interesting to the wider research community, and that information could be useful in setting up further studies and in the development of other courseware.

The eLCQC quality assurance framework

Quality assurance has been gaining attention from both educators, who adopt multimedia or e-learning approaches, and practitioners, who design and produce courseware for virtual laboratories. Pawlowski (2007) indicated that although many e-learning producers and users are aware of the importance of quality issues, they lack appropriate methods or tools for measuring the quality of their e-learning courseware. Institutions in several countries have begun efforts to promote e-learning quality. For example, the American Society of Training and Development (ASTD) has established an E-Learning Courseware Certification (ECC) program that employs a tool based on quality criteria and provides quality assurance services to organizations for their e-learning courseware accreditation (Sanders, 2001). The European Foundation for Quality in E-Learning (EFQUEL, 2007), having developed the UNIQuE certification system, issues an e-learning quality label for ICTs. FuturED, a non-governmental organization, has cooperated with the federal government in Canada to build an e-learning courseware and service quality assurance program, the Open eQuality Learning Standards (OeQLS), and the eQcheck certification, to protect e-learning consumers. Taiwan, similarly, has established the E-Learning Quality Certification Center (ELQCC) (Sung, Chang, & Yu, 2011).

The standards or criteria developed by those institutes, which all intend to promote the quality of e-learning, focus on different details of the features they check. More important is the fact that although they are used mostly for evaluating product quality, they could also be used as a framework for the design and development of e-learning materials or courseware. With this application in mind, this study adopted eLCQC 3.0 (Taiwan E-Learning Quality Certification Center, 2007) developed by the ELQCC in Taiwan as the quality assurance framework to scaffold the design and development of e-learning courseware for a nano-biotechnology laboratory. Although many quality frameworks are also available for examining e-learning courseware quality, the aspect of a specific context that learner experiences develop upon is considered the most important of all in the selection of the framework for quality assurance (Inglis, 2008). The eLCQC 3.0, which is designed for assessing the key quality factors of all aspects of e-learning courseware, was originally developed as part of a national project funded by the government of Taiwan. It consists of five dimensions (Table 1) that are intended to encompass the range of functions involved in supporting a good practice of digital courseware. In use, it is also intended to be contextualized to the setting of a courseware producer. Taking into account the value of the contributions of eLCQC 3.0 to assessment and improvement of e-courseware quality, it is recognized as a fit and adequate framework for this study. The original eLCQC was first...
issued in 2005, and the latest version, 3.0, was published in 2007. Built upon the perspectives of learner, instructor, courseware developer, and e-learning project manager, eLCQC 3.0 consists of five dimensions: content, navigation and tracking, instructional design, instructional media, and creativity. Each dimension consists of 3 to 6 standards, for a total of 19 standards. Additionally, except for those standards under the dimension of creativity, each standard has 2 to 3 criteria for the evaluation of e-learning courseware. Table 1 summarizes the specifications of eLCQC 3.0.

Table 1. Specification of eLCQC 3.0

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Standard</th>
<th>Description of Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Content</td>
<td>1.1 Accuracy</td>
<td>The e-learning courseware must provide accurate learning content with appropriate</td>
</tr>
<tr>
<td></td>
<td>1.2 Organization &amp; Completeness</td>
<td>organization and concise presentation so that learners are able to achieve successful</td>
</tr>
<tr>
<td>2. Navigation and Tracking</td>
<td>2.1 Navigation</td>
<td>The e-learning courseware must provide relevant mechanisms for guiding learners</td>
</tr>
<tr>
<td></td>
<td>2.2 Orientation &amp; Help</td>
<td>through the course smoothly so that learners can effectively control and monitor the</td>
</tr>
<tr>
<td></td>
<td>2.3 Learning Tracking</td>
<td>pace and progress of their personal learning.</td>
</tr>
<tr>
<td>3. Instructional Design</td>
<td>3.1 Instructional Objectives</td>
<td>The e-learning courseware must be consistent in the instructional design, thereby</td>
</tr>
<tr>
<td></td>
<td>3.2 Instructional Methods</td>
<td>providing learners with concise learning goals, adequate presentation of instruction,</td>
</tr>
<tr>
<td></td>
<td>3.3 Practice &amp; Formative</td>
<td>and appropriate learning strategies. The learners can benefit from this design with</td>
</tr>
<tr>
<td></td>
<td>Evaluation</td>
<td>improved understanding of the contents, good learning interactivity, and appropriate</td>
</tr>
<tr>
<td></td>
<td>3.4 Summative Evaluation</td>
<td>evaluation and feedback.</td>
</tr>
<tr>
<td></td>
<td>3.5 Facilitation Strategies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.6 Congruence</td>
<td></td>
</tr>
<tr>
<td>4. Instructional Media</td>
<td>4.1 Media Design &amp; Use</td>
<td>Effective utilization of the instructional media, aesthetically pleasing interface</td>
</tr>
<tr>
<td></td>
<td>4.2 Interface Design</td>
<td>design, and instructional media design and production are essential to the enhancement</td>
</tr>
<tr>
<td></td>
<td>4.3 Media Elements</td>
<td>of learning and comprehension.</td>
</tr>
<tr>
<td>5. Creativity</td>
<td>5.1 Content</td>
<td>To promote the “Creativity” of presentation in the e-learning courseware, creative</td>
</tr>
<tr>
<td></td>
<td>5.2 Navigation &amp; Tracking</td>
<td>examples should be highlighted, ranging from the instructional contents and learning</td>
</tr>
<tr>
<td></td>
<td>5.3 Instructional Design</td>
<td>guides to instructional design, instructional media, and other sections.</td>
</tr>
<tr>
<td></td>
<td>5.4 Instructional Media</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.5 Other</td>
<td></td>
</tr>
</tbody>
</table>

To determine the reliability of eLCQC 3.0, Sung, Chang and Yu (2011) analyzed its item difficulty, item discrimination, and generalizability. They confirmed that the eLCQC can effectively discriminate differences in quality in the courseware. As a result, eLCQC 3.0 was used in the present study, and it was coupled with DBR to refine the development process of the courseware of one unit of a virtual nano-biotechnology laboratory and to evaluate the effects of this effort in improving e-learning quality.

**Design-based research**

For research involving the development of teaching and learning innovations or the systematic design and study of instructional strategies and tools, design-based research (DBR) is commonly employed as a research method framework (Barab & Squire, 2004; Brown, 1992; Collins, 1992; Dede, 2004). DBR is an emerging paradigm for the study of learning in context, especially in technology-enhanced learning environments. The use of DBR enables educational researchers to grasp problems through its iterative development and emphasis on authentic contexts, and it allows a better understanding of the goals and implications of the research (Joseph, 2004). DBR methods focus on
designing and exploring the whole range of designed innovations to optimize the design as much as possible. The Design-Based Research Collective (2003, p5) proposes that good DBR exhibits the following five characteristics:

“(1) the central goals of designing learning environments and developing theories or “proto-theories” of learning are intertwined. (2) development and research take place through continuous cycles of design, enactment, analysis, and redesign (Collins, 1992). (3) research on designs must lead to sharable theories that help communicate relevant implications to practitioners and other educational designers. (4) research must account for how designs function in authentic settings. (5) the development of such accounts relies on methods that can document and connect processes of enactment to outcomes of interest. Therefore, DBR provides numerous opportunities for the exchange of expertise across disciplinary boundaries” (p. 5).

DBR, a pragmatic research approach, has gradually received extensive attention in various fields of engineering education, including the nanotechnology field (Middleton, Gorard, Taylor, & Bannan-Ritland, 2008). Since DBR aims to make sense of the complexity inherent in the authentic context, it is considered appropriate for research on applying knowledge to consequential problems and reflecting on that endeavor (Rick & Guzdial, 2006). In the present study, researchers, who were also instructional designers and content experts, applied DBR to the courseware development environment for setting up a virtual nano-biotechnology laboratory. In order to explore how a quality assurance framework, such as eLCQC, can be used in developing, enacting, and sustaining a better learner experience, DBR was used to inform and improve the development of e-learning courseware for nanotechnology learning.

Method

The context of the virtual nano-biotechnology laboratory

The unprecedented growth of online educational resources led Kilmeck, McLennan, Brophy, Adams and Lundstorm (2008) to affirm that in nanotechnology, as in other disciplines, the web is expanding our concept of classroom instruction, changing what is learned and how it is learned. This affects how students learn in class, and new types of nanotechnology learning activities and resources can help meet the educational needs of all groups of learners by improving their access to learning tools typically unavailable in the classroom. As part of an effort to enhance both laboratory and interdisciplinary learning, the “Nano-bio Technology Laboratory Corridor (NBTLC)” was established as the first group of laboratories to combine nano-manufacturing and biomedical science manufacturing processes in higher education in Taiwan (Sheen, 2007). The basic idea of NBTLC is a metaphor of open laboratories; it supports both independent laboratory work and integrated interdisciplinary learning laboratories. To further advance the idea and practice of NBTLC, a “Virtual Nano-bio Technology Laboratory Corridor (V-NBTLC)” was constructed as a supplemental learning space for the NBTLC. It represents a virtual portal to NBTLC with integrated multimedia design, and it offers digitized learning materials such as laboratory handbooks, conceptual lectures, instrument illustrations, step-by-step skill and procedure demonstrations, and some virtual and remote-controlled experiments of the four laboratories (Yueh & Sheen, 2009). Under the larger context of V-NBTLC, courseware on Atomic Force Microscopy (AFM) was developed in this study as a case that exemplifies the DBR approach of using eLCQC 3.0 to scaffold the design process and assure the quality of the learner experience in the virtual nano-biotechnology laboratory.

Designing the learner experience of AFM courseware

In order to facilitate the AFM courseware production process, the development team applied the “e-learning project development model” (Yueh, 2005), which includes the following phases: management and feedback, course analysis, material production, presentation and interaction, and learning and evaluation. The development team consisted of two expert groups, one group with the expertise in instructional design and the other group with expertise in the subject matter of nanotechnology. These two groups worked closely to confirm the purpose, structure, sequence and presentation of the courseware. Two researchers of the study were involved and responsible for coordination in each team, respectively. Figure 1 illustrates and emphasizes the actual tasks, process, and communication between the two groups.
Furthermore, researchers adopted eLCQC 3.0 to scaffold the design process and to evaluate the quality of AFM courseware mounted on V-NBTC. DBR enables researchers to explore possibilities for novel learning and teaching environments; develop contextualized theories of learning and teaching; construct cumulative design knowledge; and increase the human capacity for innovation. Therefore, it was adopted as the research method framework for this study.

Throughout the process of courseware development, it is necessary to conduct progressive and iterative review and revision of the instructional design plan and output according to DBR (Design-Based Research Collective, 2003). To achieve this goal, researchers employed designer self-evaluations and content expert evaluations, as illustrated in Figure 2. In the course analysis phase, instructional designers first conducted a needs analysis by collecting information from subject-matter experts. Following the needs analysis was the first designer self-evaluation. Then the second designer self-evaluation and the first expert evaluation were performed, after which multimedia producers began the material production module of courseware development. When the courseware was officially implemented, two more designer self-evaluations and expert evaluations were performed before and during the presentation and interaction phases. In the learning and evaluation phases, one more expert evaluation and the learner/user evaluation were performed. Every action in the production of the AFM courseware, including analysis, design, development, implementation, and evaluation, passed through a review based on the eLCQC framework before the next step could begin. The three basic components of enactment, review, and revision formed a dynamic courseware developmental structure in this case.

This study reports the development and research of one of the cycles of the abovementioned dynamic structure. The researchers consisted of professional instructional designers and content experts of nanotechnology, who collaborated with a multimedia production team and a group of subject-matter experts through three phases of initial needs analysis, design, and development. The first two designer self-evaluations and first expert evaluation were
conducted to serve as the progressive review and to collect comments, problems, and suggestions on refining the design plan in order to continue the design and development. Informal interviews were conducted, using eLCQC 3.0 as an analytical framework. The third designer self-evaluation and second expert evaluation integrated eLCQC 3.0 as a quality audit tool. While it was a review of the DBR process, the preliminary evaluation of the AFM courseware prototype was also considered. For data collection, the eLCQC 3.0 checklist was completed by five reviewers. The reviewers were instructional designers and experts on multimedia and content.

**Figure 2.** The DBR process of designing the learner experience of AFM courseware

### Results

#### Production of the AFM laboratory courseware

The AFM courseware based on the framework of eLCQC 3.0 was first constructed from the needs identified from interviews with content experts. These included the instructional objectives of the course unit and purposes of developing this e-learning courseware. The target audience, instructional objectives, content to be covered thoroughly, in terms of broadness and depth of learning, and the content representation and organization strategies were thus defined and reviewed before the actual design stage began. Considering the smooth navigation functionality that e-learning must provide, the AFM courseware was carefully laid out with proper guidance, orientation, tracking, and interface designs (see Figure 3). To effectively deliver the learning contents to students at a distance, the e-learning courseware of AFM used diverse instructional methods with clear media presentation and multimedia design strategies (see Figure 4). To facilitate and ensure effective learning, the practices and the formative and summative evaluations were incorporated into the interactive and feedback designs (see Figure 5). Furthermore, the designers, working with content experts, carefully checked that instructional objectives, content design, and evaluations were consistent, and they also made sure the AFM courseware included creative design that would motivate students and lead to a successful learner experience in the virtual learning environment.

The courseware development process of collaborative scaffolding encouraged interactions among instructional designers, content experts, and the media production team; facilitated joint problem solving; led to richer construction or revision of instructional planning; and took into account different and emerging roles, joint goals, and actions. It also facilitated the fusion of multiple interpretations. Under the DBR process, the cycles of enactment (such as course or need analysis, learning design, and material production), review, and revision were conducted. The outcome at the end of each cycle was an explanatory framework based on eLCQC 3.0 that specified expectations as the focus of inquiry during the next cycle. The generic courseware architecture of AFM would also be re-used to produce a new prototype courseware for V-NB TLC.
Preliminary evaluation of the AFM laboratory courseware

At completion of the prototype of AFM courseware, researchers conducted a preliminary evaluation, similarly based on the eLCQC 3.0 framework, in the form of a survey of five reviewers consisting of instructional designers and experts on multimedia and content. The survey checklist, developed according to eLCQC 3.0, included five dimensions, with 19 standards specified. Under each standard, 2 to 3 criteria were also included. For each criterion under one standard, the reviewer was asked to award 0-3 points, depending how many criteria the AFM courseware met. For example, the highest possible score of the standard “1.1 accuracy” was 9, and the lowest was 0. The score range of each standard and the preliminary evaluation results are shown in Table 2. In general, these experts awarded this courseware middle to high evaluations (a total score of 73.25 out of 106). This implies that use of the eLCQC framework, applied along with the DBR approach, enabled the AFM courseware development team to construct a common quality assurance and inspection scheme to help achieve quality in learning design.
Table 2. Preliminary evaluation results of designer and expert review

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Standard</th>
<th>Score range</th>
<th>Average score</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Content</td>
<td>1.1: Accuracy</td>
<td>0-9</td>
<td>9.00</td>
<td>17.25</td>
</tr>
<tr>
<td></td>
<td>1.2: Organization &amp; Completeness</td>
<td>0-9</td>
<td>8.25</td>
<td></td>
</tr>
<tr>
<td>II. Navigation and Tracking</td>
<td>2.1: Navigation</td>
<td>0-9</td>
<td>6.75</td>
<td>9.75</td>
</tr>
<tr>
<td></td>
<td>2.2: Orientation &amp; Help</td>
<td>0-6</td>
<td>3.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.3: Learning Tracking</td>
<td>0-6</td>
<td>0.00*</td>
<td></td>
</tr>
<tr>
<td>III. Instructional Design</td>
<td>3.1: Instructional Objectives</td>
<td>0-6</td>
<td>6.00</td>
<td>32.75</td>
</tr>
<tr>
<td></td>
<td>3.2: Instructional Methods</td>
<td>0-9</td>
<td>4.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.3: Practice and Formative Evaluation</td>
<td>0-9</td>
<td>6.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.4: Summative Evaluation</td>
<td>0-6</td>
<td>4.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.5: Facilitation Strategies</td>
<td>0-4</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.6: Congruence</td>
<td>0-9</td>
<td>9.00</td>
<td></td>
</tr>
<tr>
<td>IV. Instructional Media</td>
<td>4.1: Media Design &amp; Use</td>
<td>0-6</td>
<td>6.00</td>
<td>13.00</td>
</tr>
<tr>
<td></td>
<td>4.2: Interface Design</td>
<td>0-4</td>
<td>3.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.3: Media Elements</td>
<td>0-4</td>
<td>3.50</td>
<td></td>
</tr>
<tr>
<td>V. Creativity</td>
<td>5.1: Content</td>
<td>0-2</td>
<td>0.25</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>5.2: Navigation &amp; Tracking</td>
<td>0-2</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.3: Instructional Design</td>
<td>0-2</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.4: Instructional Media</td>
<td>0-2</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.5: Other</td>
<td>0-2</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

Note. *The prototype did not include the tracking function that would be set up in the platform.

Discussion and conclusion

Findings and discussion

Despite the results indicating acceptable quality, not all of the reviewed dimensions and standards were of good quality. In particular, the creativity in instructional design was found to be the most insufficient. While the standards under the creativity dimension had no specific criteria, it seemed to the courseware development team that the guidance was comparatively limited. Thus, to accomplish creative learning design in all aspects, namely, content, navigation and media design, was quite challenging. In other words, it was necessary to provide substantial assistance embedded within a quality assurance framework to inspire the designers towards more creative design.

One more issue should be raised in our preliminary evaluation. All of the input or feedback considered and analyzed in the present study were from the perspectives of designer practitioners or experts. If the aim is to develop more inclusive virtual learning of high quality, it is necessary to include more diverse voices, such as those of experienced senior students, laboratory assistants, or professional engineers in nanotechnology, at the beginning of courseware production. To achieve this attempt, the potential learners or experienced learners of previously developed courseware included in V-NBTLC will be identified and recruited to serve in the group of reviewers for quality checking. By doing so, a participatory approach of involving learners or content users in the quality assurance mechanism of courseware development can be promoted (Jung, 2011). In addition, reliable research data on the review and evaluation sessions of DBR have yet to be obtained. According to Hevner, March, Park and Ram (2004), five classes of evaluation methods are available for research consideration, including analytical, case study, experimental, field study, and simulation methods. To maintain the collaborative scaffolding of courseware development, using a mixed method to collect both qualitative and quantitative data under DBR may result in better outcomes.

Implications and conclusion

This study reports a new courseware development approach for the establishment of a virtual nano-biotechnology laboratory that considers the e-learning material quality assurance principles in combination with DBR. The overall results of this study should provide systematic, empirical evidence on the quality of AFM, the e-learning courseware.
evaluated. According to this study, the reflection of courseware development and the preliminary evaluation results provided by a group of experienced multimedia and e-learning specialists support that the courseware developed satisfied the standards of eLCQC in most dimensions. However, several suggestions to improve the design were made, such as providing more examples of creative learning design practices in either nanotechnology or other fields, considering the use of alternative pedagogies, adding more browsing or navigation guidelines, and providing a course map of the tutorial. As this is an ongoing study, researchers will continue the process of implementation and evaluation based on the same dynamic courseware developmental structure, which consists of enactment, review, and revision in compliance with DBR. In the future, the focus group method will be adopted for brainstorming and systematic review discussions. To be included in this focus group of quality assurance are content experts, instructional designers, multimedia producers, and courseware users or learners. In addition, feedback on the quality framework of eLCQC 3.0 will be collected and in turn will form an integral part of the process of validation of the framework. Additionally, researchers will further implement this V-NBTL in a real educational context on a large scale, and student feedback will be collected to assure the quality of this courseware for future improvement as well. It is expected that through this DBR approach, this study will not only produce quality e-learning courseware but also provide more thorough understanding of design principles of, and theoretical claims about, using technology in facilitating learning in the engineering laboratory, as well as instructional design practice and research.

Acknowledgements

The authors would like to thank National Science Council (NSC 100-2120-S-002-001-NM; NSC 100-2628-S-002-001-MY3) and National Taiwan University for their grants in support of this study.

References


Using Instructional Pervasive Game for School Children’s Cultural Learning

Cheng-Ping Chen, Ju-Ling Shih* and Yi-Chun Ma

Department of Information and Learning Technology, National University of Tainan, 33, Sec. 2, Shu-Lin St., Tainan, Taiwan // chenjp0820@gmail.com // juling450@gmail.com // cherrisma1027@gmail.com

*Corresponding author

(Submitted September 27, 2012; Revised March 3, 2013; Accepted June 17, 2013)

ABSTRACT

In the past ten years, mobile learning (m-learning) has created a new learning environment that enables learners, through active learning aids. Instructional pervasive gaming (IPG) seems to be an innovative way introduced to enhance m-learning. This study employed a theoretical IPG model to construct a cultural-based pervasive game. Individual and collaborative learning methods, learning effectiveness, attitude toward mobile devices, and satisfaction of the gameplay were explored. Forty-three fifth grade students were selected as the subjects for the experiment. Two religious historical sites were the experimental locations. Open Data Kit (ODK) form was used as the interface for gathering information from the mobile device. The results indicate that IPG is helpful for learning local culture, and students’ attitude toward mobile devices affects their learning achievement. Also, the more agreeable students with the game, the more learning achievement can be obtained.

Keywords
Instructional pervasive game, Mobile learning, Cultural learning, School children

Introduction

In the past thirty years, the instructional implementation of computer technologies have gone through the phases of “Computer-assisted Instruction (CAI),” “web-based instruction,” and “mobile learning (m-learning),” just to name a few. The education has been transformed from the teacher-directed and one-size-for-all learning materials to stand-alone, self-paced, and well-programmed instructional materials. With the assistance of digital technologies, Internet, and mobile devices, the educational implementation further break the limitation of time and space that extends the classroom to the environment around us. Learning has been beyond the physical constrain, students experience “ubiquitous learning (u-learning).” In comparison with CAI and web-based instruction, more and more research evidences have shown that m-learning or u-learning is especially helpful in facilitating learners constructing knowledge with regard to the situational information in social, cultural, and physical contexts (Laine, Sedano, Joy, & Sutinen, 2010). Mikic, Anido, Valero, and Picos (2007) also said that mobile learning has created a new learning environment that enables learners, through active learning aids.

With the help of advanced mobile devices and well-designed instructional strategies, m-learning is able to bring students out of the classroom to interact with the environment. Many local studies on m-learning repeated the concerns of establishing a learning system emphasizing on the use of mobile devices (such as PDA) with the established system, (e.g., Liang 2005; Lin, 2006) but seemed to fall short of either applying or concluding a set of effective instructional strategies with in-depth explanation of theories and practices.

For this reason, the instructional application of pervasive game seems to be an innovative way introduced to enhance m-learning (Laine, 2010). We defined the term instructional pervasive gaming (IPG) as all pervasive game involving instructional functions. Montola (2011) explained that a pervasive gaming environment is “an extension of m-learning with an emphasis on the roles of an intelligent environment and of the context”. With the aid of GPS, Chen and Shih (2011) have initiated an experimental study which employed a typical pervasive-game type activity called “geocaching” in teaching high school students geographical coordinates and map reading. The results indicated that GPS-aided geocaching helped students learn to read maps and recognize the geographical coordinates in a more effective way, and the GPS-aided geocaching group was generally acquired a better attitude toward using technology.

Such results are in coherence with some other related studies (Chavez, Schneider & Powell, 2009; Munro-Stasiuk, 2006). However, Gentes, Guyot-Mbodji and Demeure (2010) pointed out, most game strategies do not take into consideration the anthropological data pertaining to the specific context of use. They indicated that maps and
trajectories are one aspect of treasure hunts but companionship and discovery are at the core of the pleasure of pervasive game. These comments have inspired the further exploration of applying pervasive game in the m-learning with the interest to place more cultural elements into the gameplay.

Therefore, more cultural factors are included in this study. Moreover, a theoretical model of instructional pervasive gaming (IPG) is employed, in which game rules, game mechanics, game strategies, and social interaction are defined (Shih & Chen, 2012). Their model is employed for designing the local cultural-oriented IPG for the study. The objective of this study is to understand the learning effects and learning attitude on IPG. The differences between individual learning mode and collaborative learning mode are also investigated to help identify the effects of the innovative instructional strategy. At last, this study also explored the student’s attitude toward mobile devices as well as the degree of satisfaction to the game.

**Literature review**

**Mobile learning and mobile game**

Laine, Sedano, Joy and Sutinen (2010) defined m-learning as a form of informal learning where the learner traverses a physical context carrying a personal mobile device which provides learning materials and activities. Mikic, Anido, Valero and Picos (2007) believed that mobile learning has created a new learning environment that enables learners, through action learning aids, teaching materials, teachers and other learners, to learn at anytime and anywhere. The advantage of employing mobile technology in learning is not the mobile technology itself, but the unique attributes of the mobile devices that provide an environment which allow apply the innovative learning theory and learning strategies (Su, 2005). Walther (2005) recognized that a mobile game is a game that takes changing relative or absolute position/location into account in the game rules. This excludes games for which mobile devices merely provide a delivery channel where key features of mobility are not relevant to the game mechanics.

One instance of instructional system combining mobile learning and mobile game was designed by Schwabe and Goth (2005). Their instruction used college students as learners. Students were divided into groups to play the game. They obtained information from the website regarding the mission and geographical coordinates of the locations. They then used GPS to complete a task and uploaded the result back to the workstation before they could get the location information of the next mission. The winner was who completed the task first.

**Pervasive game**

According to Benford and Magerkurth (2005), pervasive games (PG) extend the gaming experience out into the real world, players with mobile computing devices move through the world. Hinske, Lampe, Magerkurth, and Röcker (2007) further explained that Pervasive Game is to employ pervasive and Mobile Computing technology in order to support, augment, and realize the game itself. Another important notation of PG has been made by Montola (2011). He argued that PG is a game that “expands the magic circle of play socially, spatially or temporally.” He further explained that magic circle of play is a social and cultural contract that separates ordinary life from play, communicating a way of understanding events that happen within the circle. Pervasive games differ from the usual games in that they break the lines of the magic circle. Pervasive games often blur the boundary of game and real life.

Walther (2005) gave a constructional framework that he called “four axes of PG.” He illustrated the four axes as follows: (1) Distribution, refers to the network which can distribute the gaming information; (2) Mobility, refers to computing mobility; (3) Persistence, refers to total availability all the time; and (4) Transmediality, refers to a media circle that multi-link the world of virtual social networks. From the educational point of view, Laine, Sedano, Joy, and Sutinen (2010) illustrated a “technology integration model” for game-based pervasive learning systems and stated that an IPG needs to meet the following requirement: (1) Pedagogical requirements, include user profile, interaction and collaboration, ownership, authenticity and relevance, and support and assessment; (2) Game design requirements, include resources—financial and human, cultural issues, technical issues, environmental issues, social issues, and temporal issues; and (3) Context requirements, include context-awareness, dynamics, interaction, and content.
As stated earlier, Gentes, Guyot-Mbodji and Demeure (2010) raised an important issue that local culture recognition should be one of the key goals for PG, especially IPG’s. They suggested that for a cultural-rich PG, three key features should be included, they are: (1) Collaborative contents: The contents should be designed by people with qualified knowledge or actually living on the premises; (2) Team exploration: The gameplay either relies on solitary errands or on collective sharing, strategies can be more “group oriented”; (3) Cultural narrative: Unlike traditional computer games which always open with a virtual scenario, PG always sufficiently describes the culture of the city therefore game players are able to have a clear idea of the cultural spots.

Although there have been numerous mixed-reality game prototypes that have been built and studied since the idea of PG developed (e.g., Montola, 2011), a comprehensive model and guidelines for developing instructional PG are still on the way. After reviewing above valuable articles, we had been able to draft a rough picture of an instructional PG. It includes a developmental foundation model (four axes and three key units), a technology integration model which emphasized the equal importance of pedagogical requirement, context requirement, and game design requirement for PG, and finally, an urban-cultural model which depicted the key elements for culture recognition. However, for building an effective and immersive IPG, these articles fall short of paying attention on the players. They are the leading character of PG’s.

Therefore, Shih & Chen (2012) proposed a conceptual model for the design and application of IPG (See Figure 1). It includes six factors that encompass the IPG learning process.

![Figure 1. Instructional Pervasive Game Learning Model](image)

The first two factors are the user and the environment. These two factors are the tangible factors in the model. The user is the student who is the central learning subject. Students play the main role in the learning process in IPG. The learning environment refers to the physical environment that is not confined in the classroom but the whole living environment where the students can physically reach and experience. In our case, Taiwan is the main location where learning happens.

The third and the fourth factors are the information and the context. The information refers to the teaching materials that are selected, designed, and organized by content experts (in most cases are the teachers). Teaching materials are taught both in class, presented on the mobile devices, and referred to in the physical environment explored by the students. It guides the students to extend their learning spaces to the environment, forms a bigger space which is the learning context. As shown in the figure 1, the information axis leads the students to “spin” around the learning space that forms the surface of context. The learning context is the abstract and conceptual part of the learning which acts as the horizon of the learning ground that is expanded into a surface. It includes both the mechanical data and the anthropological data that are the experienced, collected, and formed in the exploration process. The anthropological data are those qualitative raw materials collected by the students which include photos, interviews, field notes, and so forth. In other words, these two factors play the role of content and context.
The fifth and sixth factors are the mobility and gamist. Mobility refers to user’s ability to move around in the learning environment. Other than sitting in the classroom, users not only have the freedom but also the capability to travel and explore with learning content embedded in the mobile devices. As the mobility is the horizontal axis that extends users’ learning experience to outside of confined classrooms, gamist is the drive that enables the knowledge surface to expand. Gamist, the game rules that lead to cooperation or competition, is the driving force that creates the challenge and adventure scenario. With a metaphorical description, if mobility is the vehicle that takes the users to places, gamist is the fuel that provides the energy and motivation to surf. The two factors create the surface consist of motion and motivation.

The six factors, thereafter, altogether form the learning sphere. The whole sphere can be regarded as the knowledge body that is and to be gained by the users.

**Learning local history through inquiry-based approach**

According to Collingwood (1994), history is a kind of science that involves inquiry into the past. Providing students with opportunities to think and act like historians in a rewarding pathway to the past. Corey (2010) argued that both urban and environmental histories in particular lend themselves to hands-on approaches that link the classroom with tangible examples of historical events and trends through pedagogy of place. However, Axtell (2001) contends that the biggest mistake history teachers can make is to not engage the curiosity and imagination of students with experiences that initially drew us into history. This mistake may have severely hampered our history education for a long period of time.

Fortunately, since the early 1990s, active or inquiry-based education has received a great deal of attention. This active learning method opens a new way for social science learning. National Council for the Social Studies (1994) believed that students should “systematically employ processes of critical historical inquiry to reconstruct and reinterpret the past, such as using a variety of sources and checking their credibility, validating and weighing evidence for claims, and searching for causality.” However, Molebash (2004) found that unguided online historical inquiry does not guarantee meaningful learning. Appropriate guidance is necessary for a historical inquiry. Manlove, Lazonder and de Jong (2009) asserted that supporting tool incorporated goal-lists, hints, prompts, cues, and templates reinforce the cognitive regulation skills of students during a fluid dynamic task. In our experiment, playing pervasive game via a smart phone is our “supporting tool” for reinforcing student’s cognitive skills.

**Attitude toward mobile devices**

In 1984, Loyd developed the first version of Computer Attitude Scale (CAS). It was divided into three sub-scales: computer anxiety, computer fondness, and computer confidence. Later Loyd and his research team added the fourth sub-scale: computer usefulness. Although there were many other measuring tools valid for evaluating the computer attitude, Loyd’s CAS and its four sub-scales have been widely used for years.

In the information era, we may consider all digital-related technologies as computer technology. Mobile technology, certainly, is one of the kinds among computer technology. We employed Loyd’s CAS as the instrument for measuring student’s attitude toward mobile devices. The four sub-scales were modified to meet our research circumstance. The four modified sub-scales are: mobile device anxiety, mobile device fondness, mobile device confidence, and mobile device usefulness.

**Research methods**

**Instructional pervasive game design**

This study designed a pervasive game based on the religions of Taiwanese local culture. The design of this game matched the six-factors in the IPG theoretical model. Fifth-grade students were invited as the participants (IPG factor: users). The Ming-Zhi Temple and the Zhang-Xing Presbyterian church in Tainan were chosen to be the experimental locations (IPG factor: environment) to fit into the students’ formal lesson on local culture.
investigations. We cooperated with school teachers, local residents and the native educators to write the pre-instructional and formal teaching materials (IPG factors: information) of Ming-Zhi Temple and the Zhang-Xing Presbyterian church (IPG factors: context).

The IPG system used the Open Data Kit (ODK) as the interface. According to Anokwa, Hartung, Brunette, Borriello and Lerer (2009), ODK is a suite of tools that enables users to collect their own rich data. ODK is designed to let users own, visualize, and share data without the difficulties of setting up and maintaining servers. The tools are easy to use, deploy, and scale. Hartung, Anokwa, Brunette, Lerer, Tseng, and Borriello (2010) asserted that ODK is prevailing over many similar data collection and dissemination systems on mobile devices. They compared ODK with CyberTracker, Pendragon Forms, EpiSurveyor, and other systems concluded that ODK makes the composition easier and more deployable for non-programmers due to its structural design of modular components.

Learning contents were placed in the ODK Form with the Google App Engine (GAE). The ODK Form was uploaded to the smart phones first. The ODK Collect then gathered all the files and information from students who finished the learning process. All data were uploaded and analyzed by the ODK Aggregate server provided by ODK. A sample screen display is shown in figure 2.

![Selecting Submitted Data](image)

Figure 2. ODK Aggregate system screen display

The information collected by the students was transmitted to the cloud server for the researchers to analyze. Wireless Internet signals were available within the observation range. All data could be uploaded through the wireless network.

In the learning context, three ways of qualitative data were required to be collected by the students in the game (IPG factors: gamist) which include photographs, interviews, and observations. Students were required to use the smart phones with ODK (IPG factors: mobility) to conduct the game. Based on the speed of finishing the tasks and the learning outcome achieved by each student team, the cooperativeness and competitiveness were added into the game. The experimental design is shown in figure 3.

We spent 40 minutes to teach the students the prerequisite skills of using the smart phones. Moreover, we informed the students all the reminders to let them have the preliminary understanding. After all these steps, the IPG learning was conducted. A total of forty-three 5th grade students from class A and B of a local school in Tainan were invited as participants. In order to confirm the two classes had the same entry level, a pre-test was given to either class
before the pervasive game procedure. The result of t-test \( (p = .814) \) indicated that these two classes were homogenous in entry level. The statistical data of the two classes are shown in table 1.

Each student would experience both individual and collaborative learning due to educational fairness. Students’ learning scores of the previous semester were used for an S-type heterogeneous grouping. The two classes were separated into seven groups in total.

During the experiment, each group or individual was given a smart phone to perform the IPG. With the clues provided by the mobile device, everyone could depend on the information to finish learning. During the experimental process, all the students were placed in the same class and learned simultaneously. There were two rounds of IPG. Individual mode of IPG learning was conducted first, and then the collaborative mode. After the IPG, the teacher led the students for content discussion to conclude the whole process. Also, students were required to write the questionnaires about learning attitude as well as to the game.

### Instructional pervasive game system

The smart phones, along with the software of ODK, were used to design the game-based learning. The designers hoped the students to do the deeper observations and conduct in-depth thinking to the problems with the smart phones. The two historical places, Ming-Zhi temple and Zhang-Xing Presbyterian church, were the learning targets for the students in their social science class in the unit about local culture. The two locations present a religious contrast for the students to make the comparisons. With the system, the students could understand the religion-related details like architecture, decorations, goddess, history, religious documentations, and the activities of worshippers. The content could allow the students understand more about the historical and religious cultures of their hometowns.
The instructional strategy of the IPG

This pervasive game put two experimental themes together as one learning unit. Theme one was Ming-Zhi temple and theme two was Zhang-Xing Presbyterian church. Both class A and B adopted the individual learning mode in theme one. Conversely, they all adopted collaborative learning mode in theme two. The exploratory practice lasted for approximately two hours for each theme. In order to make the students to compare the two religions in details, the researchers divided both themes into six parts. They were architecture, valuable features, gods, religious documentations, history, and the worshippers’ activities. The students were required to play the IPG following the instructions from ODK equipped mobile devices. The task instructions were issued in the forms of action request, query, and explanatory information, which include scan, interview, observation, elaboration, photograph, and feedback. A comprehensive description of each instruction is shown in table 2. This design not only adds in the variety of investigation data that needed to be collected by students, it also allows the students to get familiar with the newest technology nowadays. Due to the amount of the observation materials in the temple were much more, so the numbers of the questions were a bit higher than the church, as shown in table 3.

<table>
<thead>
<tr>
<th>The ODK Instruction</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scan</td>
<td>Action</td>
<td>To scan the QR code in the registration station so that students were assigned with a learning task</td>
</tr>
<tr>
<td>Interview</td>
<td>Query</td>
<td>To answer the query by asking on-location people</td>
</tr>
<tr>
<td>Observation</td>
<td>Query</td>
<td>To answer the query by observing the cultural sites</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Query</td>
<td>To answer the query by elaborating</td>
</tr>
<tr>
<td>Picture-taking</td>
<td>Action</td>
<td>To take a picture of designated scene</td>
</tr>
<tr>
<td>Feedback</td>
<td>Information</td>
<td>To provide explanatory information/correct answer to the observational query</td>
</tr>
</tbody>
</table>

The registration station and the task distribution table were placed in the learning location. In the beginning of the learning process, each group or person had to scan the QR code in the registration station so that they were assigned with a learning task (Figure 4).

<table>
<thead>
<tr>
<th>Theme</th>
<th>Scan</th>
<th>Interview</th>
<th>Observation</th>
<th>Elaboration</th>
<th>Picture-taking</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ming-Zhi Temple</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>9</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Zhang-Xing Church</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>21</td>
<td>21</td>
<td>15</td>
<td>12</td>
<td>23</td>
</tr>
</tbody>
</table>

Figure 4. Login screen asking for scanning the QR code (left), and QRcode Scanning in process (right)

In this IPG, students had to complete all the tasks following ODK instructions. They could obtain the following tasks when the previous ones had been done and they returned to the task distribution table to scan the QR code again (Figure 5).
As the ODK was mainly used to collect the data, therefore, the system did not give correct answers right away, instead, explanatory information that corresponding to each observational query was provided to help students to continue explore the location, as shown in Figure 6.

After all missions were completed, students were asked to present and discuss the gaming experiences when they returned to the classrooms. In the meantime, a comprehensive assessment was done to evaluate each group’s performance in terms of speed and accuracy. Prizes were given to the winning groups. A summary of the gameplay structure is shown in figure 7.
Research instrument

The main purpose of this study was to explore the effectiveness of IPG in terms of learning gain and learning attitude. The differences between individual and collaborative learning modes, and student’s attitude toward learning with mobile game were also investigated. Three validated research instruments, including an inventory, a questionnaire, and an achievement test, were employed to collect data.

The mobile device attitude inventory

The Mobile Device Attitude Inventory was adopted from the “Inventory of Student’s Learning Attitude toward Smartphone” originally developed by Chen (2009). Modifications were made to accommodate the context of this study. There are four aspects in this inventory: mobile device anxiety, mobile device confidence, mobile device fondness, and mobile device usefulness. The inventory consists of a total of thirty-eight questions. Mobile device anxiety refers to the threats, fears, nervousness, and uneasiness of using mobile devices. The higher the score one filled out, the lower the anxiety he/she exhibited. Mobile device confidence refers to student’s confidence and the ability of learning with mobile devices. Mobile device fondness suggests the degree of appreciation toward using the mobile devices. Mobile device usefulness indicates the degree of expectations of using mobile devices in contributing to one’s daily life.

A four-point-scale was used for the questionnaire from “strongly agree” to “strongly disagree.” Negative questions were included in the inventory for reliability and were counted reversely. Content validity of the inventory was achieved by expert validation process. The inventory was modified after several discussions with content experts and the teachers involved in the experiment. The reliability was established by a pilot test, in which a total of ninety-nine sixth-grade students from the same school were attained. The Cronbach’s α was .926 for the entire inventory.

The game satisfaction questionnaire

The questionnaire investigates student’s satisfactions with the game from four game factors, including the game rules, game mechanics, game strategies, and student’s social interactions. A total of 20 questions were designed by the study. Each factor consists of 4 to 7 questions. The same scaling scheme and validation process was used. The questionnaire was modified according to the expert’s suggestions and Cronbach’s α of .892 was obtained. It confirmed that a high degree of reliability was also acquired for the questionnaire.

The achievement test

A paper-and-pencil test was performed to evaluate learning achievement after the gameplay. There were eighteen multiple-choice questions and two open-ended questions. Eight multiple-choice questions were for each gaming location, and the other four were general questions. Test items were constructed according to the learning materials embedded in the IPG and the content were about the two gaming locations. For example, the question “Where are the stone lions located?” was for Ming-Zhi temple; and the question: “Why Jesus Christ is not hung on the Cross in the Church?” was for Zhang-Xing Church. The design of the test questions was consulted with school teachers to construct expert validity. The overall reliability of the test was .665. It was obtained from a pilot test in which 104 students were attended.

Research findings

Data were collected for responding to the research objectives. The main objectives of the study are to explore the effectiveness of individual and collaborative learning modes, and the overall learning effectiveness. In addition, student’s attitude toward mobile devices, and satisfaction of the gameplay were also examined. Detailed findings after inferential data analyses are described as follows.
Findings regarding the learning modes

Two different learning methods (i.e., individual learning and collaborative learning) were applied to the two learning locations, one for each. Individual learning mode was applied while exploring the Ming-Zhi Temple; and collaborative learning mode was applied while exploring the Chang-Xing Presbyterian church. The achievement test was given to all students after IPG activities were done. An independent t-test was performed to analyze the statistical significance of the mean difference (see Table 4). It is found that learning method significantly affected the learning achievements, and individual learning mode was better than the collaborative learning mode ($t = 2.453$, $p = .016$). However, this result was incoherent with other research findings (e.g., Gentes, Guyot-Mbodji, & Demeure, 2010; Dyson, Griffin, & Hastie, 2004; Huang, Shih, & Lai, 2011). Further investigation on the differences on the instructional design and the cultural issue seem to be necessary.

<table>
<thead>
<tr>
<th>Table 4. T-test of the achievements from different learning modes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Individual mode</td>
</tr>
<tr>
<td>Collaborative mode</td>
</tr>
</tbody>
</table>

*$p < .05$.

Findings regarding overall learning effectiveness

The overall learning effectiveness was measured by the posttest-pretest gains. We could learn from table 5 that the average score of the pre-test was 45.98 while the post-test was 56.16. An independent t-test was also performed to analyze the statistical significance of the mean difference (see Table 4) from which it is concluded that the score of the posttest was outperformed the pretest ($t = 6.166$, $p < .001$). This result explains that the IPG could benefit the understanding of the historic culture.

<table>
<thead>
<tr>
<th>Table 5. T-test of pretest-posttest gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Posttest</td>
</tr>
<tr>
<td>Pretest</td>
</tr>
</tbody>
</table>

**$p < .01$.**

Results regarding the attitude toward mobile devices

Attitude toward mobile devices was measured by the Mobile Device Attitude Inventory. Table 6 displays the results in regard of each attitudinal factor.

<table>
<thead>
<tr>
<th>Table 6. Descriptive statistics of mobile device attitude variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudinal Factors</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Mobile device anxiety</td>
</tr>
<tr>
<td>Mobile device confidence</td>
</tr>
<tr>
<td>Mobile device fondness</td>
</tr>
<tr>
<td>Mobile device usefulness</td>
</tr>
</tbody>
</table>

The relationships between mobile device attitude and learning achievements were also examined. Results indicated that significant correlations were found for both individual and collaborative learning modes. This means that the higher the mobile device attitude, the higher the learning achievements. Table 7 displays detailed Pearson r’s and significant levels.

In order to further understand the details of how learning attitude would affect the learning outcome, correlation processes were done for each attitudinal factor. The results indicated that the learning achievements were correlated with mobile device anxiety, mobile device fondness, and mobile device usefulness, regardless of the learning mode.
In terms of mobile device confidence, however, the correlation was not significant. Detailed results are shown in Table 8.

**Table 7. Correlations between mobile device attitude and learning achievement**

<table>
<thead>
<tr>
<th>Mobile Device Attitude</th>
<th>Pearson r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning</td>
<td>Individual mode</td>
<td>.427**</td>
</tr>
<tr>
<td>Achievement</td>
<td>Collaborative mode</td>
<td>.326*</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.

**Table 8. Correlations between mobile device attitudes and learning achievements**

<table>
<thead>
<tr>
<th></th>
<th>Individual mode</th>
<th>Collaborative mode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pearson r</td>
<td>p</td>
</tr>
<tr>
<td>Mobile device anxiety</td>
<td>0.38 **</td>
<td>0.01</td>
</tr>
<tr>
<td>Mobile device confidence</td>
<td>0.16</td>
<td>0.15</td>
</tr>
<tr>
<td>Mobile device fondness</td>
<td>0.36 **</td>
<td>0.01</td>
</tr>
<tr>
<td>Mobile device usefulness</td>
<td>0.45 **</td>
<td>0.00</td>
</tr>
<tr>
<td>Overall</td>
<td>0.44**</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.

**Results regarding game satisfaction**

The satisfaction of gameplay was analyzed by game rules, game mechanics, game strategies, and social interaction. The means and standard deviations of each factor are displayed in Table 9.

**Table 9. Descriptive data of Pervasive Game satisfactions**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game rules</td>
<td>3.79</td>
<td>1.36</td>
</tr>
<tr>
<td>Game mechanics</td>
<td>3.39</td>
<td>1.56</td>
</tr>
<tr>
<td>Game strategies</td>
<td>3.61</td>
<td>2.51</td>
</tr>
<tr>
<td>Social interaction</td>
<td>2.81</td>
<td>2.57</td>
</tr>
</tbody>
</table>

While looking into specific questions answered on each factor, only three learners thought the game rules could not enhance learning; all others considered that clear game rules made the exploring process more motivating and interesting. We further interviewed the three learners who felt negatively on game rules that they expressed game rules were too strict, which made them feel pressures.

As the aspect of game mechanics, there were 93% of learners who thought the clues provided by the pervasive game system were clear; 97.7% of learners thought that scanning QR code was interesting; 95.3% thought photographing might help them learn; though 37% of learners thought inputting the characters to answer the questions were a little difficult. Therefore, adding more practices of inputting words into mobile devices might make the procedure of the pervasive game easier.

In regard to the game strategies, when the students were in the individual learning mode, 97.7% of the students relied on the clues; 76.6% tried to seek for help with other classmates while encountering the difficulties; 95.4% of them resort to someone from the temple or the church to solve the problems. When the students were in the collaborative learning mode, group members divided their works, shared exploring duties, and exchanged the information acquired; 67.4% of students agreed that the more fluent in operating the mobile devices, the more responsibilities would take; moreover, 95.4% felt that the group action somehow restrained each other to be more engaged in the exploring activities. Interestingly, regardless of these quantitative data, we also observed that in individual learning mode, some students insisted doing all works by themselves except when they encountered technical difficulties. Nevertheless, in collaborative learning mode, most students took the duties in turn, and they would be more likely to follow the game instructions than they were in individual mode.

Finally, in terms of social interaction, under individual learning mode, 32% of the students viewed all other classmates as competitors; under collaborative mode, 88.2% of the students liked to learn by inquiring experts;
83.7% of them would like to actively provide their own opinions to group members; 58.1% of the students felt that the collaborative learning had less difficulties; however, 79.1% of them preferred individual learning. In summary, the students felt a sense of hostility while in individual learning, but had less learning problems while in collaborative learning.

To check further about how different aspects of the game satisfactions would correlate with the learning outcomes, Pearson correlation coefficient analyses were done for all four satisfactory factors. The results are shown in Table 10.

<table>
<thead>
<tr>
<th>Learning Achievements</th>
<th>Pearson $r$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game rules</td>
<td>0.21</td>
<td>0.09</td>
</tr>
<tr>
<td>Game mechanics</td>
<td>0.18</td>
<td>0.13</td>
</tr>
<tr>
<td>Game strategies</td>
<td>0.15</td>
<td>0.16</td>
</tr>
<tr>
<td>Social interaction</td>
<td>0.09</td>
<td>0.28</td>
</tr>
</tbody>
</table>

It is found that there was no significant correlation between any of the satisfactory factors and overall learning achievements indicating that the level of game satisfaction has no relation to the level of learning achievements.

In addition, the correlations between game satisfactions and mobile device attitudes were also analyzed. Table 11 summarizes the results of the multiple correlation tests.

<table>
<thead>
<tr>
<th>Anxiety</th>
<th>Confidence</th>
<th>Fondness</th>
<th>Usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson $r$</td>
<td>$p$</td>
<td>Pearson $r$</td>
<td>$p$</td>
</tr>
<tr>
<td>Game rules</td>
<td>0.03</td>
<td>0.42</td>
<td>0.16</td>
</tr>
<tr>
<td>Game mechanics</td>
<td>0.20</td>
<td>0.10</td>
<td>0.34**</td>
</tr>
<tr>
<td>Game strategies</td>
<td>-0.03</td>
<td>0.42</td>
<td>0.22</td>
</tr>
<tr>
<td>Social interaction</td>
<td>0.05</td>
<td>0.37</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**$p < .01$.**

It is found that significant results were scattered in the table. Mobile device fondness was correlated with game rules ($r = .45, p < .01$) and game mechanics ($r = .39, p < .01$); mobile device usefulness was also correlated with game rules ($r = .53, p < .01$) and game mechanics ($r = .58, p < .01$); mobile device confidence was correlated with game mechanics ($r = .34, p = .01$). Some specific patterns were found in these results. Viewing horizontally, game mechanics was the game factor that mostly correlated with the mobile device attitudes (3 out of 4 factors were correlated); viewing vertically, mobile device fondness and mobile device usefulness were the two attitudinal factor that mostly correlated with the game factor (2 out of 4 factors were correlated). It indicated that if the agreement with the game rules were high, the fondness and usefulness toward using the smart phones would also be high. Moreover, if the level of devotion for the game mechanics were high, the confidence, fondness and usefulness toward using the smart phones would also be high. Nevertheless, the participants’ game strategies and social interactions had less correlation with learning attitudes.

**Conclusions**

The study took forty-three fifth grade students as the research subjects. The IPG was performed in two stages. Stage one applied individual learning mode and stage two applied collaborative learning mode. All subjects went through the same learning procedure. Differences on learning achievements between groups were analyzed by the independent $t$-test. Correlations among mobile device attitudes and learning achievements, and correlations among game satisfactions and mobile device attitudes were analyzed by Pearson Product-Moment Correlation process. The results of these analyses showed the IPG performance, and demonstrated the relations of the learning attitudes and gaming attitudes while playing the IPG.
It was proved in this study that it was beneficial to explore the historical religious relics through playing the IPG. Furthermore, this study identified that the individual learning mode was outperformed the collaborative learning mode in terms of learning achievements. Supplemented by observational data, it is found that the reason for better performance for the individual learning mode was that collaborative learning mode could not provide equal opportunities that allow all group members play with the mobile devices. As a result, the participants would be easily distracted from the tasks. On the contrary, students made better performance in the individual learning mode may be because everyone had his/her own device to proceed with learning tasks.

The study also found that the higher the mobile device attitudes, the higher the learning achievements, regardless of the learning mode. It is shown that if students had less anxiety and higher degree of mobile device fondness, they would have better learning achievements. Once students felt that using mobile device to learn is really beneficial, they would perform well in the IPG. However, the mobile device confidence had little to do with the learning achievements.

In terms of the relationship between game satisfaction and learning achievements, it is found that the higher the level of agreement with the game rules, the higher the learning achievements. We also found that the higher the fondness of the game mechanics, the higher the learning achievements. However, the game strategies and social interaction had little to do with the learning achievements.

With regard to the relationship between the game satisfaction and mobile device attitudes, game rules and game mechanics would affect the mobile device attitudes. A higher level of agreements with the game rules would accompany with a higher fondness and usefulness of mobile device. Meanwhile, a higher fondness toward the game mechanics would be associated with a higher confidence, fondness, and usefulness of mobile device. However, the game strategies and social interaction had little correlation with the mobile device attitudes.

In conclusion, the IPG was found to enhance learning. Facilitated by the mobile technology, IPG helped students to explore the neighborhood and facilitate them to be more involved in the context. This has made the learning process more experiential, contextual-oriented, and reflexive. Moreover, learning mode and mobile device attitudes seemed to affect the learning achievements. The most confounding result found in this study was that the individual learning mode outperformed the collaborative learning mode in the IPG due to equal opportunity of owning and controlling the mobile device. With sufficient mobile devices, the learners could immerse more in the game-based learning process enjoy the learning.

Acknowledgements

This study is supported in part by the National Science Council of the Republic of China, under contract numbers NSC 98-2511-S-024-006-MY2, NSC 100-2628-S-024-002-MY3 and NSC101-2511-S-024-009-MY3.

References


Chavez, D. J., Schneider, I., & Powell, T. (2004, June). The social-psychology of a technology driven outdoor trend: Geocaching in the USA. Paper presented at the meeting of Hawaii International Conference on Social Sciences, Honolulu, HI.


Development and Validation of the Online Instructor Satisfaction Measure (OISM)

Doris U. Bolliger1*, Fethi A. Inan2 and Oksana Wasilik1

1University of Wyoming, 1000 University Avenue, Laramie, Wyoming 82071, USA // 2Texas Tech University, Box 41071, Lubbock, TX 79409, USA // dorisbolliger@gmail.com // inanfethi@gmail.com // oksanapania@gmail.com

*Corresponding author

(Submitted October 11, 2012; Revised February 14, 2013; Accepted July 2, 2013)

ABSTRACT

The purpose of this study was to develop a self-reported measure of instructor satisfaction within the context of teaching in the online environment and to validate the instrument’s psychometric properties. Upon conducting an expert panel review and pilot study, the instrument was administered to 168 instructors who taught courses in the online environment at a large, public western university. Results of an exploratory factor analysis confirmed a hypothesized five-factor model: Instructor-to-Student Interaction, Affordances, Institutional Support, Student-to-Student Interaction, and Course Design/Development/Teaching. The 27-item Online Instructor Satisfaction Measure (OISM) proved to be psychometrically sound with good factor structure and high internal reliability coefficients for the instrument and its subscales.

Keywords

Higher education, Faculty satisfaction, Online teaching, Online course, Factor analysis

Introduction

More and more higher education institutions in the U.S. have implemented courses, degree programs, and certificates that are delivered entirely online. Leaders in academia expect enrollment to increase significantly every year, and online credit offerings are now an integral part of many higher education institutions’ long-term strategic plan (Bourne & Moore, 2005). Academic leaders at institutions assume that instructors and students are satisfied with their online course offerings (Allen & Seaman, 2004). Many leaders are convinced that the quality of online courses and student outcomes in this learning environment are at least equivalent if not better than compared to residential courses (Allen & Seaman, 2010).

One of the reasons for the increase in student enrollment in online courses is the increased need of students to have access to alternative methods of education. The student body at many universities has changed to include a high percentage of nontraditional learners (Blakely & Tomlin, 2008; Snyder & Dillow, 2011) who might be unable to attend a university campus due to many other roles and responsibilities such as work and/or family (Caffarella, 2002; van Enckevort, Harry, Morin, & Schütze, 1986). However, many individuals feel the need or wish to continue their formal education or participate in professional development opportunities. For them, online academic courses and programs provide access to education and are a good fit for individuals with busy schedules. The growth of enrollment in online courses offered by colleges and universities has continued for the past seven years—in order to meet student demand. It has by far exceeded the growth of overall student enrollment growth, and this trend is expected to continue. In fall 2009, the number of learners enrolled in at least one online course exceeded 5.6 million (Allen & Seaman, 2010).

As the number of online students and subsequently online course and degree programs offerings increase, so does the number of instructors who are being tasked to teach online. In a study conducted by Seaman (2009), 34.4% of instructors surveyed had taught at least one online course, and approximately 23.6% were teaching an online course at the time the study was conducted. Many research efforts have been devoted to investigating important elements of faculty adoption of technology in teaching (D’Silva & Reeder, 2005), participation in distance education (Clay, 1999; O’Quinn & Corry, 2002), and what motivates instructors to teach online (Panda & Mishra, 2007).

Instructors are not only important in the success of meeting university goals and outcomes, they also have an impact on the success of academic programs because “faculty play an essential role in developing and rethinking online courses” (Meyer, 2006, p. 43). The commitment of faculty to deliver quality programs and courses is documented in the literature (Curran, 2008). One important aspect in the delivery of online courses and programs, however, is
faculty satisfaction. Faculty satisfaction is so important that the Sloan Consortium includes it as one of the five pillars in the quality framework for online education (Moore, 2002).

Theoretical framework

Faculty satisfaction

In the context of this study, instructor satisfaction is defined as the perception that the process of teaching in the online environment is efficient, effective, and beneficial for the individual. It is one element in the quality framework for online education developed by the Sloan Consortium (Moore, 2002). Researchers have documented that many instructors are satisfied teaching in the online environment and are willing to continue to teach online (Conceição, 2006; Hartman, Dziuban, & Moskal, 2000; Fredericksen, Pickett, Shea, Pelz, & Swan, 2000). Some instructors, however, are critical of teaching online. Instructors are concerned about limited interaction (Bower, 2001), and they are aware that online teaching is time consuming and labor intensive which can easily lead to burnout (Hogan & McKnight, 2007).

Faculty and student satisfaction are not only two integral elements of the quality framework (Moore, 2002), they tend to impact each other. Student satisfaction was defined as an important element in faculty satisfaction by Bolliger and Wasilik (2009). Student satisfaction is defined as the value students associate with their educational experiences in formal settings (Astin, 1993). Important factors in student satisfaction are the instructor, technology, and interactivity (Bolliger & Martindale, 2004). When students are satisfied with their online learning experiences, one can argue that faculty may be more satisfied with their online teaching experience than when students are less satisfied.

Four significant themes in the literature related teaching online from the faculty perspective emerge. They include: (a) interaction between students and peers and between instructors and students; (b) instructor planning, designing, and delivering online instruction; (c) necessary institutional support; and (d) affordances of online teaching and learning. Each element is discussed below.

Interaction

When designing an online course, one needs to keep in mind that the most important element is not the content but the interaction among course participants (Simmons, Jones, & Silver, 2004). Others agree that facilitating interaction, communication, and creating connections with instructors and peers are instrumental in online courses (Duncan & Young, 2009; Fredericksen et al., 2000; Hartman et al., 2000). Three important types of interaction in distance learning were pointed out by Moore and Kearsley (2012): student interaction with instructors, peers, and content.

Faculty-to-student interaction

Moore and Kearsley (2012) believe interaction is essential in teaching online and a desirable element for instructors. Wasilik and Bolliger (2009) found that more satisfied online instructors had a “high level of interaction with online students” (p. 177) compared to less satisfied instructors. The authors point out that both quantity and quality of interaction is a critical element in faculty satisfaction. Quality interaction is also a motivating factor for faculty (Hiltz, Shea, & Kim, 2007).

Some administrators are concerned about a possible decrease of personal contact with students (Rockwell, Schauer, Fritz, & Marx, 1999), and some instructors who teach online miss the personal interaction associated with being able to meet with students (Fish & Gill, 2009). In contrast, some faculty were able to have more online personal interaction with students (Hiltz et al., 2007) and they felt that they developed strong relationships with students (De Gagne & Walters, 2009).

It is also important to online instructors that students communicate actively with them (Bolliger & Wasilik, 2009). Online instructors typically have the desire and interest to use technology (Panda & Mishra, 2007)—a variety of information communication technologies or newer media exist that can be integrated. Instructors are responsible for
responding to students and for providing feedback by gauging students’ learning outcomes and building effective interventions to improve their performance (Moore & Kearsley, 2012).

Student-to-student interaction

In settings where students do not share the same physical space, instructors or course designers should create opportunities for students to interact with peers because students find interaction with peers “to be stimulating and motivating” (Moore & Kearsley, 2012, p. 133). Instructors are satisfied when students are actively involved in their learning (Bolliger & Wasilik, 2009), engage with course materials, and discuss concepts (Moore & Kearsley, 2012).

High levels of interaction and students sharing resources with peers have been pointed out by faculty members as positive aspects in online teaching (Wasilik & Bolliger, 2009). Online instructors like to see their students share ideas, viewpoints, and experiences. As a strategy, experts recommend organizing groups or teams so that students can collaborate in small groups (Moore & Kearsley, 2012). Another aspect that attracts instructors to online teaching is the fact that in most online courses, all students are required to participate and therefore their voices are heard independently of gender and other demographics (Anderson & Haddad, 2005).

Affordances

Online learning technologies provide new possibilities for instructional use that were not inherent in traditional classroom settings. Those attributes of online learning provide potential for effective teaching and learning activities (Day & Lloyd, 2007). One of the key affordances of the online learning environment is that it can provide different benefits for different learners (Webb & Cox, 2004). For example, online communication can provide a way for shy students to participate in asynchronous discussions, while it can also provide flexibility for students with work and family commitments.

There are many positive elements with online teaching that instructors usually mention, two of which are mentioned most are flexibility and convenience. Online instructors have the opportunity to integrate a variety of resources in online courses such as external links, tutorials, audio or video files, and so forth. They reported they have easy access to online course materials and so do their students (Bolliger & Wasilik, 2009; Fish & Gill, 2009; Seaman, 2009; Wasilik & Bolliger, 2009). One major fact from which online instructors derive satisfaction is that online courses provide access to student populations that otherwise would not have access to higher education. This allows faculty to not only reach students all over the world but also in rural areas (Betts, 1998; Hiltz et al., 2007; Rockwell et al., 1999; Wasilik & Bolliger, 2009).

Online instructors attest that online learning allows for greater flexibility in their schedules (Green, Alejandro, & Brown, 2009; Hiltz et al., 2007; Wasilik & Bolliger, 2009; Young, Cantrell, & Shaw, 2001). The convenience of scheduling allows students with other family and work commitment to have the opportunity to continue their educational goals. One of the other affordances of the online tools is their support for implementing student centered-learning activities. Through effective use of the functionality of various learning tools, online instructors can provide pedagogically effective learning environments where the instruction is highly interactive, supportive, communicative, and social. However, it should be noted that the specific affordances of online communication and learning tools do not define their pedagogical usefulness (Burden & Atkinson, 2008). Therefore, instructors should receive training and instructional support to maximize the benefits of these tools.

Institutional support

Institutional support is cited frequently in the faculty satisfaction literature. These elements pertain to: release time, fair compensation, and rewards in general; adequate tools, training, and technical support; and institutional policies. Faculty in earlier studies mentioned that a lack of release time inhibited them from participating in distance education (Betts, 1998; O’Quinn & Corry, 2002), and they were aware that it is a time consuming endeavor (Bower,
“One of the most important issues . . . is the overall institutional support for the development and implementation of distance education” (Milheim, 2001, p. 538).

Fair compensation is important to instructors who teach online (Bower, 2001; Milheim, 2001; Simonson et al., 2009). Online instructors feel that rewards need to be equivalent to teaching courses on campus. However, compensation was perceived as inadequate due to high workload (Green et al., 2009; Hiltz et al., 2007) and others expressed concern about lack of monetary support, e.g. stipends (O’Quinn & Corry, 2002).

When instructors teach online, they need to have access to adequate and reliable technology–equipment and software (Betts, 1998; Fredericksen et al., 2000). Institutions that participate in online education need to provide adequate training in both, pedagogical issues and technology-related skills (Eliason & Holmes, 2010). These opportunities need to be in place not only before developing and teaching an online course for the first time but there should be ongoing faculty development (Ray, 2009). In a study conducted by O’Quinn and Corry (2002) participating faculty in distance education were concerned about their lack of technology skills. Once faculty and students are online, they need to be supported when technical issues arise. Technical support is instrumental, and therefore individuals were concerned about a lack of technical support. Some participants felt so strongly about their lack of technical skills and the lack of technical support that it prevented them from teaching in distance education environments (Betts, 1998; O’Quinn & Corry, 2002).

Good institutional policies need to be in place before online courses and programs are implemented. Researchers found that in some cases, instructional polices were not in place (Hiltz et al., 2007). The institution should value distance education programs and needs to have good policies for the evaluation and recognition of faculty who teach online; otherwise, teaching online might have a negative impact on faculty tenure and promotion decisions (Milheim, 2001). Another area of concern is the availability of proper and clear policies for copyrights and intellectual properties (Durette, 2000; Palloff & Pratt, 2001; Passmore, 2000; Simonson et al., 2009).

### Online course design, development, and teaching

One of the major concerns of potential or practicing online instructors is the design, preparation, and delivery of online courses, and their impact on workload (Betts, 1998; Green et al., 2009; O’Quinn & Corry, 2002). Opinions on whether workload increases when preparing and teaching online courses are diverse. DiBiase (2000) found that teaching an online course required less time than a campus-based course. Hislop and Ellis (2004) found that teaching online courses did not increase instructors’ workload when considering class size.

These results are contradictory to Conceição (2006) and Visser’s (2000) findings who found the development and delivery of an online course required more time and effort than a campus-based course. Seaman (2009) found that the majority of respondents thought it took more effort to teach and develop an online course, approximately 64% and 85% respectively. Others argue that the delivery of online courses alone takes more time and/or effort than campus-based courses because interactions with online students are more demanding on the instructor, and online communication – in general – takes considerably more time compared to face-to-face communication (Conceição, 2006; De Gagne & Walters, 2009; Hiltz et al., 2007; Stacey & Rice, 2002). Some online instructors feel like they are teaching nonstop (Young et al., 2001).

Instructors are concerned about quality of courses and outcomes, and they are dissatisfied when they do not have some level of control over online courses or programs (Betts, 1998; Bower, 2001). Assessment is a major part of an online course (Simmons et al., 2004) and can be time consuming. Instructor satisfaction is higher when student performance is better (Fredericksen et al., 2000), and high levels of student motivation contribute to instructor satisfaction. One positive aspect about teaching online that instructors mentioned was class management. Some instructors were of the opinion that classes can be managed more easily online than in other environments (Wasilik & Bolliger, 2009).
Purpose and significance

Because satisfaction with teaching in the online environment is a complex issue and it impacts the quality of course offerings and delivery, it was the purpose of this study to develop and validate an instrument that measures satisfaction of instructors who teach online courses. Very few instruments have been developed and validated to measure online instructor satisfaction. One of the published instruments, the online faculty satisfaction survey, designed by Bolliger and Wasilik (2009), has limitations such as: a low percentage of variance explained, only moderate reliability of two subscales, and some complex and unexpected item loadings. The validation study identified a three-factor structure which may not be comprehensive enough to represent all important constructs.

Methodology

Item development

Researchers hypothesized that five factors (instructor-to-student interaction, affordances, institutional support, student-to-student interaction, and course design/development/teaching) influenced instructors’ satisfaction with online teaching. All constructs and their definitions are listed in Table 1. Thirty items ranging from 1-strongly disagree to 5-strongly agree were developed after a review of the literature and existing questionnaires. Several items were adapted or significantly modified from a faculty satisfaction survey developed by Bolliger and Wasilik (2009). After the initial item development process, the draft version of the Online Instructor Satisfaction Measure (OISM) underwent expert review and was tested in a pilot study before administering it to the sample in this study.

Validity

In order to establish the content validity of the instrument, it was reviewed by panel of four experts: (1) three individuals with extensive experience in teaching online at a higher education institution and (2) one distance education administrator who has supported online instructors for many years. Experts received instructions and a draft of the instrument. They were asked to indicate the relevance of all items in relation to the constructs, evaluate the clarity of each item, recommend changes, and provide suggestions. Several scale items were modified based on the recommendations provided by the experts.

Pilot testing and reliability

The revised instrument was administered to 124 faculty members who teach online at a small, western research university in spring 2010. The university is a land-grant institution with an enrollment of over 13,300 students and offers over 170 academic programs. All online instructors were invited to participate in the pilot study via email sent...
by an administrative unit that coordinates online credit programs. Instructors received an invitation with an embedded link to the survey housed in an online survey tool. In order to encourage individuals to participate, they were able to register for a small gift card to a bookstore after the completion of the survey. All responses were voluntary and anonymous.

Seventy-one online instructors (57.3%) completed the survey within a 10-day time period. After the survey administration, the internal reliability coefficient for the instrument was calculated for the instrument ($a = .93$) and its sub-scales: affordances ($a = .86$); teacher-to-student interaction ($a = .79$); institutional support ($a = .79$); student-to-student interaction ($a = .78$); and course design, development, and teaching ($a = .67$). Based on the results of the pilot study, nine scale items underwent minor revision before administering the instrument to the sample in this study.

**Participants and procedures**

The instrument was administered at a master-level public university in the West. The university has seven colleges and offers over 200 undergraduate and more than 10 graduate programs. The institution has an annual enrollment of over 24,000 students. In fall 2010 and/or spring 2011 a total of 241 instructors taught online courses.

The data were collected during spring 2011. All online instructors who taught in fall 2010 and/or spring 2011 were invited to participate via an email that included the link to the online survey housed in the university’s course management system. Participation was voluntary and responses were confidential. Any identifiers were removed before the data analysis began. Participants were able to register for the drawing of four small gift cards that could be redeemed at the university’s bookstore. Nonrespondents received a reminder after two weeks.

The response rate was 71.4% ($N = 172$). All seven colleges were represented in the sample with most respondents teaching in the health profession (20.4%) closely followed by science and technology (16.8%) and arts and humanity (16.8%). Most of the respondents were employed full-time (64.3%) at the university and only a few (1.3%) were part-time faculty. A significant percentage of instructors were adjunct faculty (34.4%). Respondents’ experience with teaching in the online environment ranged from 1 to 21 years ($M = 7.5$ years). Of those who responded, 51% were female and 49% were male, and their ages ranged from 26 to 70 years ($M = 51.5$).

**Data analysis**

Four cases were deleted because one-third or more of the data were missing. Two negative scale items were reversed before descriptive statistics were generated. In order to explore the factor structure, a principal component analysis with promax rotation was performed. Prior to conducting the factor analysis, the adequacy of sample size and factorability were checked. In order to interpret factor structure, the pattern matrix and structure matrix were examined for item loadings. To determine the number of factors to retain, several methods including Kaiser’s (1960) eigenvalues greater than one rule, Cattell’ (1966) scree test, minimum average partial test (O’Connor, 2000; Velicer, 1976), parallel analysis (Horn, 1965), total variance explained, residuals (Mertler & Vannatta, 2010; Stevens, 2010), and interpretability (Mertler & Vannatta, 2010) were considered. After determining the factor structure, Cronbach’s alphas, were calculated in order to test the internal reliability of the instrument and its subscales.

**Sampling adequacy for the factor analysis**

The sample size in this study was 168 participants which is considered adequate as items per factor ratio was above 5:1 (Stevens, 2010). Bartlett’s test of sphericity and Kaiser-Myer-Olkin measure were also checked to see if the data were appropriate for factor analysis (Tabachnick & Fidell, 2007). Both Kaiser-Myer-Olkin measure of sampling adequacy (.830) and the Bartlett’s test of sphericity ($\chi^2 = 1829.5, p = .000$) suggested that the dataset was suitable for factor analysis. Furthermore, Kaiser’s Measure of Sampling Adequacy (msa) for all items was over .7 which is considered very good (Tabachnick & Fidell, 2007).
Outliers

The data was examined for univariate and multivariate outliers. Initially, an examination of $z$ scores and leverage values revealed several potential outliers. An examination of these cases did not indicate any reasons for excluding them from the data set. In order to determine whether the cases significantly impacted the results, factor analyses were performed with and without them. Examination of factor structure, communalities, and percentage of variance showed results were consistent with or without including potential outliers; therefore, they were not deleted from the data set.

Results

Factor analysis

The construct validity was examined using a principal component analysis with promax rotation. Inspection of all criteria including Eigenvalues, total variance explained, minimum average partial test, scree plot, and parallel analysis suggest that four or five components could be retained. Because the original factor structure included five constructs, principal component analysis was conducted to retain five components. The five-factor solution explained 51.5% of the variance. The factor labels proposed by researchers were well suited for the extracted factors and were retained.

Examination of the pattern matrix suggested the loading of six items for the instructor-to-student interaction and six items for the institutional support subscales. Five items loaded on each of the following subscales: affordances, student-to-student interaction, and course design/development/teaching. Three items which did not have strong loadings were excluded. Loading of variables on factors and factor labels proposed by researchers are displayed in Table 2. Variables are ordered and grouped by the size of loadings. Table 3 is a correlation matrix of all scale items, and Table 4 is a component correlation matrix. Some scale items are correlated with each other, whereas components are not correlated.

Table 2. Items, factor loadings, and communalities

<table>
<thead>
<tr>
<th>Scales</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor-to-Student Interaction</td>
<td>ISI-5. I am pleased with the quality of student work in online courses.</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISI-6. I am satisfied with students’ motivation in online courses.</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISI-2. My online students are somewhat passive in their interactions. [R]</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISI-3. My interactions with online students are satisfying.</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISI-1. My online students participate enthusiastically.</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISI-4. I do not get to know my online students well. [R]</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affordances</td>
<td>A-2. I am satisfied with the convenience of the online learning environment.</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A-1. Online courses provide a flexible learning environment.</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A-5. Online courses allow students to access a wide range of resources.</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A-3. Online teaching allows me to reach a more diverse student population.</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A-4. I am satisfied that my students can access their online course from almost anywhere.</td>
<td>0.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institutional Support</td>
<td>IS-1. At my institution, teachers are given sufficient time to design and develop online courses.</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IS-6. My institution provides the necessary technology tools (equipment and software) for teaching online.</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IS-3. My needs for training to prepare for teaching online have been met.</td>
<td>0.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IS-2. I have adequate technical support by my institution.</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IS-4. My institution provides fair compensation or incentives for</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
teaching online.
IS-5. I am satisfied with online teaching policies that have been implemented by my institution. 0.57

**Student-to-Student Interaction**

SSI-2. My online students actively collaborate. 0.82
SSI-4. My students work well together online. 0.74
SSI-1. My online students share resources with each other within the course. 0.71
SSI-3. My students appear to be part of an online community in the course. 0.68
SSI-5. In online courses every student has an opportunity to contribute. 0.40

**Course Design/Development/Teaching**

CDT-1. My online students receive quality feedback. 0.70
CDT-3. It takes a lot of time to develop an online course. [R] 0.60
CDT-2. I am accessible to students in online courses. 0.57
CDT-4. I am satisfied with how I assess students in online courses. 0.46
CDT-5. I am satisfied with the content quality of my online courses. 0.44

**Note.** [R] = reversed item

---

**Table 3. Item coefficient matrix**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SSI-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISI-1</td>
<td>.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISI-2</td>
<td>.49</td>
<td>.46</td>
<td>.53</td>
<td></td>
<td></td>
<td></td>
<td>.34</td>
<td>.41</td>
<td>.29</td>
<td>.15</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>ISI-3</td>
<td>.32</td>
<td>.42</td>
<td>.40</td>
<td>.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISI-5</td>
<td>.64</td>
<td>.45</td>
<td>.40</td>
<td>.26</td>
<td>.28</td>
<td>.28</td>
<td>.15</td>
<td>.02</td>
<td>.11</td>
<td>.11</td>
<td>.11</td>
<td>.11</td>
</tr>
<tr>
<td>ISI-6</td>
<td>.50</td>
<td>.53</td>
<td>.29</td>
<td>.33</td>
<td>.45</td>
<td>.15</td>
<td>.35</td>
<td>.09</td>
<td>.18</td>
<td>.57</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>CDT-1</td>
<td>.42</td>
<td>.38</td>
<td>.28</td>
<td>.34</td>
<td>.15</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDT-2</td>
<td>.07</td>
<td>.06</td>
<td>.02</td>
<td>.02</td>
<td>.03</td>
<td>.14</td>
<td>.44</td>
<td>.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDT-3</td>
<td>.08</td>
<td>.02</td>
<td>.10</td>
<td>.06</td>
<td>.32</td>
<td>.08</td>
<td>.21</td>
<td>.04</td>
<td>.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDT-4</td>
<td>.13</td>
<td>.21</td>
<td>.08</td>
<td>.14</td>
<td>.07</td>
<td>.55</td>
<td>.11</td>
<td>.21</td>
<td>.29</td>
<td>.11</td>
<td>.33</td>
<td></td>
</tr>
<tr>
<td>CDT-5</td>
<td>.06</td>
<td>.45</td>
<td>.40</td>
<td>.26</td>
<td>.28</td>
<td>.28</td>
<td>.15</td>
<td>.11</td>
<td>.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDT-6</td>
<td>.02</td>
<td>.50</td>
<td>.53</td>
<td>.29</td>
<td>.33</td>
<td>.45</td>
<td>.15</td>
<td>.35</td>
<td>.09</td>
<td>.18</td>
<td>.57</td>
<td></td>
</tr>
<tr>
<td>ISI-1</td>
<td>.10</td>
<td>.12</td>
<td>.06</td>
<td>.12</td>
<td>.11</td>
<td>.33</td>
<td>.03</td>
<td>.11</td>
<td>.11</td>
<td>.14</td>
<td>.12</td>
<td>.14</td>
</tr>
<tr>
<td>ISI-2</td>
<td>.19</td>
<td>.21</td>
<td>.12</td>
<td>.15</td>
<td>.04</td>
<td>.11</td>
<td>.01</td>
<td>.12</td>
<td>.27</td>
<td>.02</td>
<td>.22</td>
<td>.12</td>
</tr>
<tr>
<td>ISI-3</td>
<td>.08</td>
<td>.16</td>
<td>.14</td>
<td>.04</td>
<td>.11</td>
<td>.01</td>
<td>.12</td>
<td>.27</td>
<td>.02</td>
<td>.22</td>
<td>.12</td>
<td>.14</td>
</tr>
<tr>
<td>ISI-4</td>
<td>.16</td>
<td>.15</td>
<td>.15</td>
<td>.03</td>
<td>.06</td>
<td>.06</td>
<td>.18</td>
<td>.01</td>
<td>.07</td>
<td>.14</td>
<td>.33</td>
<td>.10</td>
</tr>
<tr>
<td>ISI-5</td>
<td>.15</td>
<td>.14</td>
<td>.06</td>
<td>.10</td>
<td>.01</td>
<td>.12</td>
<td>.02</td>
<td>.12</td>
<td>.17</td>
<td>.10</td>
<td>.24</td>
<td>.36</td>
</tr>
<tr>
<td>ISI-6</td>
<td>.16</td>
<td>.15</td>
<td>.15</td>
<td>.01</td>
<td>.06</td>
<td>.03</td>
<td>.06</td>
<td>.03</td>
<td>.07</td>
<td>.16</td>
<td>.07</td>
<td>.10</td>
</tr>
<tr>
<td>A-1</td>
<td>.10</td>
<td>.12</td>
<td>.06</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.10</td>
<td>.04</td>
<td>.03</td>
<td>.04</td>
<td>.03</td>
<td>.04</td>
</tr>
<tr>
<td>A-2</td>
<td>.10</td>
<td>.12</td>
<td>.06</td>
<td>.12</td>
<td>.10</td>
<td>.10</td>
<td>.12</td>
<td>.04</td>
<td>.03</td>
<td>.06</td>
<td>.03</td>
<td>.03</td>
</tr>
<tr>
<td>A-3</td>
<td>.10</td>
<td>.12</td>
<td>.06</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.04</td>
<td>.05</td>
<td>.03</td>
<td>.05</td>
</tr>
<tr>
<td>A-4</td>
<td>.10</td>
<td>.12</td>
<td>.06</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.05</td>
<td>.03</td>
<td>.05</td>
<td>.05</td>
</tr>
<tr>
<td>A-5</td>
<td>.10</td>
<td>.12</td>
<td>.06</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.03</td>
<td>.05</td>
<td>.03</td>
<td>.05</td>
</tr>
<tr>
<td>A-6</td>
<td>.10</td>
<td>.12</td>
<td>.06</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.10</td>
<td>.12</td>
<td>.02</td>
<td>.05</td>
<td>.02</td>
<td>.02</td>
</tr>
</tbody>
</table>

**Table 4. Subscale correlation matrix**

<table>
<thead>
<tr>
<th>Subscale</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Instructor-to-Student Interaction (ISI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Affordances (A)</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Institutional Support (IS)</td>
<td>.18</td>
<td>.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Student-to-Student Interaction (SSI)</td>
<td>.33</td>
<td>.20</td>
<td>.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Course Design/Development/Teaching (CDT)</td>
<td>.27</td>
<td>.24</td>
<td>.21</td>
<td>.05</td>
<td></td>
</tr>
</tbody>
</table>

**Reliability analysis**

In order to determine the instrument’s internal consistency reliability, Cronbach’s alpha coefficients were calculated. The final Online Instructor Satisfaction Measure (OISM) includes 27 items, and its reliability was high (.87).
Similarly, the subscale reliability was acceptable for all factors: Instructor-to-Student Interaction (.82), Affordances (.80), Institutional Support (.75), Student-to-Student Interaction (.77), and Course Design/Development/Teaching (.64).

Descriptive statistics

Overall, instructors in this study were moderately satisfied with teaching online. They were the least satisfied with interaction components. Items pertaining to both interaction factors, student-to-student and instructor-to-student interaction, received the lowest mean scores: items SSI-5 ($M = 3.01$) and SSI-7 ($M = 3.12$), and ISI-4 ($M = 3.06$) and ISI-12 ($M = 3.13$). Table 5 displays mean scores and standard deviations for all scale items. The standard deviations were relatively minor.

Table 5. Mean scores and standard deviations for scale items ($N = 168$)

<table>
<thead>
<tr>
<th>Instructor-to-Student Interaction (ISI)</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISI-1</td>
<td>3.59</td>
<td>.72</td>
</tr>
<tr>
<td>ISI-2</td>
<td>3.06</td>
<td>.94</td>
</tr>
<tr>
<td>ISI-3</td>
<td>3.90</td>
<td>.80</td>
</tr>
<tr>
<td>ISI-4</td>
<td>3.13</td>
<td>1.08</td>
</tr>
<tr>
<td>ISI-5</td>
<td>3.73</td>
<td>.85</td>
</tr>
<tr>
<td>ISI-6</td>
<td>3.30</td>
<td>.95</td>
</tr>
<tr>
<td>Affordances (A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-1</td>
<td>4.28</td>
<td>.75</td>
</tr>
<tr>
<td>A-2</td>
<td>4.26</td>
<td>.70</td>
</tr>
<tr>
<td>A-3</td>
<td>4.03</td>
<td>.91</td>
</tr>
<tr>
<td>A-4</td>
<td>4.48</td>
<td>.55</td>
</tr>
<tr>
<td>A-5</td>
<td>4.15</td>
<td>.69</td>
</tr>
<tr>
<td>Institutional Support (IS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IS-1</td>
<td>3.31</td>
<td>1.07</td>
</tr>
<tr>
<td>IS-2</td>
<td>4.22</td>
<td>.83</td>
</tr>
<tr>
<td>IS-3</td>
<td>4.16</td>
<td>.71</td>
</tr>
<tr>
<td>IS-4</td>
<td>3.30</td>
<td>1.10</td>
</tr>
<tr>
<td>IS-5</td>
<td>3.47</td>
<td>.94</td>
</tr>
<tr>
<td>IS-6</td>
<td>3.94</td>
<td>.92</td>
</tr>
<tr>
<td>Student-to-Student Interaction (SSI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSI-1</td>
<td>3.47</td>
<td>.97</td>
</tr>
<tr>
<td>SSI-2</td>
<td>3.01</td>
<td>.96</td>
</tr>
<tr>
<td>SSI-3</td>
<td>3.22</td>
<td>.97</td>
</tr>
<tr>
<td>SSI-4</td>
<td>3.12</td>
<td>.89</td>
</tr>
<tr>
<td>SSI-5</td>
<td>4.06</td>
<td>.78</td>
</tr>
<tr>
<td>Course Design/Development/Teaching (CDT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDT-1</td>
<td>4.20</td>
<td>.67</td>
</tr>
<tr>
<td>CDT-2</td>
<td>4.56</td>
<td>.57</td>
</tr>
<tr>
<td>CDT-3</td>
<td>4.42</td>
<td>.77</td>
</tr>
<tr>
<td>CDT-4</td>
<td>4.00</td>
<td>.82</td>
</tr>
<tr>
<td>CDT-5</td>
<td>4.07</td>
<td>.75</td>
</tr>
</tbody>
</table>

Table 6. Summary statistics ($N = 168$)

<table>
<thead>
<tr>
<th>Subscale</th>
<th>No. of Items</th>
<th>Cronbach's alpha</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ISI</td>
<td>6</td>
<td>.82</td>
<td>3.39</td>
<td>.65</td>
</tr>
<tr>
<td>2. A</td>
<td>5</td>
<td>.80</td>
<td>4.24</td>
<td>.54</td>
</tr>
<tr>
<td>3. IS</td>
<td>6</td>
<td>.75</td>
<td>3.73</td>
<td>.63</td>
</tr>
<tr>
<td>4. SSI</td>
<td>5</td>
<td>.77</td>
<td>3.38</td>
<td>.66</td>
</tr>
<tr>
<td>5. CDT</td>
<td>5</td>
<td>.64</td>
<td>4.25</td>
<td>.46</td>
</tr>
</tbody>
</table>
Instructors agreed most with items on the Course Design/Development/Teaching (CDT) and Affordances (A) subscales. The subscales with the highest mean scores were CDT ($M = 4.25$) and A ($M = 4.24$). Table 6 includes the summary of psychometric values for the five subscales that comprise the OISM.

**Discussion and conclusion**

The purpose of this study was to develop a self-reported measure of instructor satisfaction in the online environment building on previous research and instruments. Results indicate that the developed instrument, OISM, is a valid and reliable instrument that can be used by administrators and researchers to gauge faculty satisfaction. It can be used to assess effects of course participants’ interaction, existing institutional support, and affordances of technologies integrated in online teaching.

The final version of the validated OISM is comprised of a total of 27 items after some modifications (e.g., excluding a few items that proved to be problematic) were made that resulted in a more refined version of the instrument and confirmed the hypothesized five-factor model. Overall, the scale has a good factor structure and acceptable internal reliability. However, the reliability coefficient for one of the subscales, Course Design/Development/Teaching, was not very high. This may be due to the nature of the subscale; it is comprehensive and attempts to address various teaching elements. The findings suggest that there may be an underlying subcomponent which future research could explore.

Participating instructors reported they are generally satisfied with their online teaching experiences. Undoubtedly, some are more satisfied than others, and some individuals experience more satisfaction with certain elements that pertain to online teaching than others. The examination of descriptive results indicated that the two scales related to interaction had the lowest mean scores. These results could pertain to the nature of the online learning environment where instructors need different competencies to integrate interactive learning activities. Instructors may be unfamiliar with instructional strategies or newer media that facilitate interaction and promote student engagement and therefore could perceive online teaching as too work intensive (Conceição, 2006).

Affordances of these technologies were rated highly. However, administrators should not assume that affordances provided by properties of online technologies will be realized and utilized by instructors. Instructors should receive sufficient guidance and pedagogical support to understand how online communication and collaboration tools can be used to improve interaction and learning. Results show that instructor satisfaction with institutional support was not high. Considering that online instructors may invest more time than instructors who teach face-to-face (Reinheimer, 2005), institutional support becomes a very important component of instructor satisfaction and the quality of the instruction provided. If sufficient and timely support is not provided, some online instructors will experience burnout because of time commitments and demands that are associated with online teaching. Hogan and McKnight (2007) investigated burnout levels of online instructors and found their subjects were “on the borderline of burnout showing signs of moving to a high degree of burnout” (p. 122).

Researchers have pointed out that satisfaction of online instructors is an important topic in education because it has the potential to influence the quality of instruction and student outcomes. Therefore, administrators should monitor general levels of satisfaction of online instructors because they are important in the universities mission, goals, outcomes, and overall performance—instructors are an important asset (Green et al., 2009). If offering online programs and courses are an important element of a university’s mission and long-term strategic plan (Bourne & Moore, 2005), and universities are committed to deliver high quality online offerings and instruction, then it will be important for administrators to retain qualified instructors who are motivated to teach online, satisfied with teaching online, and deliver effective instruction.

Some limitations need to be pointed out. All data were self-reported; therefore, the scale measures perceived levels of instructor satisfaction. The sample consisted of faculty at a single university who utilized one course management system. The study is therefore geographically limited, and findings should be interpreted with caution as they may not be generalizable to instructors who utilize different delivery systems. In the future, researchers could include multiple data collection sites or institutions that provide faculty a choice of course management systems. The scale could be used to survey online instructors in different settings, e.g., K-12, community colleges, private universities,
or business and industry. The scale may also be useful to assess satisfaction of instructors who teach in blended settings and integrate a combination of delivery tools.

Acknowledgements

We would like to thank Christine Boggs, John Cochenour, Larry Jensen, and Craig Shepherd for serving on our expert panel and reviewing the instrument. We greatly appreciate their time, consideration, and support. Their critical feedback and suggestions assisted us in the improvement of the OISM.

References


Experiences and Challenges of International Students in Technology-Rich Learning Environments

Laurence Habib*, Monica Johannesen and Leikny Øgrim
Oslo and Akershus University College of Applied Sciences, PO Box 4, St Olavs Plass, 0130 Oslo, Norway // laurence.habib@hioa.no // monica.johannesen@hioa.no // leikny.ogrim@hioa.no

*Corresponding author

(Submitted November 9, 2012; Revised May 21, 2013; Accepted July 11, 2013)

ABSTRACT
This article presents a study of international students and their use of technology in a Scandinavian institution of Higher Education. A special emphasis is placed on patterns of use of a virtual learning environment (VLE) that is available to all the study programmes at the institution. Actor-Network Theory (ANT) is used as a theoretical approach to focus on the socio-material nature of the various networks that students, teachers, course designers, and artefacts make up within the realm of the institution. Qualitative data were collected through interviews with forty informants, all of them students or staff members at the studied institution. The main findings of the study are that the following factors are essential in the educational experience of international students: the students’ level of digital literacy, their degree of understanding of academic and administrative language, and the types of technology that are used in communication. The article also suggests that technology as a socio-material assemblage may encapsulate cultural codes that can be alienating for international students and that there is a need to “open the black boxes” of technology to cater for the needs of international students.

Keywords
Higher education, International students, Educational technology, Actor-Network Theory, Socio-materiality

Introduction
Trends in international Higher Education point toward an increase in technology-enhanced education (online education). In this context, the use of virtual learning environments (VLEs) has a particularly strong standing, as this type of learning technology is widely implemented throughout Higher Education. VLEs are online systems that typically allow for making course material available in a structured way and provide a platform for synchronous and asynchronous discussions. Literature on VLE use in Higher Education reports that the use of VLEs contributes to the shaping of teaching and learning activities (Blin & Munro, 2008; McGill & Hobbs, 2008), even if the technology is used primarily for making teaching material available online, for the administration of courses, or for automating time-consuming processes such as testing. However, different institutions and different disciplines and professions use the technology differently (Johannesen & Habib, 2010a). In some cases, the technology is used as a tool to support critical thinking, scaffold collaboration and facilitate processes whereby students produce their own study material and create their own learning tasks (Johannesen & Habib, 2010b). Over the last few years, social media have made their entrance on the educational scene, often raising hopes that they would deliver novel learning experiences to students (Shaltry, Henriksen, Wu, & Dickson, 2013), but questions have been raised about their capacity to facilitate debate (Friesen & Lowe, 2011).

Scandinavian institutions are attractive on the international Higher Education market, presumably due to relatively low tuition. This situation brings about challenges for Scandinavian universities and colleges, which have to cater for a growing diversity among students, in terms of language, culture, and academic background. In this article, we examine how international students use and relate to learning technologies in a predominantly monocultural context, using the theoretical and methodological lens of Actor-Network Theory (ANT).

Literature review
International students cannot be characterized as a homogenous group, but the relevant literature generally deals with one of two categories of international students: (1) students from what can be referred to as the Global South (which is also commonly labeled as the Third World or developing countries), and (2) students enrolled in institutions from the Global North on an exchange or internationalization programme.
Students from the Global South have probably experienced the so-called digital divide, a divide in terms of economy, access, knowledge and power (Ferro, 2010; Haddad, 2002; Hilbert, 2011; Wolff & MacKinnon, 2002; Carm & Øgrim, 2013). Many countries in the Global South are lagging far behind the North when it comes to technological infrastructure and penetration of personal technology (InternetWorldStats, 2012), even if the Global South as a whole experiences the highest rise in technology index (World Bank, 2009). As a result, the students from the Global South are likely to lack familiarity with technology that their Scandinavian counterparts may take for granted. In Scandinavia, those students normally enrol into the regular teaching, which is typically provided in a Scandinavian language (or, exceptionally, in English, if required by the study programme).

Students enrolled in institutions in the Global North come to Scandinavia typically for a limited period of time (one or two semesters) mostly with funding from internationalization programmes such as Erasmus. They are typically in their middle or final year of studies, and have therefore acquired some knowledge of Higher Education from their first few years of study. There may be differences in terms of level of available technology between their home institutions and the Scandinavian institutions they come to, but those are generally less significant than those that their counterparts from the Global South are likely to experience.

The growing internalisation of Higher Education has raised some concern about the integration of international students and their adaptation to an unfamiliar academic culture and environment (e.g., Guo & Chase, 2011). In the growing research literature on the subject, a number of key issues related to cross-cultural differences in Higher Education may be identified. Some of the scholarly literature focuses on variances regarding teaching and learning style (e.g., Heffernan, Morrison, Basu, & Weeney, 2010) while other research studies have a stronger focus on learning philosophies (e.g., Chen & Bennett, 2012). It is also interesting to note that technology acceptance has been identified as a cultural issue that plays a major role in today’s learning experience (Yoo & Huang, 2011).

Scandinavian teaching and learning culture can be characterized as focusing on critical thinking and reasoning, and as being to a large extent grounded in socio-cultural learning theories (Favorin & Kuutti, 1996; Illeris, 2009a, 2009b). In Scandinavia, the tradition is that students and faculty members relate to each other on a very “flat” basis, with little difference in status (Arnesen & Lundahl, 2006; Kansanenab, 1999). The students are expected to be self-regulating in their learning activities, and collaboration and peer learning are favoured as learning methods.

The Scandinavian model of education fits in a broader Western academic culture, which is generally considered to be more interactive and student-centred than Asian and African academic cultures, which have been described as more power-distant and teacher-centred (Hofstede, 2001). More generally, a number of research works on international student experience focus on describing and analysing the relationship between learning style preferences and cultural backgrounds (e.g., Charlesworth, 2008). Attributes such as perfectionism (Nilsson, Butler, Shouse, & Joshi, 2008) or reverence for teachers’ authority (Chiu, 2009) have been reported as characteristic to Asian students, also when studying in Western institutions. Nevertheless, describing the influence of culture on learning styles is not a straightforward task, and concerns have been raised about the validity and reliability of studies that propose models of culture-specific learning styles (Eaves, 2011). An important question that has been raised is whether the learning styles of international students are predetermined by their cultural background or whether their learning practices evolve as they move from one learning environment to another (Wong, 2004).

In this landscape, learning technologies (such as VLEs) emerge as key elements to understand how international students relate to technology. Our investigation has therefore focused on identifying patterns of use of learning technologies among international students.

**Theoretical framework**

The starting point of our quest for theories and methodologies was an acknowledgement of the need to find approaches that can handle the complex nature of the problem at hand, involving a large number of elements (students, academics, managers, home institutions, host institutions, native language, technology, academic language, etc.) without being reductionist. In this quest, we quickly realised that socio-material approaches (Fenwick, Edwards, & Sawchuk, 2011; Sørensen, 2009), including ANT, had much potential to support our research. Previous experience
with using ANT (Johannesen, 2013; Johannesen, Erstad, & Habib, 2012) and the literature on socio-material approaches in education (Fenwick, 2011; Fenwick & Edwards, 2011; Saito, 2010) suggest that those approaches are well suited to inform scholarly inquiry in areas where the relationships between the various elements of investigation are fast-changing and do not easily lend themselves to a clear categorization. A socio-material approach such as ANT was therefore a natural choice in an attempt to make sense of the “messiness” of the dynamics that characterised the research topic.

ANT emerged in the late 1970’s within the field of sociology of technology and sociology of innovations. The notion of “theory” might not fully encapsulate what ANT represents, as it can also be seen as an approach or perspective with both philosophical and methodological inferences. Its main tenets repose on a fundamental questioning of what at the time seemed to be a unanimously accepted view of the world as neatly divided into discrete spheres such as “the social”, “the natural” and “the technological.” ANT offered an alternative understanding of the world as made up of various networks or assemblages of interrelated entities. The originality of ANT resides in the “hybrid” character of those networks, i.e., their capacity to include both humans and non-humans (i.e., animals, technological artefacts, viruses, etc.). In order to provide a coherent framework for the purpose of social enquiry, ANT introduces the notion of “actant,” i.e., significant elements within a network that may be either human or non-human.

ANT uses a wide and often changing apparel of concepts, some of which are recurrent throughout the literature, while others are favoured more particularly by one scholar or a group of scholars. The concepts that we choose to use in this article are largely common to those that pervade the traditionally marginal, but recently growing literature within the field of educational science (Fenwick & Edwards, 2010; Fox, 2009; Nespor, 1994). From an ANT perspective, educative activities can be conceptualized as complex webs involving actants, i.e., both humans and non-human elements (including physical buildings, curricula, learning technologies, etc.) that are constantly evolving and, as they do so, are “negotiating” with new elements (Latour, 1999), and, if successful, “recruiting” or “enrolling” them into the network (Callon, 1986). Networks whose constitutive elements are tightly knit toward the same goal may become so stable and robust that they no longer are questioned and can, in ANT terms, be characterized as “blackboxed” (Lanzara, 1999). Within the realm of an ANT framework, actants, in particular technological artefacts, may be “inscribed” (Akrich, 1992; Akrich & Latour, 1992) with a certain pattern of action, as the team of designers that decide on their functionalities and appearance make such choices on the basis of a particular image they have of future users and future uses. However, with any artefact whose designer is not the end-user, there is a chance that actual use will turn out to be different from the intended use, which reflects what ANT refers to as “translation” (Hanseth, 1996), i.e., a process whereby users adapt, shape or convert an actant to fit their own needs.

Methodology

In this study, we apply an interpretive methodological approach, framing both the students’ study practice and their teachers’ teaching practice. The study has been designed as an explorative case study (Yin, 1989) of international educational programmes at a Higher Education institution. The election of the participants for this study was based on purposive selection criteria (Miles & Huberman, 1994; Patton, 1990). The main criteria used for selection were as follows: (a) being a student enrolled at the studied institution for one semester or more; and (b) having a native language other than Norwegian, Swedish or Danish (as those three languages are highly similar).

The data for this article were collected through interviews with international students at the studied institution, and with academic staff in charge of courses offered to international students. The total number of informants was 40, some of whom were interviewed in groups, while others were interviewed individually. We conducted a total of 10 group interviews of students (group size between three and seven), three individual interviews of students, and one group interview of teachers.

The interviews were semi-structured and lasted around one hour. A number of core questions were asked to all informants, while the follow-up questions differed according to the answers given to the core questions. A selection of typical core questions is provided below, categorized in thematically-related areas:

- Use of learning technologies:
  - What are your main learning activities?
  - How do you use IT/learning platforms/social media generally?
  - How do they relate to your learning activities?
Ease of use and preferences:
How do you consider the role of IT when it comes to the availability of curriculum texts?
Do you have any suggestions for any alternative learning programmes?
Where do you most want to be when carrying out learning activities?
What role does IT availability play in your choice of place when carrying out your learning activities?
What do you think would be the ideal use of IT in your education?

Skills and training:
How did you acquire your IT skills?
Who was involved in your training?
How would you evaluate your IT skills in relation to that of the other students?
How would you evaluate your IT skills according to what is required for your studies?

Writing vs. oral communication:
How would you evaluate the role of IT when it comes to your writing activities?
What do you think of digital communication compared to face-to-face communication?

Cultural differences:
Can you mention situations where you have experienced cultural differences in your education? Can you elaborate on such differences?
Do you consider such differences problematic/advantageous? In what way?

The original plan was to rely on self-recruitment to get hold of minority informants, but despite numerous ads and presentations of the study during lectures, no one volunteered to be interviewed, presumably for fear of being stigmatized. Interviews with minority student informants were therefore carried out during classes where the teacher had defined the interview as being a compulsory part of the study programme. Recruiting international students that were enrolled in an exchange programme or an international master programme was relatively more straightforward, possibly because the process of self-recruitment as international students did not conjure up the same kinds of feelings of stigmatization that minority students may have experienced.

The informants represent two distinct student groups. The first group we identified was a group of students from West or South Asia, such as Iran and Pakistan. They have typically lived in Norway for several years prior to enrolling into their study programme, and are bilingual (using both their mother tongue and the Norwegian language on a daily basis). We will further refer to these students as Asian International students (AI students), which is a sub-set of the group described above of “students from the Global South.” Those students attend either a work-place based Early Childhood programme or a bi-lingual teacher education programme at the Faculty of Education. In both programmes, the learning philosophy leans heavily on a socio-cultural approach to education, where interaction with peers plays a major role in learning (Engeström, 2001; Vygotsky, 1978). There is also a particular focus on encouraging the students to put their own reflections down in written form, and to give feedback to each other, based on those writings. The students enrolled in those programmes can be characterized as mature students, who generally already have a first degree from their home country.

The second group of student respondents in the study consists of students from either European or North-American background. We refer to this group as European or North-American International students (ENAI students), a group that is a sub-set of the above described group of “students from the Global North.” The respondents from the ENAI group all attend the “European Project Semester” (EPS) student exchange programme. These students are typically young adults from a European country, such as Germany, Spain and the Netherlands, although a few may also come from North American institutions. They are used to study in their own mother tongue, and expect the exchange programme to be a cultural and linguistic experience. Typically, a European Project Semester results in a written report, but writing texts is rarely used to support processes of learning.
Core findings: Socio-material assemblages of international students learning environments

In this section, we will present some core findings and use the conceptual framework of ANT to gain a deeper understanding of the international students’ learning environments, with a particular focus on the socio-material assemblages of these technology-rich environments.

Digital literacy

The data material indicates that all the interviewed international students use technology in their study programmes. All are satisfied with using the VLE available at the institution. The students report that all their teachers use the VLE, but with significant variations in volume and type of use. In some programmes, the VLE is used for communication between teachers and students and as a platform to hand in assignments. In others, the use of the VLE is more minimalistic, for example as a repository for links to webpages outside the VLE, often because the teachers feel pressured from their academic management to use the VLE, while they would prefer to just use the web.

Many international students are used to using VLEs from their earlier studies in their home institutions, but those typically use a number of different VLEs, while only one VLE is available at the studied institution. It appears from the interviews that only having to relate to one VLE is a considerable advantage in their studies. However, having to navigate through the system to find the relevant information is not always straightforward, and is therefore a source of dissatisfaction.

The two groups of students interviewed in this study appear to have very different perspectives on VLEs. The AI students consider the VLE as a technological artefact that they need to learn to master, just as any other technological solution. They therefore report a need for specific training in VLE use, and in the use of technological artefacts in general. Most of them report having little technological proficiency from earlier studies, presumably because, as mature students, they only have study experience from a time when technology was mostly absent from Higher Education.

One AI student reports that when asking the administrative staff for help to register required information, she was repeatedly told to “log onto the Student web.” When she explained that she had trouble finding the right form within the generic Student web, the only answer she got was to “log onto the Student web”. The above example illustrates how administrative staff may in some cases overestimate the actual level of digital literacy of some international students, which in turn may create frustrations amongst those students.

In contrast, the ENAI students appear to have a higher level of technical proficiency, presumably because, as younger students, they are more digitally literate, and have been exposed to a range of VLEs and every day technologies over the years. Although none of them have been acquainted to the studied institution’s VLE from earlier studies, they report having had no problems learning to use it, as the VLEs they are used to from their home institutions have similar functions and design. Since the academic language for the EPS programme is English, and few ENAI students use English on a daily basis at their home institutions, they report having to resort to technological help to understand lectures and express themselves correctly in writing. For example, several interviewees report using translation tools such as Google Translate or Lingui.

None of the ENAI students considers VLEs to be core technology, but rather a practical “place” to carry out and organise learning activities, such as creating study groups, handing in assignments, or finding the names of other group members. Their teachers also report having a pragmatic approach to VLE use, often as a gateway to their own webpage.

Types of technology used in communication

Several of the interviewed students—both ENAI and AI students—mention that their learning activities take place within a number of different arenas. Meetings with co-students on school premises are the preferred learning arena, while meetings on social media such as Facebook and Skype are useful supplements to physical meetings. However, they report never using social media in their communication with teachers and rather use e-mail for that purpose.
They consider social media to be arenas for private communication, in particular because they are the repository of information that can be judged inappropriate in an academic context.

Interviewer: “That is interesting, that Facebook is considered private.”
Informant 1: “Lots of privacy”
Informant 2: “You have all your pictures and I don’t know, yes”
[Extract of interview with student informants]

**Academic language**

As far as academic language is concerned, there seems to be a difference between ENAI students and AI students. The ENAI students do not report experiencing any particular difficulty when using a foreign language in an academic setting, neither for written nor for oral communication. In contrast, AI students report feeling insecure about using Norwegian, a non-native language to them, in an academic setting. This is particularly evident when they are supposed to use a VLE, i.e., a tool for publishing written work, in a language that they do not feel they master fully. The interviews reveal that having to publish their own Norwegian texts in the VLE is a source of worry and procrastination for those students. Such a situation is especially problematic in study programmes in early childhood education, as the Norwegian tradition for such studies is to use written reflections and feedback from peers as a central learning method.

“I am still very careful with using it [the VLE] because others can see it [my text], perhaps there are some grammatical errors, aren’t there? This is why I’m very careful.” [Extract of interview with student informants]

Despite their scepticism and reluctance to publishing texts online, the interviewed AI students report that they are willing to improve in that area, and that the only road to improvement is to dare to publish online, regardless of their underlying anxiety.

One interesting finding is that other technologies can be used to support the “publishing, sharing and reflecting” learning approach that is ubiquitous in the Early Childhood Education programme, and those may be more appropriate to international students as they do not require so much text. For example, some interviewees report having used digital storytelling tools within their studies, where pictures play a central role, to document their own learning experiences, and having felt more comfortable using that type of learning and documentation method.

“… Then I think [hard] about what I am going to say, it is also [a kind of] reflection. […] Because it is easier to describe [experiences] using pictures and short sentences. [Extract of interview with student informants]

**Administrative language**

In the VLE used in the studied institution, the Norwegian language pervades the system in two ways. First, the very labelling of the system is designed for Norwegian language. The students report that even when choosing English as a preferred language, several labels are still presented in Norwegian. Second, most of the general information from teachers and administrative staff to students is given in Norwegian, either via the news feature of the VLE or through emails to large groups of students. ENAI student informants describe that they regularly receive large numbers of emails written in the Norwegian language. They report that they first tried to translate those emails into their mother tongue in order to understand their content. However, because it was a time-consuming activity and because most of the information they received was not relevant to them, they ended up ignoring all news and emails, expecting that important information will be given to them orally by the teacher. This, however, was not always the case, and ENAI informants recount having missed important information, such as a visit from scholars from another university, because of the overload of incomprehensible information on the VLE.
Analysis and discussion

In this section, we revisit the data presented in the section above, using some of the core concepts from ANT as a structuring tool.

Inscription and translation

The notion of culture (and cultures) is central when trying to understand the learning experience of international students. In that context, an ANT perspective can be useful, as it offers conceptual tools that allow for a rich investigation of the various aspects of culture. For example, an important element in the broader picture of Norwegian education is the general focus on curriculum, which undergoes a series of transformations involving a number of stakeholders. Study programmes in Norway are not designed by individual academics, but are generally the result of a set national curriculum and an institutional adaptation of this curriculum. The concepts of inscription and translation can be useful to study the relationship between programmes (including curriculum, and expected learning outcome), artefacts (especially language and VLEs), and learning cultures (including work styles, student tasks, behaviour and attitudes from and among teachers and students). Former studies (e.g., Johannesen & Habib, 2010a; Nespor, 1994) have shed light onto the existence of inscriptions, i.e., situations where programmes and technologies are being designed and developed. However, those programmes and technologies are implemented within culturally laden contexts, and those contexts may prompt a number of appropriations or “translations” that are much dependent on the culture they enter. It is, however, important to note that the studies referred to here have been limited to a traditional national context, and have not taken into account the possibility of international participation to the courses.

Several elements from the data point toward the existence of inscriptions with long-lasting effects. For example, the studied institution’s VLE system appears to be unambiguously inscribed with the general institutional strategy to implement computer-supported learning at all levels and in all study programmes. However, although all study programmes have had to adopt the VLE as a platform to support teaching and learning, it is apparent that the VLE has been translated very differently from one faculty to another and from one study programme to another. In particular, in those faculties that had been using the web to support learning long before the implementation of the VLE, there seems to be a significant reluctance toward discarding old, but well-functioning solutions to adopt new, but unfamiliar and, in the eyes of some, dubious ones. The data provide a telling example of a translation of the VLE, when it is used by teachers mostly as a gateway to their own websites, where all the content used to support their teaching actually resides, in a form that is more accessible to international students.

Another institutional inscription is apparent in the way the VLE is used to support the sharing of texts created by the students. For the teachers, who have embraced a socio-cultural perspective on learning, there is a point to encouraging the students to write as much as possible and make their texts available to their peers so as to get their feedback. For the AI students, this inscription has potentially adverse effects, as having to write and publish text is experienced as strenuous and distressing. They do, however, suggest other types of uses of the VLE that would entail a translation of the very idea of the students reflecting on their own learning experience and using the VLE as a medium to share it – which seems to be central to the teaching philosophy at the studied institution. While the teachers seem to rely solely on activities involving writing texts, the students propose using other types of activities, for example digital storytelling, which makes use of pictures in combination with text, often mostly spoken text.

Those findings are in line with the literature that underlines the importance of the international students’ cultural background in their learning practices (e.g., Charlesworth, 2008). They also corroborate the idea that the AI students may carry a cultural baggage with a strong emphasis on perfectionism and high standards of achievement (Nilsson et al., 2008), which may in turn make them uneasy when having to expose their unfinished texts to the scrutiny of their peers.

Enrolment and alignment

The ANT notion of enrolment is central when analysing the interwoven-ness of technology and the social groups that use it or have a stake in it. Because VLEs are comprehensive software supporting both teaching, learning and
administrative activities, they have a strong potential as enrolling actants. Another ANT concept, that of alignment, may be useful to understand the dynamics of inter-relations with and around a VLE. One important actant in this dynamic network seems to be the socio-cultural learning philosophy that argues that learning happens between learners (Engeström, 2001; Vygotsky, 1978). Alignment then happens as technologies are used to champion and to reinforce the underlying learning philosophy, and results that corroborate the hypothesis that such a philosophy is justified are also used as bolstering elements to the network.

As mentioned in the above section, the data show that the ENAI students apply a range of different technologies, included but not restricted to VLEs. In cases where the VLE can offer functionalities that exceed and surpass other technologies, both students and teachers adopt those, such as the VLEs functionality for managing groups. In this case, the teachers, maybe unintentionally, use a certain feature of the VLE technology to enrol student into collaborative practice. In a similar way, teachers often embrace the automated hand-in facilities that make the follow-up of assignments easier and more efficient for themselves and correspondingly achieve some kind of alignment with the student groups – who, according to the data, appreciate having those follow-up routines to keep on track. It is apparent that this process of alignment plays a substantial role in making learning processes effective for both students and teachers, thereby paving the way for a strong focus on learning throughout the educational programme.

Opening black boxes

Institutions of Higher Education have typically developed their routines and systems over time, and chances are that many of them have become blackboxed, thereby becoming accepted and taken for granted by both the students and the teaching and administrative staff. The data indicate that blackboxing also happens within student groups, with little or no involvement from the institution. For example, the idea that social media such as Facebook and communication tools such as Skype are private communication technologies that may be used for communication between students, but that do not lend themselves to communication between students and teachers, is not altogether an obvious concept in the eyes of an analyst.

An example of institutional blackboxing is that of information given via the VLE. The large numbers of emails, news and guidance given by administrative and academic staff are often aimed at all the student groups and are given in Norwegian, without any consideration about the diversity in language capabilities among the receivers of that information. Consequently, the ENAI students abstain from trying to translate all this information and in practice shut down an information channel that has become irrelevant to them.

AI students also seem to be confronted by the existence of a number of black boxes, in particular the notion of writing as a central method for learning. However, because of the AI students’ reluctance to letting others read their texts, the lecturers may be compelled to question their beliefs that a focus on writing will necessarily bring about optimal levels of learning for all students. In that sense, the AI students may indirectly help their teachers to “open the black box” of “writing to learn.” This very questioning process can be useful in raising awareness about difficulties experienced by international students.

Socio-material assemblages as cultural scaffolds

The data from this study illustrate that AI students struggle to produce and publish texts in the Norwegian language as required within their study programme. However, they also report that digital technologies help them reflect on their own learning and put words on those reflections. When informants recount their positive experience with a digital story-telling tool, and how they experience this kind of technology as supportive for producing digital texts, they exemplify how technology and teaching practice can go hand in hand to support learning.

The ENAI students seem to be more technologically mature, and employ a range of technologies with a distinct and clear idea about what technology to use for what purpose. While media such as Facebook and Skype are used in both study-related and social-related affairs, the communication with and feedback from teachers is still expected to happen via email or VLEs. Such data suggest that those media do not lend themselves to be aligned directly with the general learning network, and that they might need to be translated, revisited and perhaps renamed before gaining a
legitimate place into the network. Such data concurs with the existing literature that underlines the difficulties inherent to using social media as a platform for learning (Friesen & Lowe, 2011). Our research also suggests that those difficulties may be heightened in a learning environment including several national cultures, as students belonging to minority cultures may feel insecure about what is appropriate to say and do on a social network, which may in turn become a hurdle in their learning activities, if those are meant to happen on social networks.

**Conclusion and implications**

This study points to the existence of a number of technology-related hurdles for both AI students enrolled in regular programmes, and international students enrolled in exchange or international programmes. In particular, the use of learning technologies such as VLEs may bring about a focus on reading and writing that may put non-native speakers at a disadvantage. However, such technologies also have the potential to equip AI and ENAI students with tools that are empowering to them, for example by facilitating non-text-based forms of documentation.

This study has highlighted the need to differentiate between different types of international students in a culturally diverse Higher Education context. As illustrated in the study, different learning and teaching philosophies may co-exist within the same institution and there may be variations as to how international students relate to those philosophies and appropriate the teaching and learning methods that they meet. One of the implications of the study is that there may be a need for an increased awareness amongst educators not only regarding the principles of their own educational approach but also regarding how their approach may differ from what international students are used to and are equipped to handle.

The findings from this study suggest that educators may better succeed at involving students from cultures that emphasise perfectionism if they allow for activities that do not expose the international students’ weaker points, such as the production of written texts. In that respect, technology-assisted learning, which provides a wider range of activities than traditional educational forms, can be a useful ally to teaching staff working with international students. For example, learning technologies that enable students to use images, sounds and films to express their ideas may play a role in empowering international students by facilitating their participation in dialogic learning practices that may formerly have been restricted to the exchange of written text.

Using ANT has provided us with an alternative view of learning structures such as a study programme where the focus is not on either bringing in or shutting out particular types of learning cultures, but on the negotiations that happen between several types of networks, i.e., the socio-material or cultural networks of the international students and the network that they form together with their study programme and the technology in use at their institution. An ANT-based outlook helps considering the situation not from a normative point of view, but from a more descriptive and analytical viewpoint. The process of mapping out all the networks of aligned interests, without a dogmatic stance, will give them all a place on the broader picture and will help taking seriously issues and difficulties experienced by members of different networks.

One of the main limitations of the research is that it is based on a relatively limited number of cases from a single institution. In addition, the two groups that we have identified for the study (ENAI students and AI students) follow two different types of programmes at different faculties. The EPS programme where the ENAI interviewees are enrolled does not put much emphasis on the production of written text as a method for learning. Although the end-product is generally a written report, much of the learning process seems to happen through face-to-face discussions either within the student groups or between the student groups and their project supervisor. In contrast, the interviewed AI students are all enrolled at the Faculty of Education where activities related to producing texts and providing responses to those texts are at the core of the traditional teaching philosophy.

Further research may include broadening the range of informants to include other faculties, other institutions, and other types of informants such as administrators, technical staff and software developers. Other possible avenues for research could be trying out new teaching processes within for example the realm of pilot projects, and following them closely using action research methods. Examples of such improvements may be using more picture-based narratives as a means to convey students experiences within the realm of their studies, and introducing more technological training and support for non-Western international students. For Western international students, improvements may include reducing the amount of messages in Norwegian language on the VLE.
References


A Blended Mobile Learning Environment for Museum Learning

Huei-Tse Hou¹, Sheng-Yi Wu², Peng-Chun Lin³, Yao-Ting Sung⁴, Jhe-Wei Lin⁵ and Kuo-En Chang⁶*

¹Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan // ²Department of Information Communication, University of Kang Ning, Taiwan // ³Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taiwan // ⁴Department of Educational Psychology and Counseling National Taiwan Normal University, Taiwan // ⁵Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taiwan // ⁶Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taiwan // hthou@mail.ntust.edu.tw // digschool@gmail.com // pclin@sce.pccu.edu.tw // sungtc@ntnu.edu.tw // ljwlgw@gmail.com // kchang@ntnu.edu.tw

*Corresponding author

(Submitted November 21, 2012; Revised June 23, 2013; Accepted August 5, 2013)

ABSTRACT
The use of mobile devices for informal learning has gained attention over recent years. Museum learning is also regarded as an important research topic in the field of informal learning. This study explored a blended mobile museum learning environment (BMMLE). Moreover, this study applied three blended museum learning modes: (a) the traditional museum visit accompanied by a learning website, (b) paper-based learning sheets used during museum visits accompanied by a learning website, and (c) an interactive mobile learning system used during museum visits accompanied by a learning website (i.e., BMMLE). Furthermore, the study explored the learning process through the use of each mode by museum visitors and empirically examined the differences between the learning performances and behavioral patterns of visitors. Study participants included 58 college students. A performance analysis, a behavior analysis of learners’ participation on the website and a sequential analysis of the videotaped behaviors of visiting participants were conducted. The findings showed that the BMMLE proposed in this study may enable visitors to focus on the interactions between on-site exhibits and mobile learning systems and that the BMMLE may also extend the interaction period between on-site learning and the learning website, thus facilitating the implementation of the museum’s learning activities.

Keywords
Informal learning, Blended learning, Mobile learning, Behavior analysis, Museum learning

Introduction
Applying mobile technologies to museum learning

The use of mobile devices for informal learning has aroused attention from researchers in the field of educational technology (Collins, Mulholland, & Zdrahal, 2009; Sung, Chang, Hou, & Chen, 2010b). With the development of Internet and mobile technology, research on mobile learning has gradually gained more and more attention (Kinshuk et al., 2013). Wu et al. (2012) conducted a meta-analysis of 164 studies in mobile learning which were undertook during a period of 10 years and discovered that most topics of the mobile learning studies were currently based on the learning effectiveness and the evaluation of systems. However, in terms of the research on mobile devices assisted informal learning, besides evaluating mobile systems and the learning effectiveness, the issue of how to propose a suitable mobile learning mode and further analyze the learners’ behavioral patterns based on the informal learning activity in a specific situation (e.g., visiting time, visiting place, and other learning resources) is worth a deep exploration. Museum learning is an important research topic in the field of informal learning. Currently, research on technology-assisted museum learning focuses on developing a learner-centered method and applying technology to assist learners in exploring and learning in a museum (Wishart & Triggs, 2010). There have been numerous studies of the advantages of using mobile devices as supplemental learning aids in a museum setting (e.g., Sung et al., 2010a; 2010b; Vavoula et al., 2009; Collins, Mulholland, & Zdrahal, 2009). For example, Collins et al. (2009) used the text-messaging functionality of cell phones to assist learners in observing exhibits. Sung et al. (2010b) developed a mobile guide system that provided historical narratives as backgrounds for the exhibits.
However, some studies have also found that using mobile devices to support museum learning has limitations. First, in terms of the interactions between learners and exhibits, the use of mobile devices easily compels learners to spend more time familiarizing themselves with the device interface or operating the device during the museum visit, constraining deep interactions between learners and exhibits and preventing learners from focusing on learning through the process of continuously observing the exhibits (e.g., Mantyjarvi et al., 2006; Semper & Spasojevic 2002; Hsi 2003; Klopfer et al., 2005; Reynolds et al., 2010). Based on the limitations of interactions between learners and exhibits, applying the theoretical foundations of learner-centered problem-based learning (PBL) to allow learners to solve problems or tasks only through searching for information, exploring and analyzing exhibits can support interactions between learners and exhibits and the construction of knowledge (Sung et al., 2010a). Many researchers have proposed the use of a mobile guide combined with problem solving tasks (e.g., Kwak 2004; Klopfer et al., 2005; Sung et al., 2010a; Vavoula et al., 2009). For example, Chicago History Museum’s mobile guide system provided treasure-hunt tasks for learners to use while solving problems (Kwak, 2004); Sung et al. (2010a) reported that the guide mode combined with a problem-solving strategy helped promote interactions between peers and between learners and exhibits. However, even though problem-solving tasks were used in this case, the lack of time a learner is given to familiarize himself or herself with the device interface constrains the learner’s concentration on and interaction with the exhibits. In terms of knowledge acquisition from on-site exhibits, when learners use mobile devices in a physical environment, they need to pay attention to the information on the device and the physical objects simultaneously, which may cause cognitive overload (Liu, Lin, Tsai & Pass, 2012). During museum visits, learners may fail to understand the exhibits deeply because of the abundance of exhibits and time limitations leading to information overload (Bitgood, 2009). Even learners using mobile guide systems combined with problem-solving tasks, which encourage interactions between learners and on-site exhibits, still find it difficult to familiarize themselves with the device interface and absorb the information on the devices and in the exhibits under the pressure of a time-limited museum visit. Thus, learners fail to concentrate on and deeply understand the exhibits. Therefore, beyond providing learners with problem-solving tasks, the design of a mobile museum learning environment that offers time flexibility is a key research topic. Such an environment would provide learners with more time to familiarize themselves with the device interface and to internalize the information on the devices and in the exhibits, which in turn is expected to help learners concentrate on and thoroughly understand on-site exhibits.

**Blended mobile museum learning environment (BMMLE)**

Falk and Dierking (2000) proposed a contextual model of learning (CML) that explains the factors that affect museum learning. The three contexts of the CML are personal context, socio-cultural context, and physical context. Visitors’ learning experiences in museums are the process and result of the interactions among these three contexts. The abovementioned guided activity combined with an instructional design based on problem-solving tasks can provide appropriate task situations that build on learners’ background knowledge and the interactive characteristics of effective learning to provide the required personal and socio-cultural contexts. Additionally, such a guided activity involving problem-solving tasks promotes interaction with the exhibits and directs the learners' attention toward the physical context of the museum exhibits (e.g., Sung et al., 2010a). Therefore, this study would adopt problem-solving tasks as a basis for the design of an instructional strategy in an instructional activity. Dierking (2002) added a ‘time context’ to the CML and noted that learning was the result of the mutual influences of the three aforementioned contexts coupled with the requirement of adequate time. Adding this ‘time context’ showed that museum learning requires a substantial amount of time for adequate adaptation and internalization of the knowledge presented.

From the perspective of the CML, to effectively improve learning outcomes, the design of a blended museum learning environment that integrates physical museum learning with virtual learning resources may be an appropriate approach for further research. Blended learning refers to the integration of traditional physical learning methods with virtual network technologies. Blended learning employs virtual teaching activities that complement and enrich physical teaching (Ko & Rossen, 2001). Vaughan (2007) found that students in a blended learning experience reported that its learning mode allowed them to use their time more freely and promoted their learning. Teachers reported that a blended learning activity offered a greater degree of flexibility in the instructional environment. Museum learning, which is often limited by students’ short visits, is regarded as a short-term physical teaching activity. However, this time constraint may restrict the effects of learning on learners’ cognition (Bitgood, 2009). Therefore, this study explored an instructional design in which a mobile-guided activity was integrated with the
blended learning mode of problem-solving tasks. In other words, if physical museum learning can be supplemented with virtual learning resources to form a blended learning environment, such a design could benefit students by allowing them more time for learning and a more flexible learning environment.

The interactive blended mobile museum learning environment (BMMLE) proposed in this study refers to an environment in which students can use mobile devices in conjunction with the physical learning experiences in museums and subsequently review and study what they have learned, via a learning resource website, at home. The website includes learning resources and additional information tied to the museum's exhibits. Mobile devices can be equipped with problem-solving tasks that enable students to conduct observations under the devices’ guidance while in physical proximity to the exhibits, search for relevant information on the learning website at home, and finally return to the exhibits to observe and complete the tasks using their mobile devices.

Research focus and purpose

The current study is an extension of earlier research by our team (Sung et al., 2010a). An earlier study (Sung et al., 2010a) explored a short time-span on-site instructional activity in which students visited a museum exhibition. The focus of the current study is the use of BMMLE to extend students’ learning processes and depth of learning with the aid of a virtual learning website, and this study also explores the effectiveness of BMMLE and students’ behavioral patterns through empirical research.

A new mode of blended museum learning can be created through the combination of on-site visiting and a virtual learning website. As there are various blended museum learning environments that integrate on-site museum learning with a virtual learning website, to compare and fully understand the impact, performance, characteristics, and limitations of each blended museum learning environment, this study compares the effectiveness and performance of the following three blended museum learning modes: (a) the traditional museum visit combined with a learning website, (b) the use of a paper-based learning sheet during the museum visit combined with a learning website, and (c) the use of an interactive mobile learning system during the museum visit combined with a learning website (i.e., BMMLE). Accordingly, as part of this study, a learning sheet, interactive mobile learning application software with problem-solving tasks and a learning website were developed. Students were divided into three groups for empirical analysis: one group using traditional learning with a learning website (hereafter referred to as the traditional learning group), one group using the paper-based learning sheet with a learning website (hereafter the paper-based learning group), and one group using mobile learning with a learning website (hereafter the mobile learning group).

To perform a multi-aspect evaluation, this study used a pretest and a posttest to measure differences in participants’ learning performances. This study also measured students’ participation on the learning website. Finally, to understand the students’ behavioral patterns while performing physical museum learning activities and to determine the possible causes for the observed differences in learning performance and participation on the learning website, this study further examined, via video analysis, the sequential behavioral patterns (Bakeman & Gottman, 1997; Hou & Wu, 2011) of students in each of the three groups while they were engaged in the museum learning activities. The primary contribution of this paper is the researchers’ proposal of the instructional mechanism of a blended museum mobile guide and the empirical analysis of the features and limitations of students’ performance, participation and behavioral patterns while using various blended guide modes.

The objectives of this study were as follows:

- To compare the differences in learning performances between the traditional learning group, the paper-based learning group, and the mobile learning group.
- To compare the differences in participation on the learning website between the traditional learning group, the paper-based learning group, and the mobile learning group.
- To compare the differences in behavioral patterns between participants in the traditional learning group, the paper-based learning group, and the mobile learning group during their physical museum visits.
Research design

This study used a qualitative-quantitative mixed research method (Johnson & Onwuegbuzie, 2004) to observe visiting participants’ behavior and to obtain qualitative data (the video-taped records of participants’ learning behavior) and quantitative data (participants’ frequency of participation on the learning website and their pretest and posttest scores). In addition to a quantified analysis of the participants’ learning effectiveness, participants’ video footage was also encoded for qualitative analysis to understand the differences in participants’ behavioral patterns among the three groups.

Participants

Study participants included 58 college students with an average age of 20. Participants were randomly grouped, with 20 in the traditional learning group, 19 in the paper-based learning group, and 19 in the mobile learning group. All participants were scheduled to attend a museum excursion and learn about Tang tricolor pottery, the flagship ceramic ware of the Chinese Tang Dynasty, on display at the National Museum of History in Taiwan.

Research tools

Pretest and posttest

The pretest consisted of twelve questions intended to measure the participants’ knowledge of the exhibition prior to their visit; these questions measured the participants' basic knowledge of tricolor pottery of the Chinese Tang Dynasty. The posttest consisted of 17 questions intended to measure the participants’ knowledge growth after visiting the exhibition; these questions concerned the features and related knowledge about the Tang tricolor pottery exhibits at the museum. Ten multiple-choice questions and two essay questions were included in the pretest, and 14 multiple-choice questions and three essay questions were included in the post-test. Each of the multiple-choice questions was assigned three points, and each of the essay questions was assigned ten points. The pretest and posttest were each assigned a maximum score of 100. All questions were taken from the explanatory introductory documents about the exhibition supplied by the museum. Visitors could access the documents and other related information from the learning website provided in this study.

Learning sheet

The paper-based learning group was handed a learning sheet during the experiment that included 14 questions on the Tang tricolor pottery exhibition items and related facts. The source and scope of the questions were derived from those of the pretest and posttest.

Micro-cameras

Micro-cameras were used to record the process of the museum visits. The micro-cameras had built-in lenses that recorded the process and the behavior of the participants and stored images on a micro SD card. Each participant wore a headband with a micro-camera attached, which, in turn, facilitated the analysis of the participants’ learning behavior.

Mobile learning system

The mobile learning system developed by this study provided the time indicator, exhibition location map, and a series of task-related questions to guide the visitors as they explored the exhibits in the exhibition. Much of the knowledge required to complete the task came from the exhibits and their explanatory introduction, as well as the
video or multimedia resources in the exhibition. Some of the required knowledge was not available onsite and could only be attained from the virtual learning website. For a feasible comparison, the paper-based learning group and mobile learning group shared the same 14 questions on the learning sheet.

**Learning website**

During the learning activity, the participants from the three groups were allowed to access the museum’s learning website at home or school. The content of the learning website included a variety of information, such as the historical development, applications or purposes of the exhibited items, production related data, decorative techniques, and work classifications; the learning website included an audio guide function. The scope of the information was equivalent across the pretest, post-test, and learning sheet. This study focused on a comparison of the three modes of blended learning in the museum. Therefore, all three groups of students were asked to access a learning website. Additionally, the three groups of students were able to access explanatory documents provided by the museum on the learning website to ensure the comparability among the three groups.

To attain the functions and features of the blended mobile learning system that links the mobile devices to the virtual learning website, this study offered the bonus function of “task log” on the learning website. The “task log” function was only provided to the mobile learning group. This function enabled the students in this group to review the contents of tasks, the questions that they have answered, and the answering records, which allowed students to review the test items answered incorrectly by themselves and to perform additional self-directed learning and collect more related information for further on-site visiting and exploration. The task log can be used as an auxiliary function for information gathering and self-regulating and a link between the mobile device and the learning website in BMMLE.

**Behavior coding scheme**

To explore participants’ behavior during the visiting process, the video footages recorded by the micro-cameras had to be coded. This study employed the “coding scheme of museum-learning behavior,” which has been used previously in the empirical analysis of museum-visiting behavior (Sung et al., 2010a). As this study did not interfere with participants’ discussions, after the codes of the discussion behavior were removed from the initial coding table, the remaining codes contained a total of six types of behavior, as shown in Table 1.

<table>
<thead>
<tr>
<th>Code</th>
<th>Visitor’s behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>Walk and look around in the exhibition.</td>
</tr>
<tr>
<td>P</td>
<td>Operate/look at the mobile devices or read/fill out the learning sheet.</td>
</tr>
<tr>
<td>S</td>
<td>Stop to look and observe a certain exhibit.</td>
</tr>
<tr>
<td>R</td>
<td>Make reference to the display board or media describing the exhibits in the exhibition.</td>
</tr>
<tr>
<td>H</td>
<td>Seek help from on-site staff for questions about the visiting activity.</td>
</tr>
<tr>
<td>O</td>
<td>Other actions irrelevant to the above-mentioned learning behavior.</td>
</tr>
</tbody>
</table>

**Procedures**

Set as a quasi-experimental design, in this study, the independent variables were groups with different treatments (i.e., traditional learning group, paper-based learning group, and mobile learning group), the dependent variables were the post-test scores and learning website participation, and the covariates were the pretest scores.

This experiment took place during the exhibition of Tang tricolor pottery at the National Museum of History over a two-week period. Because this study emphasizes the exploration of blended learning and museum learning in an informal learning context, it focused on observing and exploring learners’ blended learning behaviors in a natural situation, to avoid constraining students to a certain type of learning in an (unnatural) experimental environment at a specific time. The control variable in this study was the total amount of time provided for students’ learning (i.e., two weeks); also, each group had the same learning theme. Participants in the three groups could freely choose any time
slots to conduct their museum visit (each slot was between 30 and 60 minutes) during the two-week period depending on their own blended learning preference. Participants from each group took the pretest 10 minutes prior to their first visit, followed by 5 minutes of explanation and instruction by the staff, after which the experiment ensued. The mobile learning group received additional instructions on the mobile learning system (installed in HTC-HD2 smart phones).

During the two-week period of learning and the onsite visit, the researchers did not provide the learners in the traditional learning group extra resources, whereas the learners in the paper-based learning group could bring the learning sheet containing unsolved problems, which the researchers had designed beforehand; also, the learners in mobile learning group were provided the mobile guide system while visiting the museum. During their museum visits, each participant in the three groups wore a personal micro-camera during the entire process that recorded his or her visit. During the two weeks, besides the period of onsite visits (e.g., at home or school), all participants could use the learning website for information searching and learning. After the two weeks, all participants then took the post-test at a joint session at the completion of the experiment.

**Data analysis**

This study used the pretest as covariates and applied ANCOVA to examine the potential differences in learning performance. To explore students’ participation in this extended learning website activity, descriptive statistics were used to analyze the participants’ usage ratio and their average viewing time.

The lag sequential analysis (Bakeman & Gottman, 1997; Hou, 2012) was used to analyze the participants’ behavioral patterns during their learning activities at the museum. The study called on two raters with professional backgrounds in information and computer education to encode the fully documented video-camera footage of a randomly chosen participant. The video data of the sample participant was coded based on Table 1 with every five seconds as a unit, thereby accounting for a total of 351 units. The coding results revealed a kappa coefficient measured at .80, thus indicating a high inter-rater reliability between the two raters. This was followed by the main rater’s coding of the entire video data, in which each participant’s video was chronologically coded, and each behavior event was treated as a unit. The researcher then conducted the lag sequential analysis on the coding results; that is, after the above-described codes were chronologically arranged, the frequency transfer matrix was calculated. Finally, after the transfer matrix was completed, a series of significant sequences was derived, and the sequential transfer diagrams were drawn, which illustrated the sequential behavioral patterns of the museum visitors from all three groups.

**Results and discussion**

**Participants’ learning performance analysis**

This study evaluated the average and the standard deviation of participants’ pretest and posttest scores, which are illustrated in Table 2. The table shows that when compared with the traditional learning group and paper-based learning group, the posttest results of the mobile learning group are higher.

<table>
<thead>
<tr>
<th>Group</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Paper-based learning group (n = 19)</td>
<td>48.94(13.89)</td>
<td>57.32(10.73)</td>
</tr>
<tr>
<td>Traditional learning group (n = 20)</td>
<td>46.00(12.98)</td>
<td>49.45(14.63)</td>
</tr>
<tr>
<td>Mobile learning group (n = 19)</td>
<td>43.05(13.32)</td>
<td>63.47(12.13)</td>
</tr>
</tbody>
</table>

This study further examined the differences among the three groups via ANCOVA, using covariates as the pretest scores and dependent variables as the posttest scores. Before ANCOVA, the results from the homogeneity of within-class regression coefficient test showed that the slope of the three sets of data could be regarded as the same, which coincides with the basic assumptions of homogeneity of within-class regression coefficient. ANCOVA could then be
conducted ($F = 1.649$, $p = .202 > .05$). After the pretest factors were excluded, the analysis of covariance demonstrated a significant level among the three groups ($F = 5.97$, $p = .005 < .05$).

After the data of the three groups were compared and analyzed, the results showed that the mobile learning group performed significantly better than the traditional learning group ($p = .001 < .05$), while there was no significant difference found when comparing either the paper-based learning group with the traditional learning groups ($p = .064 > .05$) or the mobile learning group with the paper-based learning group ($p = .137 > .05$).

**Analysis of virtual learning website participation**

Regarding the blended learning mode, the participants’ usage statistics of the virtual learning website during the given two weeks are as shown in Table 3. The mobile learning group usage ratio is at 78.95%, and the average usage time is 23.47 minutes; the paper-based learning group usage ratio is 57.89%, and the average usage time is 14 minutes; the traditional learning group usage ratio is 25%, and the average usage time is 8.8 minutes. This indicates that the participation ratio of the mobile learning group is better than that of the other two groups (Table 3). It may also indicate that the mobile learning group equipped with mobile devices and operating on the task-based teaching model experienced success in participating in the blended virtual learning.

**Table 3. Statistical table of backend website learning**

<table>
<thead>
<tr>
<th></th>
<th>Mobile Learning Group</th>
<th>Paper-based Learning Group</th>
<th>Traditional Learning Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage ratio</td>
<td>78.95%</td>
<td>57.89%</td>
<td>25%</td>
</tr>
<tr>
<td>Average usage time</td>
<td>23.47 mins</td>
<td>14 mins</td>
<td>8.8 mins</td>
</tr>
</tbody>
</table>

**Participants’ behavioral pattern analysis**

This study further applied the lag sequential analysis on behavior to explore the behavioral patterns of the participants during their visiting experience, from which visualized pattern diagrams were presented to explore the differences in the patterns from each group. Through lag sequential analysis, the adjusted residuals table of behavior from each group is shown in Table 4, Table 5 and Table 6. The rows in the tables are the initial behavior, and the columns are the behavior immediately following. When the Z-values in the tables are greater than 1.96, it indicates that the sequences have reached statistical significance ($p < 0.05$). These significant sequences can then be deduced into behavior transfer diagrams (i.e., Figures 1, 2, and 3). The arrows in the figures indicate the direction of the behavioral sequence (for example, the arrow from Behavior A to Behavior B indicates that during the process, the behavioral continuity of ‘A to B’ reached statistical significance during the overall activity).

**Table 4. Adjusted residuals table (Traditional learning group)**

<table>
<thead>
<tr>
<th>Code</th>
<th>W</th>
<th>S</th>
<th>R</th>
<th>H</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>-13.82</td>
<td>3.46</td>
<td>7.24*</td>
<td>-0.57</td>
<td>5.84</td>
</tr>
<tr>
<td>S</td>
<td>9.75*</td>
<td>-12.70</td>
<td>4.91*</td>
<td>-1.69</td>
<td>-2.30</td>
</tr>
<tr>
<td>R</td>
<td>3.31*</td>
<td>10.60*</td>
<td>-11.96</td>
<td>-2.54</td>
<td>-2.09</td>
</tr>
<tr>
<td>H</td>
<td>-0.94</td>
<td>0.62</td>
<td>0.67</td>
<td>-0.23</td>
<td>-0.42</td>
</tr>
<tr>
<td>O</td>
<td>2.12*</td>
<td>-2.30</td>
<td>-1.00</td>
<td>8.04*</td>
<td>-2.64</td>
</tr>
</tbody>
</table>

*p < 0.05.

**Table 5. Adjusted residuals table (Paper-based learning group)**

<table>
<thead>
<tr>
<th>Code</th>
<th>W</th>
<th>P</th>
<th>S</th>
<th>R</th>
<th>H</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>-11.98</td>
<td>-2.94</td>
<td>4.91*</td>
<td>8.59*</td>
<td>0.31</td>
<td>5.44*</td>
</tr>
<tr>
<td>P</td>
<td>3.94*</td>
<td>-19.92</td>
<td>4.16*</td>
<td>13.39*</td>
<td>0.82</td>
<td>-0.99</td>
</tr>
<tr>
<td>S</td>
<td>4.55*</td>
<td>3.63*</td>
<td>-6.97</td>
<td>-1.96</td>
<td>-0.50</td>
<td>-1.76</td>
</tr>
<tr>
<td>R</td>
<td>3.78*</td>
<td>19.80*</td>
<td>-2.97</td>
<td>-19.82</td>
<td>-2.56</td>
<td>-2.47</td>
</tr>
<tr>
<td>H</td>
<td>1.93</td>
<td>-0.68</td>
<td>-0.40</td>
<td>-0.68</td>
<td>-0.10</td>
<td>-0.10</td>
</tr>
<tr>
<td>O</td>
<td>-2.27</td>
<td>0.51</td>
<td>-1.76</td>
<td>1.01</td>
<td>8.53*</td>
<td>-0.44</td>
</tr>
</tbody>
</table>

*p < 0.05.
Table 6. Adjusted residuals table (Mobile learning group)

<table>
<thead>
<tr>
<th>Code</th>
<th>W</th>
<th>P</th>
<th>S</th>
<th>R</th>
<th>H</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>-10.31</td>
<td>-0.11</td>
<td>3.39</td>
<td>4.69*</td>
<td>5.17*</td>
<td>0.23</td>
</tr>
<tr>
<td>P</td>
<td>10.44*</td>
<td>-22.32</td>
<td>4.48*</td>
<td>10.92*</td>
<td>0.02</td>
<td>0.61</td>
</tr>
<tr>
<td>S</td>
<td>0.72</td>
<td>5.90*</td>
<td>-7.42</td>
<td>-0.44</td>
<td>-1.63</td>
<td>0.54</td>
</tr>
<tr>
<td>R</td>
<td>-3.16</td>
<td>19.17*</td>
<td>-1.35</td>
<td>-15.26</td>
<td>-3.46</td>
<td>-1.22</td>
</tr>
<tr>
<td>H</td>
<td>1.93</td>
<td>0.78</td>
<td>-1.63</td>
<td>-0.91</td>
<td>-0.85</td>
<td>-0.25</td>
</tr>
<tr>
<td>O</td>
<td>1.53</td>
<td>-0.42</td>
<td>-0.85</td>
<td>-1.22</td>
<td>2.78*</td>
<td>-0.10</td>
</tr>
</tbody>
</table>

*p < 0.05.

Figure 1. Behavioral transfer diagram (traditional learning group)

As shown in Figure 1, the behavior of the participants from the traditional learning group and the three types of behavior, walking (W), exhibit viewing (S), and referencing the display board or media describing the exhibits in the exhibition (R), share a significant sequential relationship. Accordingly, this represents the basic museum-visiting behavioral pattern (i.e., S→W, W→S, S→R, R→S, R→W and W→R). It is also evident that the exhibit viewing (S) and referencing the display board or media (R) share a bi-directional sequential relationship (i.e., S→R, R→S). In addition, this study also found that when participants walk, there is a significant behavioral sequence linking the walking to other learning-unrelated behaviors (i.e., W→O). This reveals the potential restrictions embedded in the traditional museum learning system with respect to sustaining visitors’ focus.

Figure 2. Behavioral transfer diagram (paper-based learning group)

As presented in Figure 2, the four types of behavior displayed by participants in the paper-based learning group, namely, walking (W), learning sheet use (P), viewing of exhibits (S), and referencing the display board or media describing the exhibits in the exhibition (R), all shared a close sequential relationship. (i.e., S→W, W→S, P→R,
The significant behavioral sequence revealed a close sequential relationship between learning sheet use (P), exhibit viewing (S), and referencing the display board or media describing the exhibits in the exhibition (R). (This showed that the guidance offered by the problem-based learning sheet (P) may encourage learners to explore more exhibits (P→S, S→P) and exhibition-related information (P→R, R→P). In addition, the study also found that the occurrences of walking that led to other learning-unrelated behavior (W→O) corresponded to the behavior exhibited by participants from the traditional learning group.

![Behavioral transfer diagram (mobile learning group)](image)

Finally, as shown in Figure 3, the four types of behavior exhibited by the participants of the mobile learning group, namely, walking (W), learning sheet use (P), exhibit viewing (S), and referencing the display board or media describing the exhibits in the exhibition (R), shared a close sequential relationship. (i.e., W→S, P→R, R→P, P→S, S→P, P→W, W→R). The patterns also indicate that participants from the mobile learning group were benefited by the intermediary interactive function of the mobile systems (P) when viewing exhibits (P→S, S→P) and making reference to the display board or media describing the exhibits in the exhibition (P→R, R→P), thus making mobile learning systems practical tools for exhibit viewing and museum learning. The aforementioned outcome coincided with that of the paper-based learning group, indicating that the learning sheet and mobile learning system with the problem tasks may help participants in the museum learning experience. Furthermore, with the help of mobile learning systems, occurrences of walking that led to other learning-unrelated behaviors (i.e., W→O) did not reach significant behavioral sequence, which was in contrast to the behavioral results of the traditional learning group and paper-based learning group. This suggested that mobile devices may prevent the occurrences of walking leading to other learning-unrelated behavior and may help improve concentration when compared to the other two groups. Furthermore, after the participants from the mobile learning group viewed the exhibits (S) or referenced the exhibition-related information (R), significant sequential continuity was immediately linked to mobile devices (S→P, R→P), and not to other learning-unrelated behavior (O) or walking"(W), indicating that the mobile learning system may qualify as an interactive intermediary. On the other hand, both the paper-based learning group and the traditional learning group showed significant sequential behavior to engage in walking after viewing the exhibits (S→W) and exhibition-related information (R→W). This finding indicated that compared to the traditional or paper-based learning modes, mobile learning systems made it easier for visitors to focus their attention on museum learning.

**Discussion**

According to the contextual model of learning (CML), museum learning refers to the interaction between personal context, physical context, and environmental context, and the results of change over time, the deliberations, and the accumulations (Dierking, 2002). This study proposed three blended learning modes and offered a comparison of those modes. From the analyses of the learning performance tests presented herein, the mobile learning group showed significantly better results than the traditional learning group, while in contrast, the learning performance between the mobile and paper-based learning groups and the paper-based and traditional learning groups showed no significant differences. One possible reason may lie in the interactive mechanism between the problem-oriented
mobile learning group with the extra ‘task log’ function in the virtual website that connects the physical visiting and virtual learning context, which resulted in a higher learning performance among the mobile learning group and a significantly better performance than that of the traditional learning group. However, the fact that the paper-based learning group was also guided by the problem-oriented learning sheets may cause a lack of significant difference in the performance between the mobile learning group and the paper-based learning group. This conclusion not only corresponds to the advantages of the interactive strategy (such as “problem-solving tasks”) employed by past museum learning research (Klopfer et al., 2005; Sung et al., 2010a; Vavoula et al., 2009), this study further determines that the problem-solving interactive strategy within blended environments is one of the important key factors that may affect learning performance. However, unlike the mobile learning group that significantly surpassed the traditional learning group in performance, the effectiveness of the paper-based learning was clearly insufficient. Therefore, the combination of the highly portable mobile devices with the task log function suggests that the mobile learning is the better blended learning mode. The reason may well be related to its ‘task log’ mechanism that effectively assumes the intermediary role of between the two learning modes (physical and virtual) and facilitating visitors to both retrospective and introspective with respect to the learned content (Hsi, 2002; Laurillau & Paterno, 2004).

Regarding participants’ usage of the virtual learning website, the usage ratio and average time usage of the mobile learning group were higher than the other two groups, indicating that the website participation of the mobile learning group was better than that of the other two groups. This also corroborated the aforementioned results of the performance analysis; that is, the problem-based and mobile device-supported teaching mode applied by the mobile learning group yielded greater effectiveness in the virtual participation of the blended learning environment and higher usage rate of the learning website compared to the other two learning modes.

The behavioral diagram of each group revealed that both the paper-based and traditional learning groups were more prone to divert attention to non-learning behaviors during the visiting process. In addition, the diagram also showed that mobile learning devices and problem-based learning sheets effectively acted as intermediary scaffolding, linking learners’ exploratory behavior of viewing exhibits and making references to the display board or media describing the exhibits in the exhibition, thus providing valid evidence for the afore-mentioned analysis results of learning performance. In other words, the intermediary mobile devices and learning sheet allowed learners to better integrate the learning activities of the exhibition. When comparing mobile learning devices to the learning sheet, the focus level of the physical museum visit and the participation level on the virtual learning website exhibited from the behavioral patterns and web participation of the mobile learning group were better than those of the other two groups. This means that the blended museum learning environment (BMMLE) proposed by this study that combined mobile learning systems and a virtual website proved to be a superior museum learning system.

In addition to the above findings, this study also made the following recommendations from the summary of the content analysis of after-activity interviews with learners. The mobile learning system and problem-based tasks proposed by this study can efficiently guide students’ learning, as the appropriate problem-based strategy increases student interest and promotes learning (Barrow, 1996; Sung et al., 2010a). However, most of the interview feedback from the students indicated that the assigned tasks were not at all difficult. Therefore, it is suggested that the tasks be more challenging, perhaps by providing semi-structured problems or ill-structured problems that stimulate a higher level of cognitive thinking (Ertmer et al., 2008; Laxman, 2010). For example, in the task design regarding the production process of Tang tricolor pottery, questions can be designed in such a way that students are encouraged to further explore the style and characteristics of that era through the cognitive skills of comparison and analysis.

Conclusions and suggestions

This study applied the three blended learning modes to explore the learners’ learning effectiveness and the behavioral patterns. The results reveal that the knowledge gained by the mobile learning group was significantly greater than that by the traditional learning group. The system record indicates that learner participation of the mobile learning group was greater than that of the other two groups. The mobile learning mode coupled with a higher interactive mechanism allows participants to willingly spend more time learning on the website. The lag sequential pattern analysis revealed a strong similarity between the paper-based learning group and the mobile learning group, indicating that applying a mobile learning system or a learning sheet as an intermediary tool benefited the visiting process. In other words, problem-based tasks provided by a mobile learning system or a
learning sheet were similar to scaffolding which timely supported the visitors. Furthermore, the attention of the participants on learning from the mobile learning group was greater than that of the other two groups, suggesting that the BMMLE proposed in this study may allow visitors to focus on the interaction between onsite exhibits and mobile learning systems, and it may also extend the interaction period between onsite learning and the learning website, thus facilitating the implementation of the museum learning activities.

This study has certain limitations. First, the limitations regarding the limited amount of equipments (i.e., mobile devices) made a large-scale research impractical. Furthermore, the participants in this study were restricted to university students, while museum visitors usually include people of all ages. It is recommended, therefore, that subsequent research on the blended museum learning system include different age groups and present the content from a more diversified perspective.

Acknowledgments

This research was supported by the projects from the National Science Council, Republic of China, under contract number NSC-102-2511-S-011-001-MY3, NSC-100-2628-S-011-001-MY4, NSC-99-2511-S-011-007-MY3, the "Aim for the Top University Project" of National Taiwan Normal University (NTNU), and the International Research-Intensive Center of Excellence Program of NTNU and NSC, Taiwan, under contract number NSC-102-2911-I-003-301.

References


A Study of the Design and Implementation of the ASR-based iCASL System with Corrective Feedback to Facilitate English Learning

Yi-Hsuan Wang and Shelley Shwu-Ching Young*

Institute of Information Systems and Applications, National TsingHua University, Taiwan, R.O.C. // annyw12345@hotmail.com // scy@mx.nthu.edu.tw

*Corresponding author

(Submitted November 26, 2012; Revised March 22, 2013; Accepted July 1, 2013)

ABSTRACT

The purpose of the study is to explore and describe how to implement a pedagogical ASR-based intelligent computer-assisted speaking learning (iCASL) system to support adult learners with a private, flexible and individual learning environment to practice English pronunciation. The iCASL system integrates multiple levels of corrective feedback and allows learners to practice English speaking with immediate diagnosis of their utterances. The information of the multiple levels of feedback includes a list of words that are pronounced accurately and inaccurately, while audio recasting and demonstration of model utterances are adopted to facilitate speaking. To evaluate the effectiveness of the system, a total of 38 adults from Taiwan participated in this experiment, divided into an experimental and a control group. The control group practiced English speaking using the single-level-feedback system in which only a speaking score and waveform diagram was presented as feedback, while the experimental group was given the three-level-feedback iCASL system integrating implicit and explicit elements in the feedback presentation. The empirical evaluation reveals that there were significant differences between the pre-test and post-test speaking scores only for the learners with three-level-feedback iCASL system. The study is in accordance with the previous research that indicates specific feedback is needed and helpful for further enhancing their English pronunciation. Within the three-level feedback, the second level, which provides an explicit textual description with immediate recasting of the learners’ utterances, is the most useful feedback for self-correcting learning. The learners received simultaneous dual information from the audio and visual presentation and hence had better potential to improve their English speaking than the learners who were provided with single modality feedback. Moreover, the learners were satisfied with the iCASL system for self-paced learning and they mentioned that the path suggestion function was especially beneficial because it helped them to manage their learning. The adult learners used the accumulated fragmented time of the informal learning opportunities to achieve improvement in their spoken language. Some research issues and suggestions are also presented for future reference.

Keywords

CALL system design, Self-paced speaking learning, Multiple-level feedback, Speech recognition

Introduction

English is an important subject in formal education in Taiwan. Students learn English as a subject at school from third grade on, and students in junior and senior high school spend about 5 to 6 hours per week learning English (Su, 2006). Furthermore, some colleges even require students to take at least one English course when they are freshmen and to pass an English examination before they graduate. On the other hand, few adults have the chance to keep learning English after graduating from school, especially those who are not English majors. Adult learners may have insufficient exposure to English because of their time limitations and types of work. Besides, it is a common phenomenon in Taiwan that second language learners are more willing to read and write in English than to speak the language. One of the main reasons may be the limited learning materials that students are exposed to, and the cultural belief that passing exams, which are mostly paper and pen tests, is the most important goal of being a student. Moreover, the factors of large class size and limited class time make it difficult for teachers to provide every learner with equal opportunities to practice English speaking in regular classroom sessions. As a result, a student may have learned English for a number of years and yet still find it difficult to speak in basic English.

On the other hand, the advancement of computer assisted language learning (CALL) systems has facilitated learning and teaching. The repeated training of skills that computers provide increases the opportunities for exposure to the target language, while computer-based practice models reduce learners’ language learning anxiety in front of instructors or peers (Campbell, 2003; Arslan & Sahin-Kizil, 2010). Further, automatic speech recognition (ASR)
based CALL systems provide learners with an integrated environment in which they obtain immediate evaluation of their English utterances, and allow them to practice at any time to enhance their speaking (Chiu, Liou, & Yeh, 2007; Chen, 2011). However, many issues in the feedback design of ASR-based CALL programs require further research. Related research has shown that in speaking-practice CALL programs, learners tend to produce more accurate utterances when they are provided with corrective feedback, rather than just opportunities to speak (Mackey & Goo, 2007; Chwo, 2012). The importance of providing learners with corrective feedback while using CALL programs has been recognized by several studies (Neri, Micha, Gerosaa, & Giuliani, 2008; Chen, 2011). Nevertheless, few studies have evaluated the feedback design of ASR-based CALL systems because most are commercial package software which lack pedagogical strategies for the provision of feedback (Neri, Cucchiarini, Strik, & Boves, 2002; Engwall & Balter, 2007), and few developers have carefully taken targeted learners’ needs into consideration during system development. Moreover, the requirement to develop ASR technology is technically demanding and challenging, needing researchers with different domain knowledge and with both technical expertise and a learning technology background to work together.

Research questions

Understanding the abovementioned research background, we formed an interdisciplinary research team, and attempted to construct an intelligent computer-assisted speaking learning (iCASL) system in which the ASR web-service technique and pedagogical corrective feedback are integrated. Previously, we had explored our targeted learners’ requirements for using ASR-based programs for speaking learning (Wang & Young, 2012). In this study, we adopted the previous evaluation results into the system implementation and aimed to examine whether the developed system has the potential to provide learners, especially adults, with extended opportunities for practicing English speaking after graduating from formal education. The research questions of this study are as follows:

1. How do learners perceive the iCASL system embedded with ASR technology for English speaking? Does the system support their self-paced English learning?
2. Can learners achieve better English speaking after using the iCASL system? And why?
3. Can the three-level feedback iCASL system promote learning effectiveness and motivation more effectively than the one-level system? And why?

Literature review

Intelligent CALL systems: Automatic speech recognition (ASR) technology in ICALL

Over the last few decades, the variety of research on CALL has increased dramatically. Some of the topics include mobile assisted language learning (Shield & Kukulska, 2008; Cheng et al., 2010; Wu et al., 2011), tangible companions for learning conversation (Wang, Young, & Jang, 2013), virtual learning environments in Second Life (Wang, Song, Stone, & Yen, 2009) and web-based voice recognition platforms for language learning (Lambacher, 1999; Lai, Tsai, & Yu, 2009). The above studies have proved that CALL systems offer learners extended opportunities and material to simulate realistic learning interaction in a private and stress-free environment. More advanced CALL systems that involve the application of state-of-the-art computing technology such as natural language processing (NLP), artificial intelligence (AI), and automatic speech recognition technology (ASR) for language learning are referred to as intelligent computer assisted language learning (ICALL) (Figure 1). The ASR-based ICALL systems have been attracting an increasing amount of interest from researchers and English instructors alike (Kim, 2006; Lu & Jaw, 2010). The timely speaking evaluation of learners’ utterances is especially beneficial for acquiring listening and oral skills. For example, Meetei (2012) used audio-visual material embedded with a speech recognition function to diagnose learners’ pronunciation errors; while Chwo (2012) applied a package of e-learning software which embedded speech recognition technology into a college speaking and listening class to enhance the learners’ multi-language ability. Their results all indicate the effectiveness of utilizing speech recognition technology for language learning, and also show that learners can become engaged in the stress-free speaking environments.
The importance of feedback in language learning

The importance of providing learners with corrective feedback according to their learning performance rather than only giving them learning opportunities while using a CALL system for speaking learning has been recognized by several studies (Chiu, Liou, & Yeh, 2007; Chen, 2011). An awareness of the difference between learners’ utterances and the target language could hence be a factor driving learners to improve and construct their confidence in language learning, especially for low achievement learners (Hawkins, 1987). Meanwhile, providing learners with language feedback is a form of scaffolding that supports students in the accomplishment of their learning goals through step-by-step instruction, and helps them learn the content more effectively (Ohta, 2000; Chang, Sung, & Chen, 2002). Corrective feedback can be presented in implicit or explicit form. A related study (Lyster & Ranta, 1997) revealed that English teachers used recasting and implicit feedback the most to correct learners’ language errors in language classes; however, the students were less aware of their own language errors from recasting, but their awareness improved as a result of the explicit feedback from the instructor. Another two studies evaluated the use of implicit feedback in oral classes. They also reinforced the finding that explicit and segmented feedback is more efficient than implicit feedback for learners to notice their language faults (Pica, Young, & Doughty, 1987; Bigelow, Delmas, Hansen, & Tarone, 2006). In sum, we should consider pedagogical purposes and students’ responses to various types of feedback while designing language feedback, and help them improve through self-repairing their language errors with an integration of both explicit and implicit feedback.

Related CALL research on corrective feedback design

Mich, Neri, & Giuliani (2005) constructed a voice recognition system which provides three kinds of feedback, including a waveform of the learners’ uttered words along with an expression of emotion and animated characters. The comments from the teachers and learners were positive, though some learners indicated that they did not know why some of their utterances were rated as mispronounced. Similar findings were also obtained by other researchers (Ehsani & Knodt, 1998; Kommissarchik & Kommissarchik, 2000). Therefore, digital waveforms or oscillograms have not proved to be effective tools in helping learners recognize their pronunciation errors. A follow-up study was carried out by the same research team (Neri, Micha, Gerosaa, & Giuliani, 2008). Instead of using oscillograms and animated characters as feedback, they used feedback such as ‘Try it again’ and ‘Well done.’ Nevertheless, the learners still needed and expected specific feedback on their utterances to improve their speaking errors. An online program, CandleTalk, was developed to help freshmen in their conversation practice (Chiu, Liou, & Yeh, 2007). The purpose of the system was to enhance advanced learners’ conversation ability, and provided them with implicit feedback such as numeric scores, instead of correcting their pronunciation directly. However, observation of the learners’ classroom behavior showed that this implicit feedback was insufficient, and that explicit and pertinent feedback is still needed for learners to recognize their pronunciation errors. More recently, Chen (2011) implemented another evaluation of a website that utilizes speech recognition technology, and the feedback from college students and teachers was collected. Most of the students suggested that more pictures, videos and games could be added into the program to attract learners and facilitate learning. Besides, evaluation by teachers reinforced the idea suggested in the previous study, namely that students need meaningful and segmented feedback to indicate their errors in order to help them modify their language faults.
The Intelligent Computer Assisted Speaking Learning (iCASL) system

Based on the reviewed literature, we summarize a general ASR-based CALL model (Figure 2a) and then propose the iCASL system integrating ASR techniques and multiple levels of corrective feedback for English speaking (Figure 2b). The ASR technique judges the learners’ pronunciation based on a statistics model, Hidden Markov Model, that has been used as the sequence-based classifier for speech recognition (Baum & Petrie, 1966). Besides, the principles for designing the iCASL system were based on the importance of language interaction in second language acquisition and the rationale of scaffolding perspective. The language researchers suggested that through the repetition of learning input and output, the effectiveness of second language acquisition could be maximized (Gass & Varonis, 1994; Long, Inagaki, & Ortega; 1998). Hence, the iCASL system provided students with self drill-and-practice opportunities to support learner-centered speaking acquisition. Meanwhile, the study designed the corrective feedback as three levels to scaffold learners acquiring correct pronunciation gradually and to assist learners in accomplishing learning tasks at their own pace. The motivation for proposing the system is to provide Taiwanese learners, especially those who have already graduated from formal education, with flexible opportunities to practice speaking English at home. Furthermore, we aimed to explore the effectiveness of the corrective feedback in the iCASL system for addressing the actual needs of learners in Taiwan.

![Figure 2. A general ASR-based CALL model and iCALS system model](image)

The iCASL system could be regarded as a web-based speaking learning management system that assists learners in arranging their English study progress and trace their speaking performance. The system consists of four modules including the Expert module, the Instruction module, the Student module and the ASR module, each of which has its specific role in assisting learners to achieve self-paced learning and management. The proposed framework is illustrated in Figure 3.

![Figure 3. Architecture and modules of the iCASL system](image)
Expert and instruction modules: The role of the learning tutor

The Expert module and the Instruction module play the role of an English tutor and are closely related. The Expert module is a database which consists of English learning material, while the Instruction module chooses learning content from the database then organizes it into short learning units. Each learning unit includes two activities, listening and speaking (Figure 4 A-a). Meanwhile, the Instruction module automatically proposes suggested learning paths based on the progress of each student, and shows the paths in hyperlinks on the index page (Figure 4 B-a).

Figure 4. Screenshots of Expert and Instruction module

ASR module: The role of the speaking evaluator

The ASR module acts as a speaking evaluator. Students practice English speaking and get immediate feedback on their English utterances from the system (Figure 5). The ASR module applies a self-developed speech application programming interface (API) for speaking evaluation. The API adopts the Carnegie Mellon University (CMU) machine-readable pronunciation dictionary for mapping from words to their pronunciation in the given phoneme sets (Table 1). The current phoneme set contains 39 phonemes that are based on the Advanced Research Projects Agency bet (ARPA bet) symbol. The ARPA bet symbol is a phonetic transcription code set that represents the phonemes of American English in American Standard Code for Information Interchange (ASCII) characters for speech recognition use. Subsequently, a statistical process (Hidden Markov Model) graduates the textual code according to the factors of tone, speed, volume and timbre (Chen, Lo, & Jang; 2004), and the API generates a score and feedback based on the statistical results.

Table 1. Example of 39 phonemes with translation (from the CMU pronunciation dictionary)

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Example</th>
<th>Translation</th>
<th>Phoneme</th>
<th>Example</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Odd</td>
<td>AA D</td>
<td>L</td>
<td>Lee</td>
<td>L IY</td>
</tr>
<tr>
<td>AE</td>
<td>At</td>
<td>AE T</td>
<td>M</td>
<td>Ne</td>
<td>M IY</td>
</tr>
<tr>
<td>AH</td>
<td>Hut</td>
<td>HH AH T</td>
<td>N</td>
<td>Knee</td>
<td>N IY</td>
</tr>
<tr>
<td>AO</td>
<td>Ought</td>
<td>AO T</td>
<td>NG</td>
<td>ping</td>
<td>P IH NG</td>
</tr>
<tr>
<td>AW</td>
<td>Cow</td>
<td>K AW</td>
<td>OW</td>
<td>Oat</td>
<td>OW Y</td>
</tr>
<tr>
<td>AY</td>
<td>Hide</td>
<td>HH AY D</td>
<td>OY</td>
<td>Toy</td>
<td>T OY</td>
</tr>
<tr>
<td>B</td>
<td>Be</td>
<td>B IT</td>
<td>P</td>
<td>Pee</td>
<td>P IY</td>
</tr>
<tr>
<td>CH</td>
<td>Cheese</td>
<td>CH IY Z</td>
<td>R</td>
<td>Read</td>
<td>R IY D</td>
</tr>
<tr>
<td>D</td>
<td>Dee</td>
<td>D IY</td>
<td>S</td>
<td>Sea</td>
<td>S IY</td>
</tr>
<tr>
<td>DH</td>
<td>Thee</td>
<td>DH IY</td>
<td>SH</td>
<td>She</td>
<td>SH IY</td>
</tr>
<tr>
<td>EH</td>
<td>Ed</td>
<td>EH D</td>
<td>T</td>
<td>Tea</td>
<td>T IY</td>
</tr>
<tr>
<td>ER</td>
<td>Hurt</td>
<td>HH ER T</td>
<td>TH</td>
<td>Theta</td>
<td>TH EY T AH</td>
</tr>
<tr>
<td>EY</td>
<td>Ate</td>
<td>EY T</td>
<td>UH</td>
<td>Hood</td>
<td>HH UH D</td>
</tr>
<tr>
<td>F</td>
<td>Fee</td>
<td>F IY</td>
<td>UW</td>
<td>Two</td>
<td>T UW</td>
</tr>
<tr>
<td>G</td>
<td>Green</td>
<td>G R IY N</td>
<td>V</td>
<td>Vee</td>
<td>V IY</td>
</tr>
<tr>
<td>HH</td>
<td>He</td>
<td>HH IY</td>
<td>W</td>
<td>We</td>
<td>W IY</td>
</tr>
<tr>
<td>IH</td>
<td>It</td>
<td>IHT</td>
<td>Y</td>
<td>yield</td>
<td>Y IY L D</td>
</tr>
</tbody>
</table>
Student module: The role of the learning recorder

The Student module includes learners’ personal information and learning portfolios, and this module acts as a learning recorder. The learning portfolios contain learners’ study progress and speaking performance. The portfolios present the information in the form of a webpage showing learners’ audio data including practicing time and scores, allowing the learners to take a second look and to think about how they can improve in their future attempts (Figure 6).

Corrective feedback design and presentation

According to the reviewed literature, feedback could be presented either in implicit or explicit formats with text or graphs (Lyster & Ranta, 1997; Neri, Micha, Gerosa, & Giuliani, 2008). Under this framework, the study designed the corrective feedback of the iCASL system as three levels (Figure 7). The first level focuses on providing implicit feedback, showing the learner’s pronunciation score and audio waveform. At the second level, which aims to provide explicit feedback, there is a comment, a list of words that are pronounced accurately and inaccurately, and an audio toolbar for replaying the learner’s utterance. At the third level, demonstration of the accurate utterances with both full sentence and single-word form at normal and slow speed are available. The slow button is particularly useful for pronouncing multi-syllabic vocabulary (Engwall, Balter, Oster, & Kjellstrom, 2006).
Methodology

The methodology adopted in this study includes two parts: evaluating the effects of using the three-level-feedback iCASL system on pronunciation learning, and evaluating the use of the iCASL system. For the effects of pronunciation learning, a quantitative approach was employed. The study uses comparative test data and empirical experiments to report on the performance of learning English in the iCASL system with different levels of learning feedback. The control group practiced English speaking using the single-level-feedback system in which only a waveform diagram was presented as feedback to evaluate the learners’ speaking, while the experimental group was given the three-level-feedback iCASL system integrating implicit and explicit elements in the feedback presentation. The multiple-choice questions and speaking pre-test and post-test were collected, and an independent samples t-test was applied for statistical analysis. For the iCASL system evaluation, both quantitative and qualitative approaches were employed. Questionnaires regarding how the participants perceived the system in terms of its usability and interactivity were administered to the students. Furthermore, intensive observations of the learners’ learning processes through system records and analysis of the raw audio data in the system were carried out for further analysis.

The computer-based tests consisted of five multiple-choice questions for which the learners had to read the proverbs in Chinese and choose the matching English proverbs. In the speaking pre-test and post-test, the participants had to listen to the pronunciation in the system and then repeat it. The purpose of the multiple-choice questions was to understand how much language knowledge the students gained during the experiment, whereas the speaking test was designed to assess the students’ speaking accuracy before and after using the iCASL system. In addition, the questionnaire, based on a five-point Likert scale (from 5 to 1: strongly agree, agree, neutral, disagree, strongly disagree), consisted of thirty items and two open-ended questions, and had a Cronbach’s $\alpha$ of 0.91.

Participants

A total of 38 adult learners from different backgrounds and aged from 23 to 40 were invited to participate in the study. They were divided into two groups (Table 2). The experimental group consisted of 18 learners, 9 of whom were graduate students and 9 of whom were working. The control group consisted of the other 20 learners, 8 of whom were graduate students and 12 of whom were in the workforce.

<table>
<thead>
<tr>
<th>Table 2. The participants</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Background</strong></td>
</tr>
<tr>
<td>Experimental Group (E.G.)</td>
</tr>
<tr>
<td>Control Group (C.G.)</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
Learning materials: Proverbs

The learning topic chosen for this study is based on our previous research results (Wang & Young, 2012) which suggested that English proverbs are one of the most desired learning contents for Taiwanese learners of English. Meanwhile, researchers have pointed out that the goal of pronunciation teaching is not necessarily to acquire native-like pronunciation, but for “developing functional intelligibility, communicability, increased self-confidence, the development of speech monitoring abilities and speech modification strategies for use beyond the classroom” (Otlowski, 1998). In support of this argument, Pennington (1999) suggested that instructors should link pronunciation to other communicative learning goals such as putting effort into helping students raise their awareness of the contrast between the first language (L1) and the second language (L2), and the awareness of the social significance of the relationship between the L1 and L2. In accordance with the preceding point, we chose English proverbs as the learning topic for this study. Proverbs are the essence of wisdom accumulated from history and passed down from generation to generation, and can act as a bridge to help one understand the language in more depth. We hoped that the participants in this study could not only enhance their pronunciation but also become more aware of the background of the language. Besides, proverb learning is not limited by age, and could suit learners of all language levels; therefore it could fulfill the actual needs of Taiwanese learners. If language learners are familiar with proverbs, they can also apply this knowledge in their writing or oral conversation and achieve better language performance. Despite this, we found that most CALL research, especially in the Chinese area, included a wide range of topics, but did not cover English proverbs (Ma & Kelly, 2006; Chen, 2011). Hence, based on the suggestion proposed in the previous research results (Wang & Young, 2012) and cultural considerations in language learning, the iCASL system adopted English proverbs as the learning materials to fulfill the local students’ needs. The topics included Learning and Experience, Time, Chances and Success, Capability and Talent. The presentation of the learning material is illustrated in Figure 8. We integrated differing modalities including text, sound and figures into the content design.

![Figure 8. The presentation of the learning materials in the iCASL system](image)

Data collection

The duration of the data collection was eight weeks, and the participants were encouraged to arrange their schedule to practice English speaking using the system at least two times a week, and to complete all of the learning units. Their learning steps and learning processes while using the system were also recorded for further analysis.

Data analysis

Participants’ feedback and perceptions of the use of the iCASL system

The system records and data from the questionnaires regarding the learners’ reflections on using the iCASL system were analyzed in order to answer the first research question. We can infer from the system login records that the participants learned English with the system during two specific time periods, from 2 to 4 o’clock in the afternoon and from 9 to 12 o’clock at night, and those who were working often used the system from 10pm to 11pm (Figure 9). The system record shows that the average number of times each learner accessed the system was about 2.5 times per
week, and the average login time was from 30 to 40 minutes. We further analyzed the learners’ learning paths and organized them into two types (Figure 10). It was found that most of the learners scheduled their learning according to the suggested routes and followed a straight path while learning (Figure 10-a). Others, however, had a flexible schedule and tended to review the previous contents first before learning the new topic (Figure 10-b). We were pleased to find that the learners were willing to use the iCASL system for self-paced learning at the end of their working day.

The data from the questionnaire related to the learning and operational aspect of the system are shown in Table 3. It is noticed that 72% of the learners were initially embarrassed and were afraid of making mistakes when practicing English with teachers and classmates due to their prior learning experiences (Table 3, Q1). In contrast, they indicated that they felt relaxed when practicing English speaking with the system (Table 3, Q2). As many as 94% of the learners reflected that the flexibility of using the system enhanced their motivation to practice English speaking (Table 3, Q6). Moreover, most of them found that the system was helpful and they had positive reflections on the system operation (Table 3, Q8). They indicated that they would be willing to use the system to learn English in the future (Table 3, Q10). Furthermore, they claimed that using the iCASL system promoted opportunities for English speaking through acquiring proverbs, and most of them (93%) reflected that the English proverbs they had learned were immediately applicable to their daily conversation, while 72% agreed that they were more confident in speaking English after the eight weeks of on-line practice (Table 4, Q3-Q5).

From the qualitative feedback of the open-ended questions, the learners commented that “It was a good way of learning English speaking after work, and I hope it could provide me with a greater variety of learning content in the future” (EG2) and “I enjoyed this kind of personal English speaking learning. I could practice speaking repeatedly without any time limitation and I didn’t have to worry about teachers being impatient” (EG6). Suggestions about how to further improve the design of the system were also collected, such as students pointing out that they would like to know how their utterances were scored and would also like some tips for getting higher speaking grades. Other feedback from students is presented in Table 5.
Table 3. Questionnaire results of students’ feedback on the iCASL system

<table>
<thead>
<tr>
<th>Item</th>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>I feel embarrassed while practicing English with teachers and classmates</td>
<td>6.3%</td>
<td>15.6%</td>
<td>6.3%</td>
<td>31.3%</td>
<td>40.6%</td>
<td>3.84</td>
</tr>
<tr>
<td>Q2</td>
<td>I feel embarrassed while practicing English with system.</td>
<td>37.5%</td>
<td>25.0%</td>
<td>25.0%</td>
<td>9.4%</td>
<td>3.1%</td>
<td>2.16</td>
</tr>
<tr>
<td>Q3</td>
<td>I enjoy practice English with the system.</td>
<td>13.3%</td>
<td>13.3%</td>
<td>26.7%</td>
<td>13.3%</td>
<td>33.3%</td>
<td>3.84</td>
</tr>
<tr>
<td>Q4</td>
<td>I am afraid of making mistakes while practicing English with teacher.</td>
<td>3.1%</td>
<td>12.5%</td>
<td>12.5%</td>
<td>28.1%</td>
<td>43.8%</td>
<td>3.97</td>
</tr>
<tr>
<td>Q5</td>
<td>I am afraid of making mistakes while practicing English with the system.</td>
<td>50.0%</td>
<td>18.8%</td>
<td>21.9%</td>
<td>6.3%</td>
<td>3.1%</td>
<td>1.94</td>
</tr>
<tr>
<td>Q6</td>
<td>The flexibility of using the system enhance the motivation to practice</td>
<td>3.1%</td>
<td>0.0%</td>
<td>3.1%</td>
<td>50.0%</td>
<td>43.8%</td>
<td>4.31</td>
</tr>
<tr>
<td></td>
<td>English speaking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q8</td>
<td>The information of the system is clearly</td>
<td>3.1%</td>
<td>6.3%</td>
<td>25.0%</td>
<td>37.5%</td>
<td>28.1%</td>
<td>3.81</td>
</tr>
<tr>
<td>Q9</td>
<td>The loading of each week is appropriately</td>
<td>3.1%</td>
<td>9.4%</td>
<td>12.5%</td>
<td>40.6%</td>
<td>34.4%</td>
<td>3.35</td>
</tr>
<tr>
<td>Q10</td>
<td>I would like to use the system for further learning.</td>
<td>3.1%</td>
<td>0.0%</td>
<td>6.3%</td>
<td>46.9%</td>
<td>43.8%</td>
<td>4.28</td>
</tr>
</tbody>
</table>

Note. 5 to 1 points: strongly agree, agree, neutral, disagree, and strongly disagree.

Table 4. Questionnaire results of learning effectiveness

<table>
<thead>
<tr>
<th>Item</th>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Using the system enhances the opportunities of learning English proverbs.</td>
<td>3.1%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>28.1%</td>
<td>68.8%</td>
<td>4.56</td>
</tr>
<tr>
<td>Q2</td>
<td>Using the system enhances the opportunities of English speaking.</td>
<td>3.1%</td>
<td>11.1%</td>
<td>3.1%</td>
<td>25.0%</td>
<td>68.8%</td>
<td>4.59</td>
</tr>
<tr>
<td>Q3</td>
<td>Using the system enhances the English speaking ability.</td>
<td>3.1%</td>
<td>6.3%</td>
<td>18.8%</td>
<td>43.8%</td>
<td>28.1%</td>
<td>3.87</td>
</tr>
<tr>
<td>Q4</td>
<td>I could speak several English proverbs after practicing.</td>
<td>3.1%</td>
<td>3.1%</td>
<td>0.0%</td>
<td>37.5%</td>
<td>56.3%</td>
<td>4.41</td>
</tr>
<tr>
<td>Q5</td>
<td>Using the system promotes my motivation to speak English.</td>
<td>3.1%</td>
<td>3.1%</td>
<td>15.6%</td>
<td>46.9%</td>
<td>31.3%</td>
<td>4.00</td>
</tr>
</tbody>
</table>

Note. 5 to 1 points: strongly agree, agree, neutral, disagree, and strongly disagree.

Table 5. Comments and suggestion from learners

<table>
<thead>
<tr>
<th>NO.</th>
<th>Comments</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>EG4</td>
<td>I hope that the contents could be extended such as English phrases or</td>
<td>CE15 I</td>
</tr>
<tr>
<td></td>
<td>daily sentences could be included so that I could have more chance to</td>
<td>think it</td>
</tr>
<tr>
<td></td>
<td>practice English speaking.</td>
<td>would be</td>
</tr>
<tr>
<td></td>
<td></td>
<td>interesting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to add the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>stories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>in learning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>content when</td>
</tr>
<tr>
<td></td>
<td></td>
<td>learning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>proverbs.</td>
</tr>
<tr>
<td>EG8</td>
<td>I want to know the scoring mechanism for speaking and some tips for</td>
<td>EG11 The</td>
</tr>
<tr>
<td></td>
<td>getting higher speaking grades.</td>
<td>functions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>of the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>website</td>
</tr>
<tr>
<td></td>
<td></td>
<td>are quite</td>
</tr>
<tr>
<td></td>
<td></td>
<td>good but</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hope the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>textual</td>
</tr>
<tr>
<td></td>
<td></td>
<td>feedback</td>
</tr>
<tr>
<td></td>
<td></td>
<td>could be</td>
</tr>
<tr>
<td></td>
<td></td>
<td>transformed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>into human</td>
</tr>
<tr>
<td></td>
<td></td>
<td>voice as</td>
</tr>
<tr>
<td></td>
<td></td>
<td>that would</td>
</tr>
<tr>
<td></td>
<td></td>
<td>motive me</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to learn.</td>
</tr>
<tr>
<td>EG17</td>
<td>I hope the system could include the English typing function so that I</td>
<td>EG20 I</td>
</tr>
<tr>
<td></td>
<td>could remember more vocabulary.</td>
<td>would like</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to have a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>virtual</td>
</tr>
<tr>
<td></td>
<td></td>
<td>teacher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(figure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or animation)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>instructing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>me how to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>enhance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>English</td>
</tr>
<tr>
<td></td>
<td></td>
<td>speaking.</td>
</tr>
<tr>
<td>EG22</td>
<td>I hope to have speaking cloze test, for example, the system pronounces</td>
<td></td>
</tr>
<tr>
<td></td>
<td>the first three words the proverb and we have to finish the following</td>
<td></td>
</tr>
<tr>
<td></td>
<td>part.</td>
<td></td>
</tr>
</tbody>
</table>

Note. 5 to 1 points: strongly agree, agree, neutral, disagree, and strongly disagree.

Learning effectiveness of the experimental group (E.G.) and control group (C.G.)

The pre-test, post-test and raw audio data of the E.G. and C.G. were analyzed to answer the second research question. Overall, the average scores of the post-test were improved than the scores of the pre-test for the two groups. The researchers further compared the improvement rates between the pre-test and post-test for the two groups to understand learners’ learning performance. The independent sample t-test shows that there were no significant differences in the improvement rates of the multiple-choice questions between two groups (Table 6a), but, noticeably, there were significant differences in the improvement rates between the pre-test and post-test of the speaking part.

Learning effectiveness of the experimental group (E.G.) and control group (C.G.)

The pre-test, post-test and raw audio data of the E.G. and C.G. were analyzed to answer the second research question. Overall, the average scores of the post-test were improved than the scores of the pre-test for the two groups. The researchers further compared the improvement rates between the pre-test and post-test for the two groups to understand learners’ learning performance. The independent sample t-test shows that there were no significant differences in the improvement rates of the multiple-choice questions between two groups (Table 6a), but, noticeably, there were significant differences in the improvement rates between the pre-test and post-test of the speaking part.
and the learners in the E.G. achieved better improvement rates in the speaking test than those in the C.G. (Table 6b). This indicates that learners in both groups can read the proverbs in Chinese and choose the matching English proverbs correctly in the multiple-choice questions after eight weeks of on-line learning. However, only the learners in the E.G. achieved significant improvement rates in pronunciation with the help of the three-level feedback.

<table>
<thead>
<tr>
<th></th>
<th>Improvement rates between pre-test and post-test</th>
<th></th>
<th>Improvement rates between pre-test and post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Groups</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>Writing</td>
<td>E.G.</td>
<td>57.74</td>
<td>64.42</td>
</tr>
<tr>
<td>part</td>
<td>C.G.</td>
<td>33.90</td>
<td>44.84</td>
</tr>
</tbody>
</table>

Based on the audio recordings of the participants, it was found that the students tended to pronounce the vocabulary correctly after practicing speaking with the iCASL system. We randomly selected 10 learners (S1-S5 from E.G., and S6-S10 from C.G.) to analyze their audio files of utterances in the system to further understand the effectiveness of the iCASL system in terms of improving the learners’ pronunciation. Table 7 illustrates the score improvements of learners for their pronunciation of the word “live” (/l+i+v/) from the proverb “as we live, so we learn” between the pre- and the post-test. We found that the average pronunciation score for the pre-test of the word “live” was 59.1. The system records show that the learners had difficulties pronouncing the /v/ sound, as they tended to use /t/ or /d/ instead. However, the scores for the post-test were improved after eight weeks of practice, especially for the learners in the E.G. This indicates that the learners pronounced the target language better with the scaffolding process through the three-level feedback. The first level provided learners with a rough speaking score and the second level helped the learners recognize their speaking errors with detailed comments. After the learners understand their advantage and shortcomings of their utterances, they had a chance to hear their voice again and to compare it with the model pronunciation. The second level feedback enabled the learners to become aware of their language errors through the help of textual information with an audio recast. With the aid of multiple-level scaffolding, learners could re-pronounce the target words and internalize the correct pronunciation by imitating the model pronunciation gradually.

<table>
<thead>
<tr>
<th></th>
<th>Total E.G. (s1~s5)</th>
<th>C.G. (s6~s10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average score of speaking pre-test</td>
<td>59.1</td>
<td>55.6</td>
</tr>
<tr>
<td>Average score of speaking post-test</td>
<td>77.2</td>
<td>85.8</td>
</tr>
<tr>
<td>Range of two scores</td>
<td>18.1</td>
<td>30.2</td>
</tr>
</tbody>
</table>

**Learners’ evaluation of the corrective feedback**

The data from the questionnaires and the open-ended questions regarding the learners’ reflections on the multiple-level feedback were analyzed in response to the third research question. From the result of the questionnaires, it shows that the three-level-feedback iCASL system is helpful for improving pronunciation. One learner in the E.G. pointed out that she could improve her inaccurate pronunciation by reading the immediate feedback, commenting, “What a surprise that the score improved after I modified the utterances according to the system feedback!” Learners in the E.G. also claimed that they were more willing to have repeated practice with the system in order to achieve positive learning reinforcement such as the smiley face. Moreover, among the second and third level feedback given to the E.G learners, the model pronunciation of the sentence, the model pronunciation of the individual vocabulary and the audio file of learner utterances were more useful in helping the students to pronounce the target language.

On the contrary, the students in the C.G were confused about the audio waveform diagram. More than 94% reflected that they did not know how to adjust their pronunciation just by using the first-level feedback, and they suggested the need for details from the system. Besides, the results from the open-ended feedback also indicated that the learners in the C.G. had problems understanding the feedback only with numeric scores and audio waveform diagrams. One learner from the C.G. commented, “I cannot read the waveform from the website and I would prefer an explanation such as a textual description to show me how I could modify my pronunciation” (CE15). Furthermore, we collected the preferences of the C.G. students regarding the immediate feedback design to further improve the iCASL system. The statistical data suggest that the top three functions which should be included in the corrective feedback to assist
learning are: (1) model pronunciation of the English sentences; (2) an audio file of the utterance that can be played immediately; and (3) a list of the inaccurately pronounced words (Table 8).

Table 8. The results of the C.G. learners’ preferred information in system feedback

<table>
<thead>
<tr>
<th>Items</th>
<th>A: C.G.</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Model pronunciation of the sentence</td>
<td>25%</td>
<td>1</td>
</tr>
<tr>
<td>2. An audio file of utterance</td>
<td>17%</td>
<td>2</td>
</tr>
<tr>
<td>3. Waveform of students’ utterance</td>
<td>6%</td>
<td>7</td>
</tr>
<tr>
<td>4. Model pronunciation of individual vocabulary</td>
<td>14%</td>
<td>4</td>
</tr>
<tr>
<td>5. Pronunciation in slow speed</td>
<td>9%</td>
<td>6</td>
</tr>
<tr>
<td>6. A list of inaccurately pronounced words</td>
<td>17%</td>
<td>2</td>
</tr>
<tr>
<td>7. A list of accurately pronounced words</td>
<td>12%</td>
<td>5</td>
</tr>
</tbody>
</table>

Discussion

Design of the iCASL system: For self-paced and informal speaking learning

The statistical data from the questionnaires reflect that the learners had positive reflection of using the iCASL system embedded with ASR technology for English speaking. The personal speaking practice environment and the flexibility of using the system enhanced learners’ motivation to practice English speaking, especially those who were working, were engaged in the web-based English speaking scenario, despite being busy. The four modules of the system embodied the functions of learning path suggestions, immediate speaking evaluation and personal portfolios, which supported the learners to achieve self-paced language learning. The participants used their free time, during lunch or at night, to practice speaking English. Despite the fragmented time, the accumulated short periods of the informal learning opportunities made it possible for the learners to achieve great improvement in their spoken language. The learners mentioned that the path suggestion function was especially beneficial because it helped them to manage their learning. They were engaged in this self-paced speaking learning scenario, in which some of them liked to start from their latest progress instead of learning from the beginning, while others preferred to review the contents before learning a new topic. Meanwhile, they also reflected that it was a new experience in which they received immediate evaluation of their utterances, compared to their previous learning experience in the traditional English classroom in which only a few students had the chance to practice speaking with English instructors. Furthermore, they felt more comfortable and relaxed while practicing speaking in the private learning environment, and therefore were willing to continue using the system for more speaking practice. In sum, the implementation of the iCASL system for language learning promotes Taiwanese learners’ English speaking opportunities through informal education and hence may have the potential to achieve the goal of lifelong language learning.

Three-level feedback for improving English pronunciation

This study referred to the results of previous research (Mackey & Goo, 2007; Neri, Micha, Gerosaa, & Giuliani, 2008; Chen, 2011) and provided the E.G. students with three levels of learning support, feedback in an implicit format firstly and then in an explicit format. The statistical data reflect that all learners achieved better English speaking after using the iCALS system for eight weeks. While, noticeable, the learners in the E.G. achieved significant improvement rates in pronunciation between the pre-test and post-test with the help of the three-level feedback. The students in the C.G. pointed out that they had difficulty detecting their pronunciation errors from the single-level implicit feedback and in-depth analysis was needed and expected. The results of this study are in accordance with the results of earlier studies (Mich, Neri, & Giuliani, 2005; Chiu, Liou, & Yeh, 2007) that indicated adopting ASR technology into language learning had positive impact on speaking acquisition; however, the implicit feedback such as audio waveform was insufficient and meaningful and segmented feedback could be added into the program for the learners to facilitate pronunciation learning. On the other hand, the study also indicates that the second-level feedback, which involved the explicit format of immediate audio replay (recast) with a textual description, benefited the learners the most. It is our inference that the learners in the E.C. received simultaneous dual information from the audio and visual presentation and hence had better potential to improve their English speaking than the learners who were provided with single modality feedback. The result implies that the corrective
feedback proposed in the study has been proved to have a positive effect on the students, and that integrated feedback should be designed for learners, with explicit information being provided after implicit feedback.

Conclusions

The study explores how to implement a pedagogical ASR-based iCASL system to support adult learners with a private, flexible and individual learning environment to practice English pronunciation. Besides, the researchers further proposed the three-level feedback principle, consisting of feedback in an implicit format firstly and then in an explicit format. The results of the study indicate that the learners with three-level feedback made significant progress in English speaking and the iCASL system, jointly designed and implemented by a multi-discipline research team has received learners’ positive feedback. The collaboration among the Computer Science Lab and Learning Technology Lab has contributed to the successful actualization of an online automatic speech recognition based personal speaking learning program for Taiwanese learners that otherwise might not be possible. We conclude from the current study that the use of iCASL system has great potential to support self-paced learning and provide Taiwanese with more flexible learning opportunities to satisfy their learning needs. Moreover, the iCASL system serves as a helpful tool for Taiwanese learners, especially for those who have already graduated from formal education and could not afford the time to learn in fixed educational settings. What’ more, the application of iCASL system could be used for remedial teaching to reduce language teachers’ burden of taking care of every student in the given instructional conditions, also, teachers could hold speaking competition or establish speaking learning community in the iCALS system to encourage learners to change learning experience with each other and to achieve the goal of lifelong language learning. While, learners have to use the speech recognition program in a non-noisy environment, so it is important and necessary to provide learners with a quiet and individual learning scenario to ensure that the speech recognition function can work correctly. The limitations of this study are as follows. First, this study restricted itself to effects on self-learning under informal learning environments at home or at work so that the researchers did not evaluate learners’ learning behaviors and their interaction with the system. Second, the participants of the study were adults so that the results cannot be generalized to younger learners, unless further study with the younger learners will be conducted.

In the future, we plan to set up a scoring mechanism and construct learning classification to achieve a more complete ASR-based language learning system. Furthermore, we tend to extend the uses of the system for learners of different ages to address the actual needs of learners in Taiwan. It is our aim to get students to devote their own time to lifelong English practice. More research findings related to issues of design, system implementation and the effectiveness of learning English speaking with ASR-based CALL systems will be shared in the near future.

References


Chwo, S. M. G. (2012). Enhancing listening proficiency via multi modality technology an initial finding from technology university non-English major EFL learners in central Taiwan. In J. Colpaert et al. (Eds.), Proceedings of the fifteenth international CALL conference (pp. 212-215). Taichung, Taiwan: Providence University.


Meetei, N. S. (2012). The effective ways of teaching English pronunciation using audio-visual aids. In J. Colpaert et al. (Eds.), Proceedings of the fifteenth international CALL conference (pp. 553-555). Taichung, Taiwan: Providence University.


Shield L., & Kukulska-Hulme A. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL, 20*(3), 249-252.


ABSTRACT

This study examines how incorporating different electronic feedback devices (i.e., clickers versus web-based polling) may affect specific types of student engagement (i.e., behavioral, emotional, and cognitive engagement), whether students' self-efficacy for learning and performance may differ between courses that have integrated clickers and those that use web-based polling, and whether using web-based polling influences faculty members' instructional practices. The participants included six instructors and 209 students enrolled in classes at a university in the southwestern United States in which the instructors used either clickers or web-based polling. The Plenty-of-Time Teaching (PoTT) and the Just-in-Time Teaching (JiTT) approaches and their implications are presented. The results of this study highlight the benefits of using various types of electronic feedback devices to provide innovative ways to implement JiTT or PoTT, such as gauging students' understanding with pre-class polls, and offer insights that can benefit educators who wish to promote students' emotional and cognitive engagement with various types of feedback devices.

Keywords

Electronic feedback device, Polling, Just-in-time teaching, Engagement

Introduction

One main problem with the traditional lecture format is that students' levels of engagement tend to be low, which may cause their learning to suffer. In the past five years, technology has started to be applied in lecture halls and in online settings to address this issue (Deliialioglu, 2012; Koenig, 2010; Mason, 2011; Middlebrook & Sun, 2013; Sun & Rueda, 2012; Walsh, Sun, & Riconscente, 2011). Specifically, the use of electronic feedback devices (also called electronic voting systems or clickers, herein called “clickers”) is becoming more common in academic settings, especially at higher education levels (Gilbert, 2005; Martyn, 2007).

However, one limitation of clickers is that the instructors can only conduct polling during lectures, as the connection of the device is based on a live session. In addition, students can poll their answers from only one medium (i.e., the "clickers"). With web-based polling (e.g., PollEverywhere.com), students may respond before, during, and after a lecture via various devices, such as text messages, as well as mobile, desktop, and laptop browsers, which provides flexibility for instructors who want to integrate different pedagogies into the curriculum. For example, instructors can post questions on PollEverywhere.com prior to a class and modify the learning activities based on the results of the students' polling responses.

Currently, there is a lack of research empirically investigating web-based polling and its possible benefits with regard to delayed feedback. Therefore, the current study sought to evaluate the effectiveness of clickers and web-based polling for providing delayed feedback from learners and to determine how incorporating different polling strategies (i.e., clickers versus web-based polling) may affect specific types of student engagement (i.e., behavioral, emotional, and cognitive engagement) (Fredricks, Blumenfeld, Friedel, & Paris, 2005; Fredricks, Blumenfeld, & Paris, 2004). The current study also investigated whether students’ self-efficacy for learning and performance differs between courses that have integrated clickers and those that use web-based polling and whether using web-based polling influences faculty members’ instructional practices.
Literature review and research model

Clickers and polling

One way of gaining immediate feedback during classroom instruction is to employ radio frequency-based electronic feedback devices (i.e., “clickers”) (Fortner-Wood, Armistead, Marchand, & Morris, 2013). Clickers are small, portable devices that use infrared or radio frequency technology to transmit and record students’ responses to questions. Clickers provide instantaneous feedback to both the instructor and the students about the level of understanding of the material being presented. The use of clickers has revealed a variety of benefits with regard to instructional goals and objectives; for example, the anonymity of the responses encourages the participation of students who may otherwise be reluctant to do so. The anonymous, simultaneous manner of gathering responses via clickers (in comparison to traditional hand-raising) eliminates students’ tendency to conform to the answers of the academically higher-status students (Kennedy & Cutts, 2005; Stowell & Nelson, 2007). The use of clickers has also encourages an increase in student engagement (Bode, Drane, Kolikant, & Schuller, 2009; Dallaire, 2011; Lasry, 2008; Stowell & Nelson, 2007; Trees & Jackson, 2007).

Just-in-time teaching

Another benefit of electronic feedback devices has been its use with "Just in Time Teaching" (JiTT), given that the clickers facilitate the instructors’ ability to gain immediate feedback regarding what students know or do not know (Novak & Middendorf, 2004; Novak, Patterson, Gavrin, & Christian, 1999; Simkins & Maier, 2004). Given the immediacy of the aggregated responses, instructors can quickly ascertain which concepts need to be re-examined (Lasry, 2008). This teaching strategy was first developed to enhance students’ learning experiences by helping Physics faculty members learn which topics their students struggled with, and it has continued to be used in Physics (Crouch & Mazur, 2001) and in other domains, including Biology (Marrs & Novak, 2004) and Economics (Simkins & Maier, 2004). The following three core principals frame the JiTT approach: maximize the effectiveness of the class time discussion; make non-class time beneficial for students; and foster and perpetuate peer interactions. To facilitate these principals, Novak and colleagues (1999) suggest that web pages be used to house questions related to upcoming course content and that students submit their responses to these questions to the instructor a few hours before class. The latter element then manifests during class time when the instructor uses the feedback garnered from the out-of-class activities to develop a framework for the live class session. In this sense, the teaching occurs “just in time” in that students’ misconceptions, partial conceptions, or high levels of conception inform the instruction. The instructor might have planned to use a set lecture and series of learning activities that, unbeknownst to the instructor, were well beyond the students’ grasp, given their misunderstandings. Thus, delivering lectures that are based on the assumption that students understand the content actually perpetuates their misunderstandings or misconceptions. Conversely, students who possess a high level of understanding may disengage from a lecture and find the learning experience boring.

Plenty-of-time teaching and asynchronous polling

Although the immediacy of aggregated in-class responses helps instructors facilitate the JiTT strategy and quickly ascertain which concepts need to be re-examined (Lasry, 2008), one limitation of this approach is the lack of time available for an instructor to make quality adjustments to the content. One approach to mitigate this problem is the use of “Plenty-of-Time Teaching,” which is the purposeful use of pre-class activities, such as open-ended and multiple-choice questions, conducted via the Internet, that are aimed to engage students with the content prior to a class discussion. Given that students submit their responses a few hours before class, the instructor is able to modify
the content and learning activities based on the levels of students’ misunderstandings or misconceptions with the material (Novak et al., 1999).

Most studies have used open-ended questions to address the difficult problems presented in the text. For example, Crouch and Mazur (2001) typically posed three questions per unit with two question related to difficult aspects of the readings and one aimed at soliciting feedback about the overall difficulty of the concepts from the readings. Similar to this approach, two of the instructors who participated in Crouch and Mazur’s (2001) study used open-ended type questions, whereas the other three used forced response questions. However, unlike previous studies, these instructors solicited the students’ responses to the questions well before the class discussion, typically two to three days before, compared to a few hours before, as suggested by the aforementioned studies. Another consistent factor in this approach is the use of “in class” feedback through electronic polling. Bode et al. (2009), Crouch and Mazur (2001), and Lasry (2008) utilized in-class polling to gain insight into students’ understandings or misconceptions of the materials with regard to how they modified the content based on the pre-discussion feedback. Breaking from this traditional JiTT approach, the current study sought to examine how the use of web-based polling influenced instructors’ willingness to change lecture content based on data from pre-class polls. To date, no studies have examined whether the use of asynchronous pre-class polling influences an instructor’s willingness to change the class lecture or learning activities based on this feedback. To clarify, traditional clicker polls are administered in real time, or synchronously, by an instructor and with other students. The technological limitations dictate that these must be administered in a synchronous manner, such that both the students and the instructor are bound by time and place. Conversely, asynchronous administration is analogous to learning in which students and teachers are not bound by time or place. In this regard, web-based polling can be used both synchronously and asynchronously, with the purported advantage being that an instructor can obtain feedback on students’ understanding well in advance of the class discussion (as well as during and after the class discussion) with an asynchronous system. Therefore, the current study calls this strategy "Plenty of Time Teaching” (PoTT), in contrast to “Just-in-Time Teaching.” The specific focus of the current study was to examine whether the type of polling device used and the manner in which it was used significantly influence student engagement.

**Student engagement**

The notion of student engagement, particularly at the higher education level, has been examined in different capacities over the past 30 years. Astin (1984) initially described the notion of engagement as student involvement and defined it as the degree to which students expend both physical and psychological energies to achieve a particular task. Kuh (2009) similarly asserted that engagement is the amount of time and effort that students devote to achieving a desired task. Kuh (2009) also included the role that institutions play in fostering student engagement in his definition. From these two definitions, it is evident that engagement consists of both observable and unobservable characteristics. To better understand these two categories of characteristics, Fredricks and colleagues (2004) developed a framework in which engagement is comprised of the following three components: behavioral, emotional, and cognitive engagement. Behavioral engagement refers to an individual’s participation in school (Finn, 1993). Thus, behavioral engagement refers to what Astin (1984) and Kuh (2009) described as the time and effort spent on observable academic activities, such as studying, note-taking, participating in class discussions, and preparing for exams. Emotional engagement refers to an individual’s positive or negative feelings towards school (Finn & Voelkl, 1993; Fredricks et al., 2004). Although a learner may engage in observable behaviors that indicate a certain level of effort, the feelings a learner may have in relation to a certain learning activity may not readily be observable. Cognitive engagement refers to an individual’s voluntarily exertive effort to understand and master challenging tasks (Fredricks et al., 2004). Astin (1984) described this as psychological effort. Regardless of whether it is called cognitive or psychological engagement, neither is easily observed. Unlike behavioral or physical engagement and similar to emotional engagement, cognitive engagement is not easily observed, as it is essentially the degree to which a learner uses one or more cognitive processes to learn.

One way to address the problem regarding learners’ unobservable engagement characteristics is to use electronic feedback devices, which utilize questions intended to extrapolate learners’ levels of cognitive engagement, in addition to their behavioral and emotional levels of engagement. Existing studies investigating feedback devices have reported the benefits of increased overall engagement (Bode et al., 2009; Dallaire, 2011; Lasry, 2008; Stowell & Nelson, 2007; Trees & Jackson, 2007), but they have rarely examined the influence of using feedback devices on specific types of student engagement. Hence, this study aimed to determine how incorporating different devices (i.e.,
clickers versus web-based polling) affect the specific types of student engagement (i.e., behavioral, emotional, and cognitive engagement) identified by Fredricks and colleagues (2005; 2004). For the purposes of this study, engagement is defined as the extent to which a learner is cognitively, emotively, and behaviorally involved in or committed to a learning activity or goal.

Self-efficacy for learning and performance

Self-efficacy is defined as an individual’s beliefs about his or her abilities with regard to accomplishing a task. It is not concerned with the amount or quality of the skills one possesses, but rather with what a person believes he or she can achieve with the skills that he or she possesses (Bandura, 1977). Schunk and colleagues (2013) explained self-efficacy as “one’s perceived capabilities for learning or performing actions at designated levels” (p. 379). The current study examined the relationship between students’ self-efficacy with regard to learning a course’s content and their academic performance in that course. In other words, the researchers were interested in determining whether students over- or under- estimated their self-efficacy regarding the specific content of their courses, irrespective of the actual knowledge and skills that they possessed prior to and during their time enrolled in the courses. Specifically, the study examined whether there were differences between students’ perceived self-efficacy for the subject matter being taught and their ultimate performance based on the types of feedback devices used (i.e., clickers versus web-based polling).

Based on a review of previous literature regarding clickers, polling, JiTT, PoTT, student engagement, and self-efficacy for learning and performance, this study sought to examine how incorporating different electronic feedback devices (i.e., clickers versus web-based polling) may affect specific types of student engagement (i.e., behavioral, emotional, and cognitive engagement), whether students’ self-efficacy for learning and performance may differ between courses that have integrated clickers and those that use web-based polling, and whether using web-based polling influences faculty members’ instructional practices. A brief conceptual model of the current study is presented in Figure 1.

Experimental design

The current study utilized a quasi-experimental design. To capture differences between the effects of using clickers and web-based polling, data were collected from students enrolled in classes at a large research university in the southwestern U.S. The classes were as follows: (1) an educational psychology class in which the instructor used clickers during the lecture (the clicker group; control group) and (2) all other classes using PollEverywhere.com (the web-based polling group; experimental group).
Participants

Of the 209 students who participated in the survey, 45.5% \((n = 95)\) were in the clicker group, and 54.5% \((n = 114)\) were in the web-based polling group. Female students \((n = 128)\) represented 61.5% of the participants in this study. The mean age was 23.18 years \((SD = 7.07)\). All students in the clicker group were undergraduates, whereas the majority of students (64.9%) in the web-based polling group were graduate students. Although this may be a limitation of the current study, given that class levels may influence the results, this effect was minimized by the university’s grading policy, which included standardized grading criteria. Additionally, the content of all of the classes in the experimental group was related to education (e.g., educational psychology, health education, and music education), and all of the instructors were full-time professors at the university with a doctorate in their field. All courses used a minimal amount of lecture material and emphasized student participation either via question-answer format or through planned small group or individual activities. Finally, the Levene’s statistic revealed that the significance values for the engagement subscales, for total engagement, and for self-efficacy were all greater than .1, indicating that the population variances in this study were equal.

Treatments and instructional design

Figure 2 presents the experimental design for this study. The instructor in the control group conducted in-class clicker activities to gain immediate feedback from the students. The instructors in the experimental (web-based polling) group were asked to post three to five questions on PollEverywhere.com six days before class each week to facilitate the possible effect of the PoTT strategy. Figure 3 presents a screenshot of web-based polling and its result. An example question is as follows: “Read the following goal: Universal Coffee employees will know how to make vanilla café lattes with 100% accuracy. Define the level of this goal.” When creating questions, instructors in both the control and experimental groups were encouraged to incorporate the four knowledge dimensions and six cognitive process dimensions that have been identified by Bloom’s Taxonomy Structure (Anderson & Krathwohl, 2001). The former includes (1) factual knowledge, (2) conceptual knowledge, (3) procedural knowledge, and (4) meta-cognitive knowledge, and the latter includes (1) remembering, (2) understanding, (3) applying, (4) analyzing, (5) evaluating, and (6) creating. The instructors in the experimental group re-poll the same or similar questions during their lectures to measure for differences in responses. Both the control and experimental groups’ instructors reviewed the polling answers (via clickers or PollEverywhere.com). If over 75% of students chose the correct answer, the instructors asked for a brief justification. However, if a significant portion of the class (more than 30% of the students) did not choose the correct answer, the instructor initiated a discussion to understand and address the students’ misconceptions. All activities were logged to measure the instructors’ success when revising the course content based on the polling data.
Survey data were collected from students in the eighth week of the semester via an online survey (for the education psychology class; the clicker group) and with a paper-and-pencil survey (for all other classes; the web-based polling group). The survey format (online or paper-and-pencil) was determined based on the participating instructors’ preferences. At the end of the semester, a focus group meeting was held with all of the participating instructors, and all of the participating students completed an anonymous open-ended questionnaire to further explore both faculty and students’ perspectives regarding how web-based polling influenced the level of instruction and to gauge all participants’ overall satisfaction with the quality of the two different types of electronic feedback devices.

**Instrumentation**

The instruments used in this study were adapted from existing validated scales as follows: the Self-Efficacy for Learning and Performance scale from the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich, Smith, Garcia, & McKeachie, 1991) and the Engagement Scale (Fredricks, et al., 2005). The MSLQ was developed by the National Center for Research for Improving Postsecondary Teaching and Learning at the University of Michigan in 1986. The subscale called self-efficacy for learning and performance in this instrument was used to measure students’ self-efficacy in the context of the polling environment. Fredricks and colleagues (2005) developed the Engagement Scale to measure the following three types of engagement: behavioral engagement, emotional engagement, and cognitive engagement. Given that the Engagement Scale was designed to measure children’s levels of school engagement, some of the items had to be modified to measure the engagement levels of graduate and undergraduate students; for example, the item “I follow the rules at school” was revised to “I am compliant with the university’s standards of behavior.” All of these scales used a 6-point Likert rating (6 = strongly agree, 5 = agree, 4 = somewhat agree, 3 = somewhat disagree, 2 = disagree and 1 = strongly disagree).

The internal consistency coefficients (Cronbach’s α) were computed to identify the reliability of the various scales. The results were as follows: .919 for the self-efficacy for learning and performance scale; .873 for the overall engagement scale; .519 for the behavioral engagement scale; .917 for the emotional engagement scale; and .840 for the cognitive engagement scale. Given that the internal consistency coefficient (Cronbach’s α) of the behavioral engagement scale presented an unacceptable value (α = .519), exploratory factor analysis was used to analyze the 19 items on the engagement scale to determine which items loaded onto which types of engagement. The results of the exploratory factor analysis revealed that two factors (emotional engagement and cognitive engagement) most distinctly described the variance in the data. Therefore, the behavioral engagement scale was dropped. As a result, two specific types of engagement (emotional and cognitive engagement) and overall engagement were included in the data analysis.
Results

Data analysis

All of the quantitative data were coded and prepared for computerized analysis using the Predictive Analytics Software (PASW) 17.0 program. Cronbach’s alpha was computed to validate the reliability of each measurement scale. For descriptive statistics, frequencies were computed for the nominal variables, and the means and standard deviations were computed for both the interval and nominal variables in the survey. Factor analysis was used to examine the 19 items of the engagement scale to determine which items loaded onto which types of engagement. Pearson correlation coefficients were computed to examine intercorrelations among the percentages of correct answers for the pre- and end-of-class polling questions and the percentages of the lectures that were modified by the instructors. Finally, t-tests were conducted to examine the differences in the means between the clicker and web-based polling groups.

Research question 1

Are there differences in emotional, cognitive, and overall engagement between courses that integrated the clickers and those that used web-based polling? Independent sample t-tests were conducted to determine whether there were significant differences in emotional, cognitive, and overall engagement between the clicker and the web-based polling groups. The results showed that there were significant differences in emotional engagement ($t(206) = -6.469, p < .001$), cognitive engagement ($t(207) = -6.214, p < .001$), and total engagement ($t(207) = -7.931, p < .001$) between the clicker and the web-based polling groups. Specifically, students in the web-based polling group had significantly higher levels of emotional engagement ($M = 4.60, SD = .986$), cognitive engagement ($M = 4.57, SD = .845$), and overall engagement ($M = 4.75, SD = .610$) than those in the clicker group ($M = 3.71, SD = .987; M = 3.86, SD = .778; and $M = 4.08, SD = .599$, respectively). The results of the t-tests are presented in Table 1.

<table>
<thead>
<tr>
<th>Type of Engagement</th>
<th>Clicker</th>
<th>Web-Based</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Emotional Engagement</td>
<td>3.71</td>
<td>.987</td>
<td>4.60</td>
</tr>
<tr>
<td>Cognitive Engagement</td>
<td>3.86</td>
<td>.778</td>
<td>4.57</td>
</tr>
<tr>
<td>Overall Engagement</td>
<td>4.08</td>
<td>.599</td>
<td>4.75</td>
</tr>
</tbody>
</table>

Note. Clicker = Students in the clicker group; Web-Based = Students in the web-based polling group. *Significantly different at the .05 level. **Significantly different at the .001 level.

Research question 2

Are there differences in students’ self-efficacy for learning and performance between the courses that integrated the clickers and those that used web-based polling? Independent sample t-tests were conducted to determine whether there was a significant difference in students’ self-efficacy for learning and performance between the clicker and the web-based polling groups. Contrary to expectations, students in the web-based polling group had lower levels of self-efficacy for learning and performance ($M = 5.02, SD = .813$) than those in the clicker group ($M = 5.26, SD = .586$). The results of the t-test are presented in Table 2.

<table>
<thead>
<tr>
<th>Type of Engagement</th>
<th>Clicker</th>
<th>Web-Based</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Self-Efficacy</td>
<td>5.26</td>
<td>.586</td>
<td>5.02</td>
</tr>
</tbody>
</table>

Note. Clicker = Students in the clicker group; Web-Based = Students in the web-based polling group; Self-Efficacy = Students’ self-efficacy for learning and performance. *Significantly different at the .05 level. **Significantly different at the .001 level.
To better understand this result, additional t-tests were conducted to consider class standing as a possible contributing variable to students’ level of self-efficacy. The results showed that both within the web-based polling group and among all of the participants, undergraduate participants had significantly higher levels of self-efficacy for learning and performance than graduate participants ($t(111) = 2.863, p < .01$, and $t(206) = 3.989, p < .05$, respectively). Additionally, t-tests were conducted to assess whether there were differences in the levels of self-efficacy between the undergraduates who used clickers and those who used web-based polling. There were no significant differences ($t(132) = -.523, p = .602$).

**Research question 3**

*How does using web-based polling influence instructor’s success when revising the course content?* An analysis revealed that among the students who were in the web-based polling group, the majority (71.7%) responded to the in-class polls via text messaging, whereas most students (95.5%) responded to the pre-class polls via a web browser. On average, 55.57% of the students answered pre-class polling questions correctly, whereas 79.87% provided correct responses when the instructors re-pollled the same or similar questions during the lecture, which suggests that there was a 24.3% improvement after the instructor gauged students’ understanding prior to each lecture and modified the lecture for the purpose of PoTT. On average, the instructors modified 15.8% of the lecture and learning activities prior to the day of class based on the students’ answers to the pre-class questions that were posted on PollEverywhere.com each week.

A summary of the means, standard deviations, and Pearson correlation coefficients for the measured variables are listed in Table 3. Contrary to expectations, the percentage of correct answers for the pre-class questions was not significantly correlated with the percentage of the lectures modified by the instructors ($r = .080, p = .606$). Thus, the amount of lecture modification did not appear to match students’ pre-class polling results. However, results from the focus meetings showed that instructors applied the PoTT technique and helped students address the initial gaps in their learning by modifying the lectures each week, which indicates that our pilot quantitative measure for PoTT may need to be revised to improve its accuracy. Moreover, most instructors believed that posting the same or similar questions at the end of the class provided them with a greater awareness of the amount of knowledge that students retained after each class. Self-reports from the instructors also revealed that they had a positive perception of the web-based polling tools and experienced educational benefits when using the pre- and end-of-class polls, as was evident in the following two instructors’ comments:

> Web-based polling tools facilitated in-class discussion. Asking students questions via PollEverywhere.com before the lecture allowed them to think about the subjects before class and be more curious and eager to figure things out. Most students read the materials ahead so they were more prepared. (Instructor A)

> Students liked the pre-class questions. When they came to the class, they were attentive to find the answers. Web-based polling tools helped develop students’ meta-cognitive awareness! (Instructor B)

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Student Answer Pre</td>
<td>0.56</td>
<td>.19</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>2. Student Answer Post</td>
<td>0.80</td>
<td>.16</td>
<td>-.216</td>
<td>---</td>
</tr>
<tr>
<td>3. Lecture Modified</td>
<td>0.16</td>
<td>.08</td>
<td>.080</td>
<td>-.366*</td>
</tr>
</tbody>
</table>

*Note. Student Answer Pre = Percentage of correct answers for the pre-class questions; Student Answer Post = Percentage of correct answers for the end-of-class questions; Lecture Modified = Percentage of the lecture modified by the instructors based on the students’ answers to the pre-class questions

$p < .05$

**Discussion and implications**

The results of this study both confirm that the use of in-class polls results in higher levels of engagement (Bode et al., 2009; Dallaire, 2011; Lasry, 2008; Stowell & Nelson, 2007; Trees & Jackson, 2007) and suggest that web-based polling, with its use of pre- and post-class questions, and use of the PoTT strategy during class were associated with
higher levels of specific types of student engagement, including emotional and cognitive engagement, than the exclusive use of clickers during class. Given that emotional and cognitive engagement refer to students’ feelings and their voluntary efforts with regard to challenging tasks (Finn & Voelkl, 1993; Fredricks et al., 2004), these results suggest that the use of the PoTT strategy creates an environment that facilitates students’ positive emotions and helps students concentrate on the classroom instruction.

Contrary to expectations, the results showed that the students in the clicker group had significantly higher levels of self-efficacy for learning and performance, which is an important learning factor in academic settings (Bandura, 1977), than those in the web-based polling group. However, the fact that there were no differences in the levels of self-efficacy between the undergraduates who used clickers and those who used web-based polling implies that the higher levels of students’ self-efficacy in the clicker group may not necessarily have resulted from the use of the clickers. The difficulty level of graduate courses may be higher than that of undergraduate courses, resulting in lower levels of self-efficacy for learning and performance for the graduate participants in this study. One additional consideration is that the graduate students had more time to read, think about, and respond to web-based polling, whereas the undergraduate clicker group responded in real time to the questions that were posed in class.

Pedagogically speaking, instructors may want to use in-class polling, such as with clickers, to review factually based information or for discussing attitudes and beliefs about a topic using a Likert-scale. In these cases, the cognitive processing demands are not as taxing given that students are recalling factual or conceptual knowledge or have most likely already formed opinions for Likert-scale polls. Conversely, questions that are more cognitively demanding, such as those that ask students to analyze or evaluate a concept (Anderson & Krathwohl, 2001), may be better disseminated through a web-based polling system to allow students more time to process the information.

According to the information gathered from the faculty focus group, the instructors reported favorable experiences with web-based polling tools with regard to gauging students’ understanding before the lecture, although, according to our pilot quantitative measure, the instructors did not appear to modify their lectures accordingly. This may be due to the novel use of web-based polling and PoTT, in that the instructors may not have felt equipped to modify content or instruction based on the students’ feedback. Instructors noticed that students tended to be better prepared for class as a result of the pre-class polls. This finding indicates the possible benefit of using web-based polling tools to increase students’ motivation and their knowledge retention of class materials, which was demonstrated by the results of the quantitative measure and the faculty focus group meeting. Given that previous research has shown little to no relationship between polling and academic performance (Anthis, 2011; Elicker & McConnell, 2011; Kennedy & Cutts, 2005; Stowell & Nelson, 2007), the findings of the current study suggest that instructors who wish to incorporate polling strategies should expect to see benefits with regard to students’ motivation and their learning factors, but not in their academic outcomes. One significant implication for instructors is that the use of the PoTT approach demands the willingness to modify and change instruction to meet the students’ needs. Although course syllabi are created prior to the instruction, students’ learning of the content may fall behind or pass the indicated pace of the syllabus. This means that instructors need to be more flexible in the ultimate scope and sequence of a course and understand that web-based polling in advance of a class may reflect a wide range of students’ understanding.

Finally, although instructors can use clickers to facilitate Just-in-Time Teaching (JiTT), gauging students’ understanding with in-class clicker polls, simultaneously modifying lectures may result in instructors’ cognitive overload (Sweller, 1988; van Merrienboer & Sweller, 2005), which may in turn affect their teaching performance (Feldon, 2007). Plenty-of-Time Teaching (PoTT) takes advantage of web-based polling tools by allowing instructors to gauge students’ understanding before the actual class, which in turn provides them with sufficient time to modify their lectures.

Conclusion

Given that educational technologies change rapidly over time, the results of this study signify the role of polling tools, provide innovative ways to implement JiTT or PoTT, such as gauging students’ understanding with pre-class polls, and offer insights that can benefit educators who wish to promote students’ emotional and cognitive engagement with various types of feedback devices. Future research should utilize in-depth qualitative data analyses to improve the validity of the findings and make it possible to understand the learning processes and benefits of polling strategies from the students’ perspective.
Acknowledgements

This study is supported in part by the National Science Council of the Republic of China under contract numbers NSC 100-2511-S-009-012 and NSC 101-2511-S-009-010-MY3 and the Center for Scholarly Technology at the University of Southern California. The authors would like to thank the instructors and the students who participated in this study.

References


The Relationships among Chinese Practicing Teachers’ Epistemic Beliefs, Pedagogical Beliefs and Their Beliefs about the Use of ICT

Feng Deng¹, Ching Sing Chai¹, Chin-Chung Tsai² and Min-Hsien Lee³*

¹ National Institute of Education, Nanyang Technological University, Singapore // ² Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan // ³ Center for Teacher Education, National Sun Yat-sen University, Taiwan // feng.deng@nie.edu.sg // chingsing.chai@nie.edu.sg // cctsai@mail.ntust.edu.tw // leemh@mail.nsysu.edu.tw

* Corresponding author

(Submitted December 26, 2012; Accepted June 10, 2013)

ABSTRACT
This study aimed to investigate the relationships among practicing teachers’ epistemic beliefs, pedagogical beliefs and their beliefs about the use of ICT through survey methodology. Participants were 396 high school practicing teachers from mainland China. The path analysis results analyzed via structural equation modelling technique indicated that the systemic relationships among these three types of beliefs were nested. Specifically, teachers’ sophisticated beliefs about the source of knowledge were aligned with constructivist pedagogical beliefs and constructivist use of ICT, with one belief highly related to another.

Keywords
Teachers’ beliefs, Information and communication technology (ICT), Constructivism

Introduction
The developments of information and communication technology (ICT) have brought forth many affordances that can be of great pedagogical relevance (Jonassen, Howland, Marra, & Crismond, 2008; Tsai, 2001). In the hands of skillful teachers, these affordances could enable constructivist-oriented student-centered learning, which is the focus of many current education reforms. However, while access to ICT both in school and at home has improved considerably in most developed countries, and teachers have been reported to be using ICT more frequently (Greenhow, Robelia, & Hughes, 2009), the change in pedagogical practices has yet to be established. Many research studies continue to highlight the challenges in reforms towards ICT supported constructivist teaching which include the need for time to prepare ICT-based lessons, the examination requirements, and the teachers’ personal beliefs (Ertmer, 2005; Hermans, Tondeur, van Braak, & Valcke, 2008). In recent years, researchers tended to use the belief profiles or beliefs system to examine teachers’ various aspects of educational belief. For example, regarding to teachers’ science teaching, Tsai (2002) found that teachers’ beliefs about teaching science, learning science, and their understanding about nature of science are closely aligned, and he called these closely aligned beliefs as “nested epistemologies.” Such studies may help us to understand how teachers’ various beliefs intertwined and also provide potential insight for teacher educators when they design courses. In the field of historical education, Stoddard (2010) found that teachers’ epistemic beliefs affect their pedagogy with historical media through qualitative study. However, in the field of ICT enhanced instruction, studies regarding how teachers’ belief systems may be related to their pedagogy of integrating ICT are still few. Teaching with ICT, teachers’ beliefs about knowing and knowledge (i.e., epistemic beliefs), teaching (pedagogical beliefs), and ICT use (beliefs about the ICT use) may play important roles in the ICT use in the classroom. In this paper, we study the relationships between Chinese practicing teachers’ epistemic beliefs, pedagogical beliefs and their views about the constructivist or traditional use of ICT. The paper aims to contribute to the literature by examining this set of beliefs through structural equation modeling (SEM).

Most studies on how epistemic beliefs are related to other forms of beliefs and how various types of beliefs are connected have been proposed by researchers, especially among practicing teachers in Western countries (Hofer, 2010). In addition, reports of Eastern Chinese practicing teachers’ beliefs, especially for teachers’ epistemic beliefs, are arguably rare (e.g., Chai, Deng, Qian, & Wong, 2010; Sang, Valcke, van Braak, & Tondeur, 2009; 2010). Hofer (2008) and Feucht and Bendixen (2010) argued that personal epistemology is sensitive to the cultural context. Recent research studies on Chinese teachers’ epistemic beliefs were targeted at preservice teachers (Chai et al., 2010). Sang et al.’s (2009; 2010) studies on Chinese teachers were targeted at primary school teachers’ and preservice teachers’ pedagogical beliefs and their relationships with teachers’ views about the use of ICT. However, very few studies have investigated the relationships among teachers’ epistemic beliefs, pedagogical beliefs, and their views about the
use of ICT, especially by involving Chinese high school teachers. In mainland China, the teaching practices in high school generally differ from those in either primary school or college, at least in terms of the class size, the number of subjects taught, the pressure from the national examinations, the professional development of the teachers themselves, etc. This study thus examines Chinese high school teachers’ epistemic and pedagogical beliefs and how they are related to their beliefs about ICT use. It may also expand our current understanding of the various aspects of personal beliefs, given the differences in the cultural context and how these beliefs are connected.

**Literature review**

In the following sections, we provide a brief overview of teachers’ beliefs and how relevant aspects of teachers’ beliefs may be connected to the use of ICT, either for constructivist teaching or traditional teaching. The review helps us to form the hypothesized model of how these beliefs are interrelated, which is tested empirically using SEM.

**Teachers’ epistemic beliefs**

Teachers’ beliefs have been an important area of research for teacher educators as they are generally viewed as the core mental structure that influences what teachers learn and how teachers make instructional decisions (Nespor, 1987). Owing to their complex nature, there are numerous ways of conceptualizing teachers’ beliefs. For example, Woolfolk-Hoy, Davis and Pape (2006) articulated an ecological model that organizes the different aspects of teachers’ beliefs into broad categories of self (e.g., self-efficacy), immediate context (e.g., beliefs about students), state and national context (e.g., beliefs about assessment), and cultural norms and values (e.g., beliefs about the meaning of schooling). How the many aspects of teachers’ beliefs are interconnected structurally to form a belief system that facilitates teachers’ decisions has continuously attracted researchers’ attention to this day (Nespor, 1987; Sang et al., 2010).

Among the myriad beliefs, we argue that teachers’ epistemic beliefs are one of the core beliefs that need to be taken into consideration when studying teachers’ beliefs in the current context of education reforms, which emphasize cultivating the ability of knowledge creation (Bereiter & Scardamalia, 2006). This is essentially a shift in the epistemology and it calls for a constructivist/relativist epistemological stance (Windschitl, 2002). The study of personal epistemology or epistemic beliefs from a psychological perspective is concerned with individuals’ views about the nature of knowledge and knowing (Hofer & Pintrich, 1997). Common dimensions for the nature of knowledge include certainty and simplicity of knowledge. As for the nature of knowing, viewing knowing as receiving knowledge passed down by authorities and experts, as contrasted against viewing knowing as constructing understanding based on personal experience and contextual evidence, is commonly employed as the key dimension (Chan & Elliott, 2004; Feucht & Bendixen, 2010). Generally, the epistemic positions that one adopts can be broadly classified as dualist, multiplist/relativist and evaluativist (Greene, Azevedo, & Torney-Purta, 2008). The dualist sees knowledge as either true or false and it mainly comes from authority such as experts or textbooks. A relativist, however, sees knowledge as evolving in nature, constructed by self, and anyone’s claim of knowledge may be as good as another. The evaluativist sees knowledge as constructed by the self and assesses knowledge claims based on contextual evidence (Feucht & Bendixen, 2010).

Studies of teachers’ epistemic beliefs, which are more common among preservice teachers, indicate that most teachers hold a relativist position, with fewer teachers holding dualist/naïve or evaluativist/sophisticated beliefs. This pattern is common to both preservice teachers in the East and the West (Cheng, Chan, Tang, & Cheng, 2009). Schraw and Olafson (2002) surveyed 24 Caucasian teachers (22 females) and they reported that 23 of them adopted a contextualist epistemological worldview, which is akin to the evaluativist position. Other studies of practicing teachers’ epistemic beliefs are mostly carried out from a domain-specific perspective, in conjunction with their practice of instruction. They are reported together with teachers’ pedagogical beliefs in a later section.

Recent research of Chinese preservice teachers from Guangzhou indicates that while they are disinclined towards relying on authority/expert as source of knowledge, their views on certainty of knowledge are largely neutral (Chai et al., 2010). For the authority/expert aspect, their views are similar to those of their Asian counterparts from Hong Kong and Taiwan. For the aspect of certainty of knowledge, they are different from those of Hong Kong and Taiwanese preservice teachers who view knowledge as clearly evolving (Chai, Hong, & Teo, 2009; Cheng et al.,
2009). Specifically, they expressed a “compromise” viewpoint that knowledge is changeable but also highly stable (e.g., Good & Shymansky, 2001). With regards to the epistemic profile of Chinese practicing teachers, we were unable to locate any literature.

To a certain extent, how teachers view knowledge is related to how they see teaching (Maggioni & Parkinson, 2008; Sinatra & Kardash, 2004; Tsai, 2006). Teachers’ pedagogical views of teaching are reviewed next.

**Teachers’ pedagogical beliefs**

Studies of teachers’ pedagogical beliefs generally portray these beliefs as either traditional or constructivist. The traditional pedagogical beliefs are characterized by teacher-centeredness with its theoretical foundations more associated with behaviorism. The constructivist pedagogical beliefs, on the other hand, are characterized by child-centeredness with constructivism and social constructivism as its theoretical grounding (e.g., Chan & Elliott, 2004; Sinatra & Kardash, 2004). Teachers holding the traditional beliefs tend to see teaching as a process of transmitting knowledge to the students. To achieve that, they assume control of the classroom environment as well as students’ behavior and the instructional content and sequences. They act as the authority to assess the correctness of students’ learning outcomes. The students are treated as passive recipients of verified knowledge. Teachers who are inclined towards constructivist teaching would see teaching as a process of facilitating students’ construction of meaning and understanding of the phenomena they experience. These teachers structure the learning environment to promote active sense-making among the students and they are responsive rather than prescriptive in deciding what and how to learn. Assessment in such a setting is formative rather than summative. While the above literature shows the popularity of conceptualizing teaching in a dichotomous way, which seems to be theoretically consistent, empirical studies challenge such dichotomous views.

Many studies showed that while some teachers’ pedagogical beliefs can be classified as either traditional or constructivist, many teachers are reporting eclectic beliefs (Cheng et al., 2009; Sang et al., 2009; Tondeur, Hermans, van Braak, & Valcke, 2008). In particular, both Sang et al. (2009) and Tondeur et al. (2008) employ cluster analysis and report four broad groups of teachers with different belief profiles among Chinese and Flanders elementary school teachers. They are: constructivist; constructivist and traditional; traditional; and neither constructivist nor traditional. Recent developments in the study of personal beliefs suggest that there is a possibility that teachers could see beliefs as cognitive resources and activate different forms of beliefs for different contents and students to be taught (Maggioni & Parkinson, 2008).

The connections between teachers’ epistemic beliefs and their pedagogical beliefs have been investigated by several researchers. Generally, when teachers view knowing as an accumulation of facts, they tend to adopt the traditional pedagogy. On the other hand, when teachers see knowing as construction of progressive understanding, they are more inclined towards the constructivist approach (Maggioni & Parkinson, 2008). Most studies investigating teachers’ epistemic beliefs and their instructional approaches are qualitative in nature and they adopt domain specific epistemology (Maggioni & Parkinson, 2008; Tsai, 2007). Quantitative studies adopting the domain general approach to the study of personal epistemology and its relation with teachers’ pedagogical beliefs are more common among preservice teachers. Chan and Elliott (2004) and Wong, Chan and Lai (2009) reported that preservice teachers who believe in authority as sources of knowledge and certainty of knowledge tended to believe in traditional teaching. In short, most research indicates that naïve epistemic beliefs are associated with traditional pedagogical beliefs, while sophisticated epistemic beliefs are more associated with constructivist views of teaching. However, Cheng et al.’s (2009) study indicates that preservice teachers who view knowledge as uncertain but believe that authority can transmit knowledge can also be inclined towards constructivist teaching.

**Teachers’ beliefs about the use of ICT**

As a pedagogical tool, ICT can be deployed under different teaching approaches. Many electronic drill-and-practice and tutorial software packages have been created to support traditional teaching (Chen, 2010). Application software such as spreadsheets and discussion forums, on the other hand, can be used as cognitive tools to support the constructivist/social constructivist approach in teaching (Jonassen et al., 2008). Ertmer (2005) suggest that teachers’ use of ICT is influenced by their beliefs. Research to date indicates that teachers may use ICT in both ways (Tubin,
There are studies that associate teachers who are inclined towards constructivist teaching with use of ICT in support of students’ sense making (Hermans et al., 2008). However, constructivist-oriented teachers also seem to support the traditional use of ICT (Chai, 2010; Lee, Chai, Teo, & Chen, 2008; Peeraer & Van Petegem, 2011). Sang et al.’s (2010) recent investigation indicates that Chinese preservice teachers’ constructivist beliefs predict their computer self-efficacy and positive attitudes towards ICT, which in turn predict prospective use of ICT. Constructivist beliefs also have a direct effect on the preservice teachers’ prospective use of ICT. However, for Sang et al.’s study, the scale depicting prospective ICT use merges both constructivist and traditional use of ICT. Teacher use of ICT to support traditional teaching may run contrary to the intended outcomes of education reforms.

In summary, it seems reasonable to assume that teachers holding a relativistic epistemic belief may be more inclined towards the constructivist pedagogy. They also tend to use the computers as cognitive tools to support students’ personal construction of knowledge. On the other hand, teachers who view knowledge as certain and unchanging and that knowing can be achieved through knowledge transmission are more likely to practice traditional teaching. They would use computers, mainly the tutorial software and drill-and-practice packages, to support students’ acquisition of knowledge (Chen, 2010; Jonassen et al., 2008). While such depiction may be theoretically sound, in practice, other contextual variables may interact with and influence the teachers’ beliefs. The teaching practices that are manifested in the classroom are the outcome of the complex interplay between the various forms of beliefs and contexts.

Aim of the present study

This study aims to study the profile of Chinese practicing teachers’ beliefs about knowledge, pedagogy and their preferred use of ICT. It will also examine the relationships among Chinese teachers’ epistemic beliefs, pedagogical beliefs and their preferred use of ICT to support either traditional teaching or constructivist teaching. A concurrent examination of these beliefs is to our knowledge very rare. In relation to the literature reviewed, the hypothesized framework is shown in Figure 1.

Figure 1 illustrates the hypothesized relationships among the teachers’ epistemic beliefs, pedagogical beliefs and their beliefs about the use of ICT. Similar to Tsai’s (2002) “nested epistemologies,” this study also suggested that, regarding to teaching with ICT, teacher’s beliefs about knowledge, pedagogy and their preferred use of ICT were closely aligned.

Method

Participants

A total of 396 high school teachers (Chinese, English, Mathematics, Sciences, Arts, Social Science subjects) from Guangdong province of China participated in this study. About 60% of the participants were female. Of these teachers, 106 had taught for less than 5 years; 101 between 5-10 years; 93 between 10-15 years; and 96 of them more than 15 years. The sample can be regarded as being comprised of beginning and experienced teachers. The samples were selected based on convenience. After obtaining the permission from the three school principals through personal connection, the authors went to the school to distribute the questionnaires. Participation was voluntary and the response rate was around 90%. Regarding the ICT provision issue, every classroom in the three high schools is equipped with one computer for teachers’ uses. Similar to many other regions in China, the participants have been greatly encouraged to integrate ICT into their daily teaching practices.
Procedure and Instruments

A composite paper-based survey questionnaire in Chinese was created and the teachers were invited to respond to it. The translation of the items was subjected to review by Chinese language experts. Demographic data such as gender, age, subject and level taught were asked, but the survey was anonymous. The items were scored based on 5-point Likert scales with 1 indicating “strongly disagree” and 5 indicating “strongly agree”. Higher scores indicate that the participants are inclined to agree with the subscales. The scales involved are:

**Teachers’ epistemic beliefs scale (TEBS)**

This scale consists of two subscales: authority/expert knowledge (AEK) and certainty of knowledge (CK). Six items were selected for AEK (e.g., “Advice from experts should not be questioned”) while the CK dimension has 3 items (e.g., “Scientific knowledge is certain and does not change”). These items, which were created by Chan and Elliott (2004) were selected based on reported studies for Chinese pre-service teachers (Chai et al., 2010). Higher scores in these two subscales suggested more dualist-aligned and less evaluativist-oriented epistemic beliefs. On the contrary, lower scores indicated a more evaluativist or relativistic position for epistemic beliefs.

**Teachers’ pedagogical beliefs scale (TPBS)**

The TPBS measures teachers’ pedagogical beliefs in terms of Traditional Pedagogical Beliefs (TPB) (5 items) and Constructivist Pedagogical Beliefs (CPB) (5 items). The items were created initially by Chan and Elliott (2004). Only 10 items were selected from the original 30 items based on the two main reasons. First, these 10 items were reported to demonstrate satisfactory reliability by a previous study (Chai, 2010). Second, it has been suggested about 4-5 indicators are enough for one latent construct (e.g., Kline, 2005). Sample items in the TPB and CPB subscales are “Learning occurs primarily through drill and practice” and “The ideas of students are important and should be carefully considered” respectively.

**Instructional use of ICT scale (IUICTS)**

The IUICTS represents two ways of using ICT in the classroom: The traditional use of ICT (TUI) and the constructivist use of ICT (CUI). Comprising nine items, there are five items of how teachers would employ ICT in a traditional manner (e.g., “I prefer to use ICT to remediate skills not learned well”) and four items in a constructivist manner (e.g., “I prefer to use ICT to analyze information critically”). These nine items were selected from a previous study that had documented satisfactory psychometric properties (Chai, 2010).

**Reliability of the scales**

To obtain the factor loadings, a principal axis factoring analysis was conducted. As shown in Table 1, the eigenvalues of all constructs were greater than 1.00 and these six constructs jointly explained about 69.02% of the cumulative variance. The factor loadings (in bold) of all the items exceed .50, and this indicates an acceptable convergent validity at the item level. As shown in Table 2, the Cronbach’s alpha values for the scales were all above .70, indicating adequate internal consistency of all scales.

An average variance extracted (AVE) of .50 or higher, or a CR of .70 or above, can be a good rule of thumb suggesting adequate convergence at the construct level (Hair et al., 2006). According to Table 2, all scales consistently demonstrate satisfactory convergent validity.

Considering that very few cross-loadings were observed, a satisfactory level of discriminant validity at the item level was established. At the construct level, it is regarded as adequate when the square root of the average variance extracted (AVE) for a specific construct is greater than the correlation estimates between that construct and all other constructs (Fornell & Larcker, 1981). Table 3 shows the correlation matrix for the six constructs. The diagonal elements (i.e., square roots of AVE) were larger than the off-diagonal elements (i.e., correlation coefficients) in the
corresponding rows and columns. This demonstrates that each construct shared more variance with its items than it
does with other constructs. That is, discriminant validity was also acceptable at the construct level.

### Table 1. Results of principal axis factor analysis

<table>
<thead>
<tr>
<th>Factor</th>
<th>CPB</th>
<th>AEK</th>
<th>TUI</th>
<th>TPB</th>
<th>CK</th>
<th>CUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPB4</td>
<td>.895</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPB5</td>
<td>.887</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPB2</td>
<td>.805</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPB1</td>
<td>.688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPB3</td>
<td>.683</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AEK2</td>
<td></td>
<td>.879</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AEK4</td>
<td></td>
<td>.842</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AEK3</td>
<td></td>
<td>.831</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AEK6</td>
<td></td>
<td>.812</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AEK1</td>
<td></td>
<td>.799</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AEK5</td>
<td></td>
<td>.784</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUI2</td>
<td></td>
<td></td>
<td>.910</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUI3</td>
<td></td>
<td></td>
<td>.907</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUI1</td>
<td></td>
<td></td>
<td>.802</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUI5</td>
<td></td>
<td></td>
<td>.798</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUI4</td>
<td></td>
<td></td>
<td>.769</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPB3</td>
<td></td>
<td></td>
<td></td>
<td>.830</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPB2</td>
<td></td>
<td></td>
<td></td>
<td>.804</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPB4</td>
<td></td>
<td></td>
<td></td>
<td>.799</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPB1</td>
<td></td>
<td></td>
<td></td>
<td>.727</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPB5</td>
<td></td>
<td></td>
<td></td>
<td>.682</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CK2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.939</td>
<td></td>
</tr>
<tr>
<td>CK3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.899</td>
<td></td>
</tr>
<tr>
<td>CK1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.796</td>
<td></td>
</tr>
<tr>
<td>CUI4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.828</td>
</tr>
<tr>
<td>CUI3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.822</td>
</tr>
<tr>
<td>CUI2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.691</td>
</tr>
<tr>
<td>CUI1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.628</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>6.585</td>
<td>5.366</td>
<td>3.671</td>
<td>2.289</td>
<td>2.004</td>
<td>1.203</td>
</tr>
<tr>
<td>% Variance</td>
<td>22.377</td>
<td>18.087</td>
<td>12.113</td>
<td>6.940</td>
<td>6.341</td>
<td>1.203</td>
</tr>
</tbody>
</table>

Notes. Extraction method: principal axis factoring; rotation method: Oblimin with Kaiser Normalization. AEK: Authority/Expert Knowledge; CK: Certainty of Knowledge; TPB: Traditional Pedagogical Beliefs; CPB: Constructivist Pedagogical Beliefs; TUI: Traditional Use of ICT; CUI: Constructivist Use of ICT. Factor loadings less than .50 were not presented.

### Table 2. Index of reliabilities for all scales

<table>
<thead>
<tr>
<th>Factor</th>
<th>α</th>
<th>AVE</th>
<th>CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authority/Expert Knowledge (AEK)</td>
<td>.93</td>
<td>.68</td>
<td>.93</td>
</tr>
<tr>
<td>Certainty of Knowledge (CK)</td>
<td>.91</td>
<td>.77</td>
<td>.91</td>
</tr>
<tr>
<td>Traditional Pedagogical Beliefs (TPB)</td>
<td>.88</td>
<td>.59</td>
<td>.88</td>
</tr>
<tr>
<td>Constructivist Pedagogical Beliefs (CPB)</td>
<td>.92</td>
<td>.64</td>
<td>.90</td>
</tr>
<tr>
<td>Traditional Use of ICT (TUI)</td>
<td>.92</td>
<td>.70</td>
<td>.92</td>
</tr>
<tr>
<td>Constructivist Use of ICT (CUI)</td>
<td>.87</td>
<td>.56</td>
<td>.83</td>
</tr>
</tbody>
</table>

Notes. AVE: Average Variance Extracted=(∑λ²)/n; CR: Construct Reliability=(∑λ²)/(∑λ²+∑(1-λ²))

### Table 3. Inter-factor zero-order correlations (2-tailed)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.Authority/Expert Knowledge (AEK)</td>
<td>(.83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.Certainty of Knowledge (CK)</td>
<td>.18**</td>
<td>(.88)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Traditional Pedagogical Beliefs (TPB)</td>
<td>.41**</td>
<td>.08</td>
<td>(.77)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.Constructivist Pedagogical Beliefs (CPB)</td>
<td>-.04</td>
<td>.26**</td>
<td>-.17**</td>
<td>(.80)</td>
<td></td>
</tr>
</tbody>
</table>
5. Traditional Use of ICT (TUI)
6. Constructivist Use of ICT (CUI)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>No. of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEK</td>
<td>2.00</td>
<td>.85</td>
<td>1.10</td>
<td>.92</td>
<td>6</td>
</tr>
<tr>
<td>CK</td>
<td>3.15</td>
<td>1.06</td>
<td>-.16</td>
<td>-.98</td>
<td>3</td>
</tr>
<tr>
<td>TPB</td>
<td>2.31</td>
<td>.86</td>
<td>.61</td>
<td>-.17</td>
<td>5</td>
</tr>
<tr>
<td>CPB</td>
<td>4.24</td>
<td>.78</td>
<td>-.74</td>
<td>3.72</td>
<td>5</td>
</tr>
<tr>
<td>TUI</td>
<td>2.90</td>
<td>.97</td>
<td>.16</td>
<td>-.89</td>
<td>5</td>
</tr>
<tr>
<td>CUI</td>
<td>3.89</td>
<td>.78</td>
<td>-.71</td>
<td>.28</td>
<td>4</td>
</tr>
</tbody>
</table>

Notes. AEK: Authority/Expert Knowledge; CK: Certainty of Knowledge; TPB: Traditional Pedagogical Beliefs; CPB: Constructivist Pedagogical Beliefs; TUI: Traditional Use of ICT; CUI: Constructivist Use of ICT.

In general, the above results indicate acceptable convergent validity and discriminant validity at both the item and construct levels. Therefore, the measurements of the six constructs in the proposed research model are considered to be adequate.

Results

The descriptive data of the scales used, such as mean, standard deviation, skewness and kurtosis, are presented in Table 4. No items show a skewness or kurtosis value greater than the cutoffs of |3.0| or |8.0| (Kline, 2005), respectively, and this supported univariate normality in the items. In the following paragraphs, we mainly report (a) the results of testing the measurement model comprising all scales, and (b) the structural model proposed in the study.

Test of the measurement model

According to Anderson and Gerbing (1988), testing the originally specified theory (structural model) may not be meaningful unless the measurement model holds. Thus, a confirmatory factor analysis (CFA) with AMOS 7.0 was conducted to validate the measurement model (see Figure 2) consisting of the aforementioned six scales. The CFA results imply a model fit. The fit indices revealed $\chi^2 = 825.12$, $df = 333$, $\chi^2/df = 2.48$; GFI = .873; TLI = .931; CFI = .939; RMSEA = .061; SRMR = .046. The values of these indices were regarded by most researchers as indicative of a good model fit to the data (Hair et al., 2006; Kline, 2005). All items were found to have significant parameter estimates with standardized estimates greater than .50. This also suggests a satisfactory convergent validity of each scale, since each scale can explain the items it measures better than the items from another scale (Fornell & Larcker, 1981).

Test of the structural model

The test of the structural model includes examining the statistical significance of the path coefficients from one latent variable to another. In the current study, a structural model (see Figure 3) based on the hypothesized framework of this study (i.e., Figure 1) with six latent variables was specified to ascertain the relationships among epistemic beliefs, pedagogical beliefs, and beliefs about the use of ICT. The overall goodness of fit can be considered satisfactory for this structural model ($\chi^2 = 838.74$, $df = 335$, $\chi^2/df = 2.48$; GFI = .872; TLI = .930; CFI = .938; RMSEA = .062; SRMR = .055). As shown in Figure 3, a total of seven path coefficients are significant at the .05 level (AEK→TPB; AEK→CPB; AEK→CUI; CK→CPB; CK→CUI; CPB→CUI; CPB→TUI). These significant path coefficients are also highlighted by the bold lines in Figure 3. The $R^2$ values of TUI and CUI are .028 and .543, respectively. This indicates that the four exogenous variables (AEK, CK, TPB, and CPB) explained 2.8% and 54.3% of the variance of TUI and CUI respectively. Meanwhile, the AEK and CK jointly accounted for 19.6% and 8.2% of the variance of TPB and CPB respectively.
Figure 2. The measurement model among the constructs in the study

Figure 3. Path coefficients of the proposed research model for the constructs considered in this study

* p < .05. ** p < .01. *** p < .001

Notes. Both the bold and dotted lines represent the significant path coefficients. The bold lines indicate the anticipated predictions which supported the hypothesized framework, while the dotted lines also support the hypothesized framework but with alternative predictions.
Discussion and conclusion

Generally, the structural equation model obtained in this study attests that the systemic relationships between the teachers’ epistemic beliefs, pedagogical beliefs and their beliefs about ICT use, to a certain extent, are nested. That is, the model is consistent with other models in two Asian studies (e.g., Chai, 2010; Lee et al., 2008). This implies that while China teachers work in a rather different sociopolitical environment, the underlying cultural beliefs about teaching (e.g., originated from the Confucius philosophy) is shared. Besides, the model also supports findings from other Western studies (e.g., Ertmer, 2005; Maggioni & Parkinson, 2008; Stoddard, 2010). This suggests that the difference between the East and West may be insignificant for the variables we have investigated in this study. The rhetoric for the inclusion of ICT in classrooms is similar regardless of sociocultural differences. This study found that teachers’ sophisticated beliefs about AEK, constructivist pedagogical beliefs (CPB), and constructivist use of ICT (CUI) are aligned, which one belief was highly related to another. As shown in Figure 3, the relationships between teachers’ beliefs about CK, CPB, and CUI were also identified in this study. However, the positive path coefficients between naive CK and CPB and CUI were obtained in this study. This finding indicates that more theorizing and empirical testing of the model of beliefs is necessary, especially for teachers working in different sociocultural contexts (Hofer, 2008). Such work would enable educators to have deeper knowledge of the various ways in which the investigated beliefs are related. In the following paragraphs, we examine and compare the results of this study in more detail.

The profile of AEK we obtained among the participants (mean = 2.00, Table 4) is consistent with previous research on Chinese preservice teachers (mean = 2.20) from the same region (Chai et al., 2010). It also seems to be consistent with Schraw and Olafson’s (2002) study that indicates that many practicing teachers are evaluativist. Both preservice and practicing teachers in China are not inclined to rely on authorities and experts as the sole source of knowledge. The relationships between AEK and the teachers’ pedagogical beliefs are comparable with those reported by Wong et al. (2009) who employed a similar instrument with Hong Kong preservice teachers. However, another study (Chai, Teo, & Lee, 2010) which also employed a similar instrument for Singaporean preservice teachers revealed a different pattern of relationships, with sophisticated AEK being positively related to CUI and negatively related to TUI. The Singaporean preservice teachers were inclined towards AEK (mean = 3.46), which is different from the current study. According to figure 3, the Chinese practicing teachers who believe in AEK are inclined towards traditional pedagogy and are less inclined towards constructivist teaching. AEK is also negatively related to constructivist use of ICT, although its association with traditional use of ICT is insignificant. On the whole, it suggests that AEK is an important variable that is likely to influence teachers’ pedagogical practices related to ICT. Moreover, the results of this study also imply that teachers’ beliefs about knowing, rather than beliefs about knowledge (i.e., CK), may play an important role on their teaching. Teachers believed knowing as constructing understanding based on personal experience and contextual evidence may have constructivist beliefs about teaching and teaching with ICT constructively.

The teachers’ beliefs about CK (mean = 3.15, Table 4) in this study are also consistent with the profile of Chinese preservice teachers obtained by Chai et al. (2010) (mean = 3.33). Chinese practicing teachers’ profiles of AEK and CK, as compared to the Chinese preservice teachers’ profile, seems to suggest that the teachers become more relativistic in their epistemic outlook when they enter the service. In this study, we obtained positive path coefficients between naive CK and CPB and CUI. This may be due to the fact that the teachers’ belief of CK is close to a neutral position. Greene and his colleagues (2008) have pointed out that one problem of adopting a domain general stance in personal epistemology research could be the difficulty involved in interpreting such neutral responses. Chai et al.’s (2010) study reported a similar CK mean score (mean = 3.12) among Singaporean preservice teachers, and a significant negative path coefficient was obtained from CK to TPB with no significant relations between CK and CPB. Wong et al.’s (2009) study, however, reported a lower mean score for CK among Hong Kong preservice teachers (mean = 2.79), and a significant positive path coefficient was obtained from CK to TPB, while a non-significant negative relationship was reported between CK and CPB. All three studies indicated some significant relationships between CK and CPB or TPB, which affirm CK as an important variable within this system of beliefs. As for the relation between teachers’ beliefs about the nature of knowledge and their use of ICT, teacher’s naïve beliefs about CK are positively associated with constructivist use of ICT.
A re-interpretation of the construct CK may provide a possible alternative interpretation of the positive significant relationships that we have obtained between CK and CPB and CUI. The three items of CK are (a) If scientists try hard enough, they can find the truth to almost anything; (b) Scientists will ultimately get to the truth if they keep searching for it; and (c) Scientific knowledge is certain and does not change. Out of the three items, two may be interpreted as the attainability of scientific truth. If this is how the teachers interpreted the items, then the positive path coefficients may indicate that the teachers are in favour of employing the constructivist pedagogy and the use of computers for students to construct scientific models. In other words, the teachers are advocating that constructivist use of computers can help students to attain scientific truth. While such a stance may not be consistent with the epistemic positions of the constructivist philosophy, the teachers may be seeing the items from a pragmatist’s point of view. The dominance of Marxism in China, which advocates that objective scientific laws of nature can be discovered through scientific research, could provide further support for this interpretation (Chai et al., 2010). This also shows that further research adopting the qualitative method is needed to test whether this alternative interpretation corresponds to the participants’ views. It also indicates that further item refinement and development with more precise use of language is needed (see also Hofer, 2010).

The profile of pedagogical beliefs of the Chinese practicing teachers is comparable with the preservice teachers in Asia. In most studies conducted in Asia, the preservice teachers surveyed are inclined to subscribe to constructivist teaching beliefs (CPB) rather than traditional teaching beliefs (TPB) (see Chai et al., 2010; Cheng et al., 2009; Sang et al., 2010; Wong et al., 2009). One possible explanation of such a similarity between preservice and inservice teachers in China may be attributed to the explicit promotion of the constructivist pedagogy in recent years. This however does not seem to be in total agreement with Sang et al.’s (2009) study among Chinese practicing elementary teachers. Sang et al.’s study indicates that around 60% of the teachers they surveyed are inclined towards constructivist or constructivist and traditional pedagogical beliefs. The other 40% are less inclined towards constructivist teaching. The difference may be due to the instrument employed, the different statistical procedures and the grade level that the teachers are teaching (primary vs. secondary). We suggest that the current study could be re-analyzed employing cluster analyses as used by Sang et al. (2009).

The path coefficients obtained supported the relationships between pedagogical beliefs and views of ICT use, and the results are in general agreement with recent studies in that constructivist beliefs are more compatible to promoting the use of ICT (Chai, 2010; Teo, Chai, Hung, & Lee, 2008; Hermans, et al., 2008; Tubin, 2006). In particular, Chai (2010) has also obtained positive significant path coefficients from CPB to CUI and TUI. The current study, along with others reported in the literature, affirms that teachers’ constructivist-oriented pedagogical beliefs are a significantly positive predictor for their use of computers across different cultures. This implies that it is desirable to foster a constructivist view of teaching in teacher professional development activities for the integration of ICT in classrooms. The teachers may see the traditional use of ICT as a means of enhancing students’ grasp of basic facts before they embark on more adventurous knowledge construction activities.

Traditional pedagogical beliefs in this study are not related in any significant way to the use of ICT. Given that traditional pedagogy was developed without any reference to the use of ICT, this is a probable outcome. While traditional pedagogic views are at times reported to be significantly related to traditional use of ICT (Tondeur et al., 2008), studies have also reported negative or no relations between TPB and the use of computers (Hermans et al., 2008; Teo et al., 2008).

In conclusion, this study has contributed to the current literature studying how teachers’ beliefs may play a role in their use of ICT. This study confirms that teachers’ beliefs about knowing, knowledge and teaching are related to their preference of the teaching practice with ICT. The associations between epistemic beliefs about the nature of knowing and knowledge and teachers’ views about the use of ICT have been found to be significant, with AEK negatively associated with CUI and CK positively associated with CUI. It seems that epistemic beliefs are related to the constructivist use of ICT, but they are not related to the traditional use of ICT. The associations between Chinese practicing teachers’ pedagogical beliefs (CPB) and their views of use of ICT (TUI and CUI) have also been established. We suggest that future research can further confirm the findings through both qualitative and quantitative methods. Studies focusing on how contextual variables influence the evolution of teachers’ beliefs are particularly valuable. In addition, given China’s population and size, numerous studies would be necessary to ascertain our findings. Furthermore, we believe that employing a domain-specific approach (such as in the domain of science or language) to the study of teachers’ beliefs could contribute to more refined recommendations for teacher education.
Acknowledgements

Funding for this research work was supported by National Science Council, Taiwan, under grant numbers NSC 100-2511-S-011-004-MY3 and NSC 100-2511-S-110-010-MY3.

References


Developing Learners’ Second Language Communicative Competence through Active Learning: Clickers or Communicative Approach?

Alaba Olaoluwakotansibe Agbatogun
University of Edinburgh, Moray House School of Education, EH8 8AQ, Edinburgh, UK // alabaagbatogun@yahoo.com

(Submitted January 5, 2013; Revised April 12, 2013; Accepted June 21, 2013)

ABSTRACT
The purpose of this study was to compare the impact of clickers, the communicative approach and the lecture method on the communicative competence development of learners who were taught English as a second language (ESL). Ninety nine pupils from three primary schools participated in the study. Quasi-experimental non-randomised pre-test posttest control group design was adopted for the study. A battery of English Language Listening Tests and English Language Speaking Tests were used to measure pupils' communicative competence. Study’s data were analysed using boxplot, paired samples t-test, Analysis of covariance and multiple regression analyses. Findings indicated that, there was a significant difference between the communicative competence pre-test and post-test scores of pupils in each of the groups. Furthermore, across the groups, there was a significant difference in pupils’ communicative competence post-test scores based on the teaching strategy. Multiple regression analysis results revealed that 84.9% of the variance of pupils' communicative competence was accounted for by a combination of the predictor variables. Speaking skills was the potent contributor while gender did not make a significant contribution to the prediction of pupils’ communicative competence in ESL classrooms.

Keywords
Second language, Active learning, Clickers, Interaction, Communicative competence

Background
Different pedagogical strategies have varying degrees of success. Students’ academic performance may be influenced positively by their active engagement in the classroom (Emerson & Taylor, 2004; Johnson, 2005). In developing countries like Nigeria, teacher-talk, and the persistence of triadic initiation-response-feedback (IRF) mode of discourse dominate classroom instructional process (Oluwole, 2008; Onukaogu, 2001). In traditional classrooms, students engage in recitation of scripts, minimal interaction, and less involvement in productive thinking. Interaction between the students, the learning materials, other students, and the teacher are significant to learning outcomes (Singh & Mohammed, 2012; Smith, Hardman & Higgins, 2006).

Second language (L2) learning requires that learners take ownership of learning activities through interaction, active participation and the use of the target language in a more authentic context (Lantolf, 1994; Tabber & deKoeijer, 2010). Despite English being the medium of instruction in Nigerian schools, many students are academic underachievers because of their low level of communicative skills in English caused by teachers’ reliance on the lecture method (Adesemowo, 2005; Oluwole, 2008). The traditional “chalk and talk” method which involves the teacher talking to students and writing notes on the chalkboard results in rote learning, learners’ low level of retention, and passive learning. Onukaogu (2001) remarked that the traditional method of teaching provided learners fewer opportunities to participate actively in class; hence learners are less confident to express themselves.

Interaction is a key element to successful instructional process. According to Singh and Mohammed (2012), knowledge is best constructed when learners involve in negotiation of meaning. In the recent time, most educational theories as exemplified in Figure 1 emphasise social learning and learner-centred learning in knowledge construction. Studies have shown that classroom interaction promotes improved learning outcomes, and critical thinking (Chou, 2003; Kay & LeSage, 2009), and captures students’ attention and interest (Sims, 2003). Individual learning styles influences interaction and participation in the classroom (Debourgh, 2008). There are active learners (learn by doing), sensing learners (learn by discussing possibilities and relationships), visual learners (learn when they see things), and the sequential learners who gain understanding in linear steps (Felder & Spurlin, 2005). The multimedia learning principle of Mayer (2001) proposes that auditory information is less contributory to effective learning than
when text is combined with visual images. Therefore, the multidimensional nature of an interactive and a communicative classroom suits learners of different learning styles.

**Constructivism** relies on the learner selecting and transforming information and making decisions to construct meaning.

**Whole-class teaching** brings the entire class together, focuses their attention and provides structured, teacher-focused group interaction.

**Active learning** learners actively engage in the learning process through reading, writing, discussion, analysis, synthesis and evaluation, rather than passively absorbing instruction (e.g., lecture model of instruction).

Figure 1. Student active engagement in social learning
(Adapted from SMART Technologies, 2006)

The three learning theories in Figure 1 emphasise the importance of student’s active participation in the instructional process (Beeland, 2002; Singh & Mohammed, 2012). Students would be motivated to learn when they are actively engaged in learning activities than they would have when they are passive in the classroom. Ensuring interactivity in the traditional classroom is challenging (DeBourgh, 2008).

In the last two decades, one of the most influencing developments in language learning is the introduction of digital technology. The introduction of interactive teaching approaches into schools has had an increasing impact on the way teacher teach, and the process students learn (Facer, Sutherland, & Furlong, 2003). Communicative approach (CA) is directed towards enhancing classroom interaction and learners’ participation in communication during the instructional process (Menking, 2002; Qinghong, 2009). CA is a classroom strategy that involves pairing and grouping of learners to enhance negotiation of meaning, development of confidence by engaging in tasks and activities that are fluency-based. The role of a CA teacher is more of a facilitator of learners’ task performance because learners do more of the talking than in the traditional classroom. With CA, activities and tasks set up by the teacher include real life situations which involve games, role-playing, simulations and problem-solving.

Some strategies employed to promote learners’ active engagement in a second language (L2) classroom have been criticised. For example, the use of flash cards and students’ thumbs to signify responses have been criticised to lack the privacy that builds students’ confidence in the class (Caldwell, 2007). Moreover, communication between interlocutors is either distorted or interrupted due to low bandwidth and unreliable Internet network when Mobile phones, MP3 players and Smartphones are used for the learning process (Huffman, 2011). However, one technology, which facilitates students’ active engagement during the instructional process, is clickers (Lantz, 2010; Lea, 2008; Wu & Gao, 2011).
Clickers are devices similar to the TV remote control used by the audience to respond to questions on a TV programme known as “1 vs 100”. Clickers provide students the opportunity to answer questions anonymously in class (Caldwell, 2007; Kelly, 2007; Lantz, 2010). Clickers’ handsets transmit students’ responses to the teacher’s questions unto the receiver which is attached to the Universal Serial Bus (USB) port of the teacher’s computer. The device provides immediate feedback as the distribution of students’ responses in the form of a bar graph is displayed on a projection screen (Johnson & Lillis, 2010). Clickers’ questions may be in the form of true/false or yes/no answers, multiple-choice responses, or short answers.

One of the prominent advocates of clickers’ use in teaching and learning is Eric Mazur; who employed the technology for peer-instruction in physics education. Eric Mazur’s peer instruction involves the short presentation of key points, presentation of a Concept Test (short conceptual questions on subject being discussed), allowing students to formulate answers, and providing students the opportunity to discuss their answers with peers (Fies & Marshall, 2006; Mazur, 1996; Simelane & Skhosana, 2012). The essence of concept test is to prompt students’ interaction and critical thinking, as well as assess their understanding of concepts based on peers’ views. Previous research report that clickers do not only wake students from lethargy periodically to answer questions, but trigger learners’ critical thinking and active engagement (Fies & Marshall, 2006; Mintzes & Leonard, 2006). Pedagogical use of clickers encourages self-directed learning (Carnevale, 2005; Duncan, 2006) and sustains students’ attention (Hoffman & Godwin, 2006). Furthermore, clickers have been reported to provide Spanish language learners opportunities for more interactive activities, active engagement, retention and improved learning outcomes (Fritz, n.d; Pennestri, n.d). Recent research supports the effectiveness of clickers among French learners, as a tool that provides immediate feedback and promotes students’ interaction and critical thinking.

**Theoretical framework**

This study was based on the active learning theory. Active learning theory has been well discussed in education, especially with respect to the adoption and integration of technology in the classroom (Hoffman & Godwin, 2006). Active learning is a subfield of machine learning which occurs when a learning algorithm is given access to a pool of unlabelled examples and is also allowed to request the label of specific examples from the pool. By this, the function that perfectly predicts the label of new examples is learned as much as possible in the process of few labels. On the contrary, with passive learning, requested examples are chosen randomly (Hanneke, 2009).

Active learning is anything course-related that all students in a class session are called to do other than simply watching, listening and taking notes. It keeps students awake and provides the opportunity for high-level of learning and retention unlike what happens in the traditional lecture classroom. Authentic communication in the classroom is a basic element of active learning (Felder & Brent, 2009). The theory of active learning can be linked with the quote of the Confucius “I hear, and I forget, I see, and I remember, I do, and I understand” (Braxton, Jones, Hirschy & Hartley, 2008; Nguyen & Trimarchi, 2010). Unlike in the traditional classroom, active learners use more opportunities to decide about aspects of the learning process; they move beyond mere acquisition of information to getting engaged in higher order thinking tasks of analysis, synthesis and evaluation. In this study, the interventions were introduced to stimulate a two-way interaction in the classroom. Rather than being involved in memorisation and regurgitation of sentences, the intervention groups were involved in the development of their speaking skills by talking about what they learned by using the target language during discussions.

**Statement of the problem**

Although the apparent benefits of clickers in teaching and learning is being reported, research on the use of clickers in ESL classrooms, its adoption and integration into Nigerian education system is yet to receive adequate attention because clickers’ adoption is still at the infancy stage. Moreover, the use of clickers at the elementary level of education is not well documented. Majority of research on the effectiveness of clickers in the classroom only compared the use of clickers to the lecture method; this study was undertaken to compare L2 learners’ communicative competence development based on the use of clickers, communicative approach and lecture method.
Research questions

Two hypotheses and a research question were raised to guide this study:
1. The communicative competence pre-test and post-test scores would not be significantly different for pupils in each of the groups (the communicative approach, the clickers and the control groups).
2. The post-test communicative competence scores of pupils across the groups would not be significantly different.
3. What are the relative and joint contributions of gender, classes of pre-test score, listening and speaking skills to pupils’ communicative competence in ESL classrooms?

Method

Design

The study used a pre-test, post-test quasi-experimental non-randomised control group design with two experimental groups exposed to two conditions (the communicative approach and clickers) and a control group taught with the lecture method.

Participants

The population of this study was all primary six pupils in a local government in Ogun State, Nigeria. Multi-stage sampling technique was employed to select three schools. A sample of 99 pupils (10 and 13 years) from the three schools which participated in the study conducted between September, 2010 and April, 2011 was selected based on convenience. There were 32, 41 and 26 pupils in the communicative approach, the clickers and the control groups respectively. All the groups were similar with respect to socio-economic background; school location, language of the environment and school type (public schools). In each group, pupils with pre-test scores below the group’s mean were treated as low pre-test scorers while those with pre-test scores above the group’s mean were treated as high pre-test scorers. The results of the Levenes test of homogeneity of variances for the groups show no significant differences in their English language pre-test scores ($F(2, 96) = .51 > .05$); hence the assumption of homogeneity of variances was not violated.

Instruments

To measure pupils’ communicative competence, performance scores in English Language listening and Speaking Tests (developed by the researcher and a 7-man review committee made up of primary school English teachers) were used. All test items were derived from the content of the English textbook used within the context of the study. The instruments were pilot tested on a sample of pupils who were not involved in the present study. During the pilot-study, all instruments were re-administered to the same sets of pupils as post-test two weeks after the initial administration.

English language listening tests

The English Language Listening Tests (ELLTs) 1, 2, 3 and 4 were administered in all the groups to assess pupils’ listening ability in ESL classroom. Each of the ELLTs comprised of a short comprehension passage and five short questions. The comprehension passages were summaries of selected comprehension passages contained in the pupils’ English textbook. Some of the multiple-choice questions for the passages in the textbook were changed into short answer questions and sentence completion questions. To each comprehension passage question, obtainable marks ranged between 0 and 3. The test re-test reliability of English Language Listening Tests 1, 2, 3 and 4 were .94, .93, .86 and .87 respectively.
English language speaking tests

English Language Speaking Tests (ELSTs) 1 and 2 were administered to assess pupils’ English speaking ability. The English Language Speaking Test 1 consisted of ten items (nine mini-guided-situation and a picture-description test items) while English Language Speaking Test 2 comprised of seven mini-guided-situation and one picture-description test items. All items of the English Language Speaking Tests 1 and 2 were generated to prompt pupils’ use of the target language. Selected exercises in the pupils’ English textbook were changed into sentences that reflect guided real-life situations, which required pupils’ responses. The clickers’ questions were used to trigger interaction and discussion after the first round of voting, so as to provide learners opportunities for speech practice as they argued out their initial ideas with peers before answering the questions again. Testing the pupils’ speaking skills followed the one-to-one interviews. Pupils’ performances in each item were rated on a scale of 0 to 5. The test re-test reliability of English Language Speaking Tests 1 and 2 were .87 and .88 respectively.

Twelve copies of pupils’ listening and speaking scripts were selected as a representative sample of pupils’ responses. The selected scripts were subjected to double blind review by an independent rater and the teacher to each group. The inter-rater reliability coefficient of the raters’ judgement was 0.99. The range of possible scores for listening test was from 0 to 60 and speaking test ranged between 0 and 90. For the overall communicative competence, the range of possible value was from 0 to 150. Obtained marks in all the listening and the speaking test items were summed up to estimate the total communicative competence score for each pupil. Scores above 59 indicate high communicative competence and scores below 59 indicate low communicative competence.

Validity of instruments

All instruments used for this study were reviewed by the review committee, two e-learning lecturers, one quantitative researcher and two English language lecturers before the final drafts were produced.

Equipment

The equipment used for this study was the eInstruction’s clickers. 48-piece of Radio Frequency mode of clickers provided by eInstruction to support the study was distributed and collected at the end of the lesson on a daily basis to pupils in the clickers’ group.

Procedure

Approval to conduct the study in all the participating schools was granted by the Local Government Education Authority and the head-teachers of the schools. Participating teachers and parents of pupils signed the consent form. On the form, parents had the opportunity to decline their child’s participation or thereafter withdraw in the process of the research without the pupils being penalised. 100% of the consent forms were returned with positive responses. In the first week of the study, the purpose of the research was introduced to the pupils before administering all the English language tests to assess pupils’ initial level of language skills. Thereafter, the groups were exposed to different instructional conditions. During the research, the three groups were taken through the same content in oral and written comprehension, composition, and grammar activities.

In the clicker’s group, grades were not attached to pupils’ responses. The clickers were used to trigger learning by doing instruction, and catalyse interaction, negotiation of meanings, and the use of the target language in oral communication during discussions. The clickers’ displayed feedback and oral output in a social context contribute to learners’ improved oral communication. The procedure about the use of clickers in anonymous mode was practised with the pupils before its continued daily use during English lessons. Two or three questions were posed by the teacher during the lesson; students respond through the wireless clickers’ keypads; teacher prompts group or peer discussion after the display of responses with no clue to the correct answer, students respond a second time through the clickers’ keypads, and correct answers are indicated, followed by the teacher’s explanation, comments and contributions.
Pupils in the communicative approach group worked in groups. During lessons, the teacher assigned different tasks of the lesson content and gave instructions on how to accomplish the assigned tasks. The teacher went around prompting meaningful discussions and making clarifications. Groups’ representative(s) presented a summary of their activities to the class. Groups’ discussion were summarised by the teacher on the chalkboard. In the control group, the teacher used the lecture method. Pupils raised their hands to signify their willingness to answer the teacher’s questions. At the eleventh week of the study, the three groups were post-tested with the same sets of tests used at the pre-treatment stage in order to determine the effects of the teaching strategies on pupils’ language skills development.

**Data analysis**

The t-test statistics was used to determine whether a significant difference existed between the pre- and post-test scores of each group while the multiple regression analysis was performed to test the contribution of the independent variables to the prediction of the pupils’ L2 skills development. Analysis of Covariance (ANCOVA) was conducted to control for differences in pre-test scores while analysing the significant differences in the means of groups’ language skills development.

**Results**

**Table 1. Comparison of communicative competence pre-test and post-test scores**

<table>
<thead>
<tr>
<th>Group</th>
<th>Type of Test</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>df</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comm. Approach</td>
<td>Pre-test Scores</td>
<td>32</td>
<td>49.6</td>
<td>21.8</td>
<td>31</td>
<td>-8.982</td>
<td>p &lt; .05</td>
</tr>
<tr>
<td></td>
<td>Post-test Scores</td>
<td>32</td>
<td>69.4</td>
<td>19.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clicker’s</td>
<td>Pre-test Scores</td>
<td>41</td>
<td>61.2</td>
<td>24.3</td>
<td>41</td>
<td>-11.232</td>
<td>p &lt; .05</td>
</tr>
<tr>
<td></td>
<td>Post-test Scores</td>
<td>41</td>
<td>88.1</td>
<td>17.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>Pre-test Scores</td>
<td>26</td>
<td>67.8</td>
<td>21.4</td>
<td>25</td>
<td>.991</td>
<td>p &gt; .05</td>
</tr>
<tr>
<td></td>
<td>Post-test Scores</td>
<td>26</td>
<td>63.8</td>
<td>17.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the paired samples t-test in Table 1 compared the English communicative competence pre-test and post-test scores of pupils in the communicative approach, the clicker and the control groups. The results indicated that there was a statistically significant difference between the communicative competence pre-test scores and post-test scores for pupils in the communicative approach and the clicker’s groups. The results also revealed no statistical significant difference between the communicative competence pre-test scores and post-test scores for pupils in the control group.

The results suggest that pupils in the communicative approach and clicker groups recorded higher communicative competence scores at the post-test than in the pre-test. Pupils in the control group had very similar communicative competence pre-test and post-test scores. The results suggest that unlike in the traditional classroom, pupils’ communicative competence in the ESL classroom would improve when learners are exposed to the communicative approach, and the clickers. The hypothesis “the communicative competence pre-test and post-test scores would not be significantly different for pupils in each of the groups” was thus rejected.

The results of the Analysis of Covariance (ANCOVA) in Table 2 show that the overall model is statistically significant \(F_{(2, 92)} = 54.93, p < .05\) and that after adjusting for the pre-test language skills development scores, there was a significant effect of treatment on pupils’ post-test language skills development scores \(F_{(2, 95)} = 38.28, p < .05\). The results further show that 62.3% of the total variance in pupils’ post-test communicative competence scores was accounted for by the three levels of teaching strategies after controlling for the effect of pupils’ pre-test communicative competence development scores. The outcome thus indicates that the strength of the relationship between the teaching strategies and pupils’ posttest communicative competence scores in the ESL classroom was very strong.
Table 2. Effect of treatment on pupils’ communicative competence

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>26769.72</td>
<td>3</td>
<td>8923.24</td>
<td>54.93</td>
<td>.000</td>
</tr>
<tr>
<td>Intercept</td>
<td>20213.07</td>
<td>1</td>
<td>20213.07</td>
<td>124.43</td>
<td>.000</td>
</tr>
<tr>
<td>Pre Comm. Competence</td>
<td>15462.22</td>
<td>1</td>
<td>15462.22</td>
<td>95.18</td>
<td>.000</td>
</tr>
<tr>
<td>Treatment</td>
<td>12435.90</td>
<td>2</td>
<td>6217.95</td>
<td>38.28</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>15432.28</td>
<td>95</td>
<td>162.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>609021.00</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>42202.00</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 presents the outcome of the univariate analysis of covariance for the effect of the clicker technology, the communicative approach and the lecture method on pupils’ communicative competence post-test scores in the ESL classroom. The results showed that there was a significant effect of the teaching strategies on the pupils’ communicative competence development in the ESL classroom ($F(2, 95) = 38.28, p < .05$). The results suggest that pupils’ levels of communicative competence differed across the three teaching strategies. In other words, pupils’ levels of communicative competence would be determined by the type of teaching strategy they are exposed to in the ESL classroom.

Table 3. ANOVA for the effect of treatment on pupils’ communicative competence

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast</td>
<td>12435.896</td>
<td>2</td>
<td>6217.948</td>
<td>38.28</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>15432.282</td>
<td>95</td>
<td>162.445</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In view of the statistically significant difference in the English language communicative competence post-test scores across the groups, a follow-up test was conducted to evaluate pairwise differences among the adjusted means for communicative competence post-test scores. The Bonferroni procedure was used to control for Type I error across the three pairwise comparisons (see Table 4).

Table 4. Comparisons of differences in communicative competence post-test scores by group

<table>
<thead>
<tr>
<th>(I) Treatment</th>
<th>(J) Treatment</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comm. App.</td>
<td>PRS</td>
<td>-12.3*</td>
<td>3.08</td>
<td>.000</td>
<td>-19.8 -4.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control Group</td>
<td>15.8*</td>
<td>3.52</td>
<td>.000</td>
<td>7.2 24.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRS Group</td>
<td>Comm. Approach</td>
<td>12.3*</td>
<td>3.08</td>
<td>.000</td>
<td>4.8 19.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control Group</td>
<td>28.0*</td>
<td>3.22</td>
<td>.000</td>
<td>20.2 35.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Group</td>
<td>Comm. Approach</td>
<td>-15.8*</td>
<td>3.52</td>
<td>.000</td>
<td>-24.4 -7.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PRS</td>
<td>-28.0*</td>
<td>3.22</td>
<td>.000</td>
<td>-35.9 -20.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p < .05$.  

263
Table 4 shows the results of the evaluated pairwise differences among the adjusted means of communicative competence post-test scores in all the groups. The Bonferroni pairwise comparisons results indicated that the mean communicative competence post-test scores difference between the communicative approach group and the control group was statistically significant. Moreover, the results indicated that the mean communicative competence post-test scores difference between the PRS group and the communicative approach group was also statistically significant. Similarly, the results further indicated a statistically significant difference between the mean communicative competence post-test scores of the PRS and the control group.

![Boxplot of communicative competence post-test scores by group](image)

*Figure 2. Boxplot of communicative competence post-test scores by group*

In Figure 2, the boxplot shows that the clickers’ group had the highest median communicative competence post-test score while the control group recorded the lowest median communicative competence post-test score. The interquartile range of the pupils’ communicative competence post-test scores differed from one group to another. The overall range of the data showed that the control group had the highest range while the clickers’ group had the smallest range. The range for the communicative approach group was higher than that of the clickers’ group.

Meanwhile, the spread of communicative competence post-test score scores in the clickers’ group was slightly higher than that of the communicative approach group. The communicative competence post-test score distribution in the control group was positively skewed, indicating that many learners in the group had lower communicative competence post-test scores as compared to those with high communicative competence post-test scores. The clickers’ and the communicative approach groups appear negatively skewed indicating that a majority of the pupils had improved communicative competence scores. Meanwhile, the communicative competence post-test scores of the communicative approach group were slightly more negatively skewed than that of the clickers’ group. The results thus revealed that more pupils experienced improved communicative competence at the post-test in the clickers’ group when compared with those in the communicative approach group. There were no outliers in the distributions across the groups.

Overall, the results suggest that, pupils’ English language communicative competence would improve if they are exposed to communicative approach and clickers. However, there may be no improvement over the time scale of these observations in the pupils’ communicative competence development if ESL is taught with the lecture method. Moreover, the results also show that pupils taught with clickers would experience more of communicative competence than those exposed to the communicative approach and the lecture method. The hypothesis “the communicative competence post-test scores of pupils across the groups would not be significantly different” was therefore rejected.
The results of the multiple regression analysis in Table 5 show that a combination of gender, classes of pre-test, listening and speaking skills contributed a coefficient of multiple regression of .925 and a multiple correlation square of .849 towards the prediction of pupils’ level of communicative competence in English language. The results thus suggest that 84.9% of the total variance of the communicative competence attained by pupils in the ESL classroom is accounted for by the combination of the independent variables. The results further reveal that, the analysis of variance of the multiple regression is significant ($F = 138.46$, $p = .000$). Moreover, while pupil’s speaking ability provides the best contribution while pupils’ class of pre-test scores is the least significant contributor to the prediction of pupils’ communicative competence development. Gender did not make a significant contribution to the prediction of pupils’ communicative competence development.

Table 5. Multiple regression analysis of the predictor variables and pupils’ communicative competence

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>16.38</td>
<td>4.17</td>
<td>.390</td>
<td>.000</td>
</tr>
<tr>
<td>Gender</td>
<td>.65</td>
<td>1.70</td>
<td>.02</td>
<td>.38</td>
</tr>
<tr>
<td>Classes of pretest</td>
<td>-18.84</td>
<td>2.53</td>
<td>-.46</td>
<td>-7.40</td>
</tr>
<tr>
<td>Listening skills</td>
<td>.63</td>
<td>.06</td>
<td>.47</td>
<td>11.90</td>
</tr>
<tr>
<td>Speaking skills</td>
<td>.64</td>
<td>.04</td>
<td>.97</td>
<td>16.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R (adjusted) = .925</td>
</tr>
<tr>
<td>Multiple R2 (adjusted) = .849</td>
</tr>
<tr>
<td>Standard error estimate = 8.07</td>
</tr>
<tr>
<td>$F = 138.46$</td>
</tr>
<tr>
<td>Sig. = .000</td>
</tr>
</tbody>
</table>

Discussion

The purpose of this study was to investigate whether pupils’ communicative competence in the traditional ESL classroom would be better improved with the teacher’s adoption of the communicative approach only or the use of the clickers in a communicative approach context. This study reveals that, unlike in the control group, the English language communicative competence of the clickers’ group increased from the pre-test to the post-test. This outcome is confirmed by the findings of FitzPatrick, Finn and Campisi (2011) which indicate that the use of clickers is associated with increased learning performance. The control group’s low communicative competence may be associated with discipline anxiety, often prompted by teacher’s negative criticisms, or corporal punishment meted out for pupils who gave wrong answers. In such a learning environment, pupils are likely to be unwilling to talk in class. The interactive element of clickers enables learners to showcase their levels of understanding of the lesson and to develop new knowledge while they test out their knowledge by sharing information with others. As the pupils in the clickers’ group discussed and shared opinions collaboratively, after casting their votes to respond to teacher’s questions in the first instance, they learned a lot from one another. Moreover, the process of negotiation of meaning could have increased the quantity of language practice opportunities which the pupils needed to improve their oral communication fluency and academic performance. Moreover, the clickers’ group peer-interaction was with less intrusion from the teacher. The teacher mostly acted the role of what Giri (1996) and Littlewood (1981) referred to as a facilitator who offers suggestions regarding solutions to the assigned tasks. The results of this study confirm the relevance of interaction in the L2 learning process; where it is essential for the learners to practice the use of the target language in an authentic context. That the clickers motivated the pupils to oral communication skills’ development through practice more than they would have in the traditional classroom implies that the technology created new pathways to language learning.

Findings of this study also reveal that the communicative approach group improved more at the post-test when compared with their pre-test performance. Earlier research outcomes also show that the communicative competence
level of students who were taught with the lecture method was low when compared with those who were exposed to electronic board (Zha, Kelly, Ko Park & Fitzgerald, 2006) and task-based learning (Livingstone, 2010, Liqun & Xiubo, 2011). Pupils in the communicative approach group were exposed to a series of interactive tasks, such as role-play, drama, dialogues, games, and game-like activities during their English lessons. Besides, tasks were sometimes supported with pictorial illustrations, which could have enhanced pupils’ understanding and comprehension of concepts. The tasks could have provided the pupils with the opportunity to make more input and increased time of oral production of the target language. Such opportunities were lacking in the traditional classroom, where the teacher did most of the talking and the pupils sat and passively listened. Where pupils had the opportunity to talk, their utterances were well tailored by the teacher for grammatical correctness. Informally, the researcher observed that pupils in the traditional classroom mostly acted the teacher’s scripts by doing whatever she wanted in the way the teacher desired. Long and Porter (1985) remarked that learners’ perceptions of the teacher as a judge constitute a limitation to learners’ speech confidence and speech practice in the target language.

This study reveals an outstanding improvement in the communicative competence of pupils in the clickers’ classroom as compared to their counterparts in the communicative approach and the control groups. The outcomes of this study match up with the findings of earlier researchers (Basoglu & Akdemir, 2010; Gok, 2011), which reveal that, learning achievement is better improved when clicker is integrated within discussion session than the adoption of non-technology interactive pedagogies or the use of flashcards. However, Morgan (2008) reported no significant difference in the academic performances of learners exposed to clickers and those taught with the lecture method. Verkler (2004) opines that a language is best learned when a child engages in rich and authentic communication with peers, when appropriate technology is employed to enhance the interactive session. Exposing the pupils to graphic illustrations (bar charts) and activity-oriented learning experience, combined with the creative integration of the clickers in class, may have made language learning more appealing to the pupils. The facilitated zest might have contributed to the improved communicative competence experienced by pupils taught with clickers in their ESL classroom.

The outcome of this study underscores the importance of learners’ active engagement and ownership of learning in L2 classroom through dialogic communication. Moreover, the fun from using clickers could have triggered pupils’ interest in ESL learning more that they would have in a non-technology communicative approach context. The ANCOVA analysis rules out the assumption that the difference observed across the groups was due to teacher’s effect or the conduct of the research. Pupils in the communicative approach group experienced a non-threatening classroom atmosphere, and teacher’s non-interference in group discussions, but the anonymity of clickers group’s responses might have minimized the degree of learners’ exposure to embarrassment. Such non-embarrassing learning environment may have encouraged the less confident pupils in the clickers’ group to use the target language than those in the communicative approach only context, let alone those in the traditional classroom. The researcher thus contend that learners’ language oral communication skill is best developed when clickers is combined with peer instruction than when pupils learn language communicatively without the technology.

On the findings regarding the contributions of the independent variables to the prediction of pupils’ communicative competence development in the ESL classroom, all the four variables were significant joint contributors to the criterion variable. Relatively, pupils’ classes of pre-test scores, listening and speaking skills contributed significantly to the prediction of the criterion variable. The outcome of this study tallies with earlier findings (Bozorgian, 2012; Bahrani & Sim, 2012). Perhaps the low pre-test scorers were more attentive and ready to have a good grasp of the subject content in order to ensure that they contribute meaningfully during the group discussion. Moreover, the non-significant contribution of gender to the prediction of pupils’ English language communicative competence is in harmony with the findings of Huang (2010), and MacIntyre, Baker, Clement, and Donovan (2002). The outcome of this study may not be unconnected with the fact that both boys and girls in this study were introduced to English language at the same stage of education and grew in a social setting where communication outside the school was in the native language.

**Conclusion and implications**

The interactive features and reduced cost have added to the popularity of clickers in the education sector. Non-native speakers require interactive learning environment in order to improve their proficiency in English. As the use of clickers in education is gaining ground in various disciplines, the outcome of this study underscores the efficacy of
clickers in ESL classrooms. Moreover, the use of clickers with peer instruction serves as an important strategy of using technology to enrich learners’ learning experience and improve their communicative competence in the ESL classroom. Moreover, the study has contributed to the fledging literature on whether clickers can make a significant difference in improving L2 learners’ oral proficiency. It is high time teachers explore the potential of the technology to make ESL learning more attractive to learners, provide learners opportunity to use the target language in oral communication, and improve their communicative skills. It is recommended that further studies should investigate the effects of clickers on the four language skills in ESL classroom. Moreover, further research with post-primary school students should be conducted within the context of Nigeria and across countries in order to make the findings more generalisable.

References


An Automatic Caption Filtering and Partial Hiding Approach to Improving the English Listening Comprehension of EFL Students

Ching-Kun Hsu, Gwo-Jen Hwang and Chih-Kai Chang

Department of Technology Application and Human Resource Development, National Taiwan Normal University, No. 162, Sec. 1, Heping E. Rd., Taipei City 10610, Taiwan // Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Taipei, 106, Taiwan // Department of Information and Learning Technology, National University of Tainan, No. 33, Sec. 2, Shu-Lin St. Tainan, 700, Taiwan // ckhusu@ntnu.edu.tw // gihwang.academic@gmail.com // chihkai@mail.nutn.edu.tw

Corresponding author

(Submitted January 10, 2013; Revised June 20, 2013; Accepted June 28, 2013)

ABSTRACT

Fostering the listening comprehension of English as Foreign Language (EFL) learners has been recognized as an important and challenging issue. Videos have been used as one of the English listening learning resources; however, without effective learning supports, EFL students are likely to encounter difficulties in comprehending the video content, leading to frustration and reducing learning interest. In this study, a learning system with an automatic caption filtering and partial hiding mechanism was developed to improve the English listening comprehension of EFL students. An experiment was conducted to evaluate the effects of the proposed approach on students’ learning achievements and perceptions. The participants were 76 freshmen from two classes of a non-language-related department of a university. Each class contained 38 students. The two classes of students were situated in the proposed learning context and the conventional technology-enhanced learning context with a counterbalanced experimental design. The experimental results verified that the proposed caption filtering approach effectively improved the listening comprehension of the students in comparison with the conventional approach which provides full captions and an e-dictionary. Moreover, from the data collected by the eye movement tracking device, it was found that more than 76% of the students relied on the captions during the learning activity, reflecting the importance of developing effective caption filtering mechanisms for supporting EFL learners.

Keywords

Computer-assisted language learning, Caption filtering, English as foreign language, English listening comprehension, Eye movement tracking device

Introduction

Learning English as a foreign language has been heavily emphasized in Asia owing to its predominance in international communications. Among listening, speaking, reading and writing, listening has been recognized as the most essential aspect of language development (Chung, 1999; Liu, Chen, & Chang, 2009). Due to increasing needs, it has become an important and challenging issue to develop innovative and effective approaches to improving EFL students’ English competence. In recent years, the advancement and popularity of computer, communication and multimedia technologies has provided an effective way to facilitate English listening comprehension training (Vanderplank, 2010). Various web-based or computer-assisted learning systems have been developed for conducting English listening activities (Chapelle, 2009; Liu & Chen, 2007; Liu, Liu, & Hwang, 2011).

Among various English learning sources, videos are the most popular form for training listening comprehension. Researchers have indicated that simultaneous audial and visual input can benefit foreign language learners (Seo, 2002); therefore, many studies related to EFL learning mainly employ videos as learning materials rather than audios or texts (Chapple & Curtis, 2000; Vanderplank, 2010; Williams & Thorne, 2000). Chapple and Curtis’s (2000) have further reported that videos not only help learners improve their listening comprehensions, but also promote their confidence in speaking English, in particular, for those who are not native English speakers, such as EFL students.

Furthermore, to assist EFL students in comprehending the learning content, most English learning systems provide subtitles or captions on videos. Subtitles are the on-screen text in the students’ native language combined with a second language soundtrack in the video, while captions are the on-screen text in the same language as the soundtrack (Markham, Peter, & McCarthy, 2001; Pujolá, 2002). In this study, the term subtitles refers to on-screen
Chinese text combined with an English soundtrack, while captions refers to on-screen English text combined with an English soundtrack. There are many advantages of providing captions when using videos as foreign language listening training materials (Yang, Huang, Tsai, Chung, & Wu, 2009). For example, Garza (1991) indicated that the use of captions could bridge the gap between the students’ competence in reading and listening; Chung’s (1996) study reported that videos with captions helped students associate the spoken and written forms of words more easily and quickly than videos without captions. In the meantime, many studies have shown that subtitles can enhance the reading, vocabulary, and listening competences of the students who learn with videos (Danan, 1992; Markham et al., 2001; Hayati & Mohmedi, 2011; Danan, 2010).

On the other hand, researchers have reported that relying heavily on subtitles when watching audio-visual materials is not conducive to improving listening proficiency (Latifi, Mobalegh, & Mohammadi, 2011). Some researchers have further indicated that videos without subtitles or captions are more beneficial as they induce students to pay attention to various pronunciation features, such as reduced forms, assimilation, elision, and re-syllabification (Hulstijn, 2003; Field, 2003; Vandergrift, 2007). In sum, although captions and subtitles could be helpful to learners in comprehending learning materials, it is important to provide learning supports that meet individuals’ knowledge level at the right time to avoid possible negative effects. Miller (1956) has noted that people’s capacity for processing information is limited, implying the need of providing caption-filtering mechanisms in developing learning system for English listening comprehension by taking into account what the learners need based on their knowledge levels. On the other hand, Mayer and Moreno (2003) have indicated the importance of presenting relevant learning materials (e.g., texts, videos, images) at closer space and time to improve learners’ comprehensions; therefore, it is interesting to investigate the helpfulness of filtered captions by analyzing students’ fixation status during the video watching process using eye tracking devices.

In this study, an automatic caption filtering and partial hiding mechanism is proposed to cope with this problem. A learning system is developed based on the proposed mechanism to assist EFL students in improving their English listening comprehension. Moreover, an experiment has been conducted to evaluate the performance of the proposed approach with several measuring tools, including achievement tests, questionnaire surveys and an eye-tracking machine.

**Literature review**

**Captions and listening comprehension**

Researchers have indicated that simultaneous audial and visual input can benefit foreign language learners (Seo, 2002); therefore, many studies related to language learning mainly employ videos as the learning materials (Chapple & Curtis, 2000; Vanderplank, 2010; Williams & Thorne, 2000). On the other hand, educators have noted that foreign language learners usually require vocabulary support when learning with video materials (Chung & Huang, 1998), implying the need to provide captions to help students understand the materials. Guillory (1998) showed that video-embedded keywords or full-text captions could benefit students more in terms of comprehending the learning content than non-captioned videos.

Several studies have also reported consistent findings related to the effectiveness of using captions in language learning (Chai & Erlam, 2008). For example, Winke, Gass, and Sydorenko (2010) indicated that captions are able to increase learners' attention and improve their learning achievements by linking with their previous knowledge.

The effect of captions or subtitles on video-based learning is highly related to the Information Processing Theory (IPT) (Miller, 1956), which refers to how humans process information via their short-term memory and store the processed results in their long-term memory. Short-term memory is limited, implying the importance of providing effective prompts to help students pay attention to important and critical information (Card, Moran, & Newell, 1983). Wang and Liu (2011) indicated that, with proper captions, students could improve their English learning interest and efficiency, showing the potential of filtered captions in language learning.
Eye-tracking and the noticing hypothesis

Attention and noticing have been recognized as important factors that affect students' learning performance in second language acquisition (Izumi, 2002; Li & Iribe, 2010; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Schmidt, 2001). The advancement of eye tracking technology provides an effective way to investigate the noticing behaviors of students by recording and analyzing their eye focusing tracks on learning materials (Scheiter & van Gog, 2009; van Gog & Scheiter, 2010). Researchers have indicated that eye focusing tracks are related to cognitive processing (Rayner, 1998; West, Carlson & Cohen, 2007); moreover, analyzing eye fixation data is an effective means of determining attentive selections (e.g., Deubel & Schneider, 1996; Henderson & Hollingworth, 1998; Hoffman & Subramaniam, 1995; Kowler et al., 1995; Irwin & Gordon, 1998).

In the past decade, eye tracking equipment has been used in various studies for analyzing the learning attention and cognitive processes of students (Jacob & Karn, 2003; Li & Iribe, 2010; Ozcelik, Arslan-Ari, & Cagiltay, 2010). For example, Ozcelik, Arslan-Ari and Cagiltay (2010) analyzed eye movement tracking data and found that when students watched and listened to learning materials, proper prompting messages that guided them to pay attention to useful or important learning information helped improve their learning effectiveness and efficiency. The study of Loboda and Brusilovsky (2010) further showed the effectiveness of adaptive visualization, that is, the provision of a clear visual metaphor for concepts or comprehension by selectively showing supportive information for important content while hiding other information. The findings from these previous studies have not only confirmed the eye-mind assumption of reading proposed by Just and Carpenter (1980), but have also shown the usefulness of using eye movement equipment in investigating the learning behaviors and cognitive processes of students (Eberhard et al., 1995; Griffin, 2001; Griffin & Bock, 2000; Tanenhaus et al., 1995).

Development of a learning system with automatic caption filtering

In this study, a learning system with an automatic caption filtering mechanism is proposed, as shown in Figure 1. The system was developed with Visual Basic Studio 2010 programming language and Microsoft Access 2010 database. The video playing functions provided by the learning system included "play" and "pause." When the learners press the "pause" button, the vocabulary included in the previous classes they had attended was hidden, while that which was new to them was displayed based on their learning profiles. This software is installed in a laptop for broadcasting movies. The captions of the movies are filtered before showing in the caption area of this application software when the students pressed the button of “Pause” function, and hiding the easier words from the original captions.

Figure 1. Architectural diagram of the automatic caption-filtering system

Figure 2 shows the learning scenario of using the developed learning system (left) and a sample figure of eye focus distributions (right). The bluish color denotes the area to which the user pays less attention, while the reddish color represents where the user gazes more.
Research questions

To evaluate the effectiveness of the automatic caption filtering and partial hiding mechanism proposed in this study, an experiment was conducted in the National University of Tainan, Taiwan. The students' learning performance and learning behaviors were analyzed using a Computer Usability Satisfaction Questionnaire (CUSQ), two tests, and an eye movement tracking device to investigate the following research questions.

1. What do the students pay attention to during the learning process?
2. Does the proposed approach enhance the listening comprehension of EFL students in comparison with the conventional computer-assisted language learning approach?
3. What do the students feel about the proposed approach and the conventional computer-assisted language learning approach in terms of the usability, ease of use, and satisfaction with the captions provided by the two approaches?

Experimental design

An experiment was conducted to evaluate the effectiveness of the caption filtering mechanism. In this experiment, the video soundtrack and the corresponding captions were in English; that is, the participants were prevented from watching subtitles in their first language during the video play; in the meantime English captions were displayed at the bottom of the screen.

Participants

In this study, the participants were two classes of freshmen from a non-language-related department of a university in Taiwan. Each class consisted of 38 students. The participants’ average age was eighteen. All of them had learned the 2,000 most frequently used words as defined by the Ministry of Education of Taiwan, and had passed the primary level of a national English proficiency test.

Learning materials

Two videos of the same language level and length were selected by two experts who had more than five years’ experience of teaching English. To avoid interference between the two videos due to familiarity after watching one of them, the two videos represented different themes aimed at introducing computer science. One introduced PC hardware components, while the other was related to web searching functions. Both videos used colloquial style English to convey the basic concepts as real life characteristics. The two videos were used in the learning activities based on the original curriculum design of the sample course. The aim of the course was to introduce the structure, functionalities and applications of computers and networks. That is, the learning activity as well as the learning content reflected the teaching reality of that course. For example, the participants of this study had acquired the 2,000
most frequently used words defined by the Ministry of Education in Taiwan; therefore, in Figure 3 which is the snapshots of the learning system, those words were hidden, so that they could pay more attention to the target vocabulary presented in the videos. When the students pressed the button of “Pause,” all the translations of the other harder words are shown in the instant translation area. In the following, we shall call the 2,000 most frequently used words “familiar words,” while referring to the others as "unfamiliar words."

![Image](image_url)

**Figure 3.** Hiding the easier words and instantly translating the harder words after pausing

**Measuring tools**

For each video, a test was developed by three experienced English teachers, and was pronounced by two native English speakers who had more than five years of English teaching experience. Each of the tests included two parts: that is, an auditory test and a listening comprehension test. In the auditory test, the questions and choices for testing the video content were presented via audios; in the listening comprehension test, the questions and choices for testing the video content were presented in texts. Both the test parts contained four multiple-choice items with a perfect score of 100. To examine the reliability of the tests, the study conducted a preliminary examination on 14 students in the same department of that university. In terms of the two auditory tests, the Cronbach's alpha value of reliability statistics is .928, and the correlation coefficient value of Spearman's rho is .866 from the equivalent-forms method of pretesting. As for the two listening comprehension tests, the Cronbach's alpha value of reliability statistics is .980, and the correlation coefficient value of Spearman's rho is .962. From the statistical results, it is inferred that the two tests can be viewed as equivalent.

The CUSQ was developed by Lewis (1995). It consists of three dimensions (i.e., “ease of use,” “usability” and “satisfaction”) with a total of 19 questions using a seven-point Likert scale. The Cronbach's alpha values of the three dimensions were 0.75, 0.71 and 0.79, respectively.
Experimental procedure

The treatment of this study is an experimental counterbalanced design, which is one of the formal educational research methods (Gall, Gall & Borg, 2007), and is also used in language learning studies (e.g., Freiermuth, 2001; Snellings, van Gelderen & de Glopper, 2004). The choice of this approach is to reduce the influence of ordering when comparing the effectiveness of two treatments, while allowing the students in both groups to experience the two approaches. This design meets the request from the teacher that all of participants have the same opportunity to receive the new learning approach. As shown in Table 1, the participants in both classes experience both the experimental group treatment and the comparison group treatment. That is, 38 students in class A learned with the experimental group treatment first and then experienced the comparison group treatment thereafter. On the other hand, the other 38 students in class B learned with the comparison group treatment first and then experienced the experimental group treatment later.

Table 1. Counterbalanced experimental design

<table>
<thead>
<tr>
<th>Order</th>
<th>N</th>
<th>First</th>
<th>Latter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class A</td>
<td>38</td>
<td>X₁O₁</td>
<td>X₂O₂</td>
</tr>
<tr>
<td>Class B</td>
<td>38</td>
<td>X₂O₂</td>
<td>X₁O₁</td>
</tr>
</tbody>
</table>

Note. X₁: Comparison group treatment; X₂: Experimental group treatment; O₁: Post-test of the auditory test and listening comprehension test; O₂: Post-test of the auditory test and listening comprehension test.

During the learning process, the students in the comparison group learned with the conventional technology-enhanced learning approach; that is, they were instructed to push the pause button of the broadcast software whenever they could not comprehend what they had heard, allowing them to consult an e-dictionary embedded in the system for finding the Chinese translation of the unfamiliar words. When the students paused the videos and pointed at an unfamiliar word in the caption, possible Chinese translations of the word were displayed.

On the other hand, the students in the experimental group learned with the proposed system, where the familiar vocabulary was automatically filtered, while the unfamiliar vocabulary (e.g., the new vocabulary to be learned in the present unit) and the corresponding Chinese translations were displayed when the video player was paused.

While watching the videos, all of the students were observed by an eye movement tracking device. The eye-tracking system used was MangoldVision, developed by Mangold International GmbH. The portable MangoldVision Eyetracker comes in a case the size of a small briefcase (25x40x10 cm).
Figure 4 shows the duration of each treatment stage and the distribution of the participants. The participants took an auditory test (15 minutes), and then took a listening comprehension test (15 minutes) after watching the video with each treatment. The scores were collected for the paired-samples t-test because all of the students participated in the counterbalanced design of the experiment to prevent the impact of sequence between the two treatments. After the counterbalanced treatments, the students filled in the usability questionnaire to give feedback regarding their opinions. Finally, the results were analyzed for future reference.

Results and discussion

Students’ eye fixation analysis

From the analysis of the eye focusing tracks, it was found that most of the students in both groups focused on the captions during the video play. The eye-focusing positions were categorized into four types based on the data collected from the eye tracking device, as shown in Table 2. It was found that more than 76% of the participants (N = 59) paid much attention to the captions when watching the videos; that is, they tended to rely on the captions to comprehend the learning content, indicating the importance of developing more effective caption mechanisms for EFL students.

<table>
<thead>
<tr>
<th>Stare position</th>
<th>Example figures</th>
<th>Class A</th>
<th>Class B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>X1</td>
<td>X2</td>
</tr>
<tr>
<td>Mainly on the captions</td>
<td></td>
<td>39.47%</td>
<td>38.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(N = 30)</td>
<td>(N = 29)</td>
</tr>
<tr>
<td>Mainly on the video</td>
<td></td>
<td>5.26%</td>
<td>3.95%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(N = 4)</td>
<td>(N = 3)</td>
</tr>
<tr>
<td>Both captions and video</td>
<td></td>
<td>5.26%</td>
<td>7.89%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(N = 4)</td>
<td>(N = 6)</td>
</tr>
<tr>
<td>Distracted from captions and video</td>
<td></td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(N = 0)</td>
<td>(N = 0)</td>
</tr>
</tbody>
</table>

Note. % is of total number (i.e., 76); X1: Comparison group treatment; X2: Experimental group treatment.

Results of the auditory tests and listening comprehension tests

Table 3 shows the statistical results of the scores of the two groups. It was found that, for the students in both groups, the auditory test scores after experiencing the experimental group treatment were always higher than those after experiencing the comparison group treatment. The difference between the means of the two test scores was more than 7 points. Moreover, for both groups, the listening comprehension test scores after receiving the experimental
group treatment were higher than those after receiving the comparison group treatment. The difference between the means of the two test scores was more than 5 points.

The t-test result further showed that the two treatments differed significantly in the listening outcomes ($t = 2.66^*; p = .01$), but not in the comprehension degrees ($t = 1.51; p = .13$). From the t-test result, it is concluded that the first research hypothesis is validated; that is, the learning system with the caption filtering approach is more effective than the conventional approach in terms of facilitating English listening comprehension ability.

Table 3. Paired-samples t-test of the two treatments in the auditory tests and listening comprehension tests

<table>
<thead>
<tr>
<th>Post-test</th>
<th>Treatments</th>
<th>Samples</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Tests</td>
<td>X1</td>
<td>Class A</td>
<td>38</td>
<td>49.34</td>
<td>29.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Class B</td>
<td>38</td>
<td>45.39</td>
<td>24.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>76</td>
<td>47.37</td>
<td>26.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Class B</td>
<td>38</td>
<td>57.89</td>
<td>27.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Class A</td>
<td>38</td>
<td>55.26</td>
<td>30.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>76</td>
<td>56.58</td>
<td>28.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paired sample t</td>
<td></td>
<td></td>
<td>30.23</td>
<td>2.66*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>test (X2_ALL-X1_ALL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening comprehension</td>
<td>X1</td>
<td>Class A</td>
<td>38</td>
<td>57.24</td>
<td>25.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Class B</td>
<td>38</td>
<td>58.55</td>
<td>24.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>76</td>
<td>57.89</td>
<td>24.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Class B</td>
<td>38</td>
<td>62.50</td>
<td>23.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Class A</td>
<td>38</td>
<td>63.82</td>
<td>25.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>76</td>
<td>63.16</td>
<td>24.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paired sample t</td>
<td></td>
<td></td>
<td>30.36</td>
<td>1.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>test (Sum of X2 - Sum of X1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05. X1: Comparison group treatment; X2: Experimental group treatment.

To understand more deeply the relationship between the learning outcomes and eye-gazing, the study selected those students whose performance was in the top 27% of the listening and listening comprehension test outcomes as the high-achievement group, and those in the bottom 27% as the low-achievement group (Ebel, 1972). The fixation distributions of the students in different achievement groups were compared.

Table 4 shows the statistical data of four categories of fixation distributions (i.e., paying attention to captions, videos, both, or none) for the low- and high-achievement students. From the Pearson Chi-Square analysis, it was found that different achievement students had similar fixation distributions. For example, in terms of the comparison group treatment students’ behaviors during the auditory test, 75% of the high-achievement students and 87.5% of the low-achievement students mainly focused their eyes on the captions during the learning activity. From the results of the Chi square tests, the dependent degrees on the captions are similar ($X_2 = 0.65; p > 0.05$), indicating that most students, no matter whether they were high-achievement or low-achievement, needed the assistance of the captions to comprehend the learning materials.

Table 4. The fixation distribution of the high- and low-achievement students

<table>
<thead>
<tr>
<th>Tests</th>
<th>High-achievement students</th>
<th>Low-achievement students</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Caption</td>
<td>Video</td>
<td>Both</td>
</tr>
<tr>
<td>Auditory tests</td>
<td>C 75.00%</td>
<td>15.00%</td>
<td>10.00%</td>
</tr>
<tr>
<td></td>
<td>E 68.18%</td>
<td>22.73%</td>
<td>9.09%</td>
</tr>
<tr>
<td>Listening comprehension</td>
<td>C 83.33%</td>
<td>11.11%</td>
<td>5.56%</td>
</tr>
<tr>
<td>tests</td>
<td>E 65.22%</td>
<td>21.74%</td>
<td>13.04%</td>
</tr>
<tr>
<td>Note</td>
<td>C: Comparison group treatment; E: Experimental group treatment.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In comparison with the comparison group treatment, it was found that, on average, the students who learned with the experimental group treatment, no matter whether they were high- or low-achievement, spent less time on reading the captions. For example, in the listening comprehension tests, only 65.22% of the high-achievement students who
learned with the experimental group treatment spent most of the time reading the captions, while 83.33% of those who learned with the comparison group treatment kept focusing on the captions. Moreover, only 70.83% of the low-achievement students who learned with the experimental group treatment spent most of the time reading the captions, while 88.89% of those who learned with the comparison group treatment comparison group treatment kept focusing on the captions. This finding implies that filtered captions were able to ease the load of the students during the learning process.

This study further analyzed the learning performance of the students who looked at the captions, to see if their learning behaviors were helpful in improving their listening or comprehending performance. It was found that most of the students receiving the experimental group treatment revealed better performance than those who learned with the comparison treatment. That is, the students had significantly better performance when reading the selected words in the captions than when reading the full captions in terms of their listening comprehension. Consequently, it can be inferred that the automatic caption-filtering system was helpful to the students.

Results of the usability evaluation

Table 5 shows the statistical results of the students’ ratings on the questionnaire items of the “ease of use,” “usability” and “satisfaction” dimensions.

| Table 5. Adaptive caption-filtering system usability questionnaire (Revised from CUSQ) |
|--------------------------------------|------|------|
| **Perceived Ease-Of-Use** | mean | SD  |
| 1. It was simple to use this system. | 5.24 | 1.33 |
| 2. I can effectively complete my work using this system. | 4.63 | 1.21 |
| 3. I am able to complete my work quickly using this system. | 4.87 | 1.02 |
| 4. It was easy to learn to use this system. | 5.78 | 1.07 |
| 5. Whenever I make a mistake using the system, I recover easily and quickly. | 4.28 | 0.99 |
| 6. It is easy to find the information I need. | 4.80 | 1.23 |
| 7. The information provided by the system is easy to understand. | 5.50 | 1.05 |
| **Average** | **5.01** | **1.23** |

| **Perceived Usability** | mean | SD  |
| 8. I am able to efficiently complete my work using this system. | 4.74 | 1.02 |
| 9. I believe I became productive quickly using this system. | 4.71 | 1.09 |
| 10. The system gives error messages that clearly tell me how to fix problems. | 4.17 | 0.97 |
| 11. The information (such as online help, on-screen messages, and other documentation) provided with this system is clear. | 4.59 | 1.07 |
| 12. The information is effective in helping me complete the tasks and scenarios. | 5.20 | 0.95 |
| 13. The organization of information on the system screens is clear. | 5.09 | 1.07 |
| 14. This system has all the functions and capabilities I expect it to have. | 4.84 | 0.95 |
| **Average** | **4.76** | **1.06** |

| **Perceived Satisfaction** | mean | SD  |
| 15. Overall, I am satisfied with how easy it is to use this system. | 4.92 | 1.10 |
| 16. I feel comfortable using this system. | 4.76 | 1.04 |
| 17. The interface of this system is pleasant. | 4.99 | 1.08 |
| 18. I like using the interface of this system. | 4.83 | 1.00 |
| 19. Overall, I am satisfied with this system. | 5.34 | 0.95 |
| **Average** | **4.97** | **1.05** |

In terms of the “ease of use” dimension, the average rating was quite positive (mean = 5.01; SD = 1.23). Among the seven items in this dimension, one item, “It was easy to learn to use this system,” reached 5.78 on average, indicating that the students perceived that the English listening competence training system with the adaptive caption-filtering mechanism was easy to use.

As for the dimension of usability, the average rating was 4.76. In this dimension, the average rating of the item “The information is effective in helping me complete the tasks and scenarios” is 5.2, implying that most students confirmed the usefulness of the proposed approach.
With regard to the dimension of perceived satisfaction, the average score was 4.97; moreover, the average rating of the item "Overall, I am satisfied with this system" was 5.34, revealing that most of the students were highly satisfied with the learning approach.

Figure 5 summarizes the results of the 76 questionnaires. According to this figure, 65% of the students indicated that the proposed system was usable and easy to use; more than 62% of the participants were satisfied with the learning approach. Moreover, no student selected the 7th category (i.e., strongly disagree) in answering the questionnaire items. To sum up, the caption filtering mechanism enabled the students to concentrate on the words that were difficult for them, which made the learning process more effective; furthermore, the provision of instant translations of the target words helped the students comprehend the new vocabulary in a more efficient way. As a consequence, the students gave quite positive feedback in terms of ease of use, usability and satisfaction.

Conclusions and implications

When learning a second or foreign language, students often encounter difficulties in watching or listening to those learning materials which involve a rapid rate of presentation and unknown vocabulary; moreover, redundant information (e.g., subtitles) can also confuse learners (Smidt & Hegelheimer, 2004; Kellerman, 1990). In this paper, an English listening comprehension training system with an automatic caption filtering approach is presented. To evaluate the performance of the proposed approach, the study compared the effects of the proposed learning system with the caption filtering mechanism and the conventional technology-enhanced learning approach with an e-dictionary on the auditory test and listening comprehension test. Moreover, an eye movement tracking device was used to collect the students’ eye-fixation data during the learning process.

For the first research question, the eye-tracking data analysis showed that 78% of the students in class A, and 76% of the students in class B focus their look at the captions, revealing the importance of providing effective captions in English listening training programs. This finding conforms to some previous studies reporting that participants’ eyes tended to fixate more on the text than on the illustrations; moreover, text seems to provide more helpful information than pictures (Liu & Chuang, 2010). In the meantime, the eye-tracking data also showed that, when learning with the videos with filtered captions, the students were able to focus more on the target words (new vocabulary) in comparison with the full captions, which could be the reason why the experimental group treatment benefited the students more in terms of their listening comprehension. This finding conforms to the study of Danan (2004), who indicated that captions visualize the information of the foreign language that students hear in videos, and showing only unfamiliar words can usually provide sufficient information to assist students with their listening comprehension. Several studies have also reported that hiding familiar words in captions enables students to pay
more attention to the target words, which is of great benefit to the development of their listening ability (Hulstijn, 2003; Field, 2003; Mitterer & McQueen, 2009; Stewart & Pertusa, 2004).

For the second question, it was found that the students who learned with the proposed caption-filtering approach showed significant better learning achievements in the auditory tests and listening comprehension tests than those who learned with the conventional technology-enhanced learning approach.

Moreover, for the third research question, the questionnaire results showed that most of the students highly accepted the proposed approach in terms of “ease of use,” “usability” and “satisfaction.” The average ratings of the three dimensions were 5.01, 4.76, and 4.97, respectively.

In the future, we plan to conduct several continuous studies. First, a large-scale experiment will be conducted using videos of various lengths and with more vocabulary. Second, we plan to employ the proposed approach in other English training programs in high schools. Third, in addition to applying the eye-tracking machine to detect and record the learning behaviors of students, we plan to record and analyze the time and frequency with which students press the “pause” button, such that a more effective caption mechanism might be proposed based on the analysis results. Fourth, we plan to expand the learning system by providing more facilities to help the students comprehend the learning content, as well as enhancing their foreign language listening proficiency, such as some Mindtools, context-aware facilities, or discussion forums developed in some previous studies (Chu, Hwang, Tsai, & Chen, 2009; Hwang, Chu, Lin, & Tsai, 2011; Hwang, Tsai, Chu, Kinshuk, & Chen, 2012).

Acknowledgements
This study is supported in part by the National Science Council of the Republic of China under contract numbers NSC 98-2511-S-024-008, NSC 99-2631-S-011-002 and NSC 101-2511-S-011-005 -MY3.

References


282


Assessment of Charisma as a Factor in Effective Teaching

Yun-Chen Huang¹ and Shu-Hui Lin²*

¹Department & Graduate school of Accounting Information, National Taichung University of Science and Technology, 129, Sec. 3, Sanmin Rd., Taichung, Taiwan, R.O.C. // ²Department of Insurance and Finance, National Taichung University of Science and Technology, 129, Sec. 3, Sanmin Rd., Taichung, Taiwan, R.O.C. // minnie@nutc.edu.tw // suelin@nutc.edu.tw

*Corresponding author

(Submitted February 6, 2013; Revised May 11, 2013; Accepted June 4, 2013)

ABSTRACT

In order to advance the effectiveness of higher education, we concentrated attention on teachers’ classroom behaviours and aimed to develop an inventory for evaluating teacher’s charisma in college settings. This study identified teacher characteristics that are indicators of teaching charisma and then developed a teaching evaluation questionnaire based on these indicators. The exploratory factor analysis and confirmatory factor analysis resulted in a 23-item Inventory of Teaching Charisma in College Classroom (ITCCC) that comprises four factors: knowledge, character traits, teaching techniques, and humor. As opposed to the questionnaire widely used in higher education institutions in Taiwan that mainly focuses on the evaluation of teacher’s teaching method or attitude, the new measure enables college students to evaluate teacher’s teaching in terms of their charisma. An advantage of this instrument is that it has just a few items and is therefore easy to administer and to complete. It results in good response rates from students and the analysis is straightforward, since the scores on the items are simply summed to get the subscale scores. The results can serve as a basis for teachers to improve these skills or qualities and may provide a framework to help teachers to increase their teaching efficiency.

Keywords

Character traits, Humor, Knowledge, Teaching techniques

Introduction

This paper describes the development of framework and a student evaluation instrument for use in college settings that can accurately measure a teacher’s charisma. While many studies have attempted to identify factors that contribute to teacher effectiveness and teaching quality, the majority are strictly theoretical. Empirical studies are also necessary to advance our understanding of this field. In particular a teacher’s charisma is often recognized as an important factor of his or her effectiveness in the classroom, but by its very nature this characteristic is regarded as something that is difficult to define or quantify. Thus a reliable instrument which can accurately measure a teacher’s charisma deserves to be developed.

When an instructor establishes a supportive social climate in the classroom, students are more likely to be receptive to learning. The behaviours and attitudes of teachers in teaching are the primary determinant on students’ perceptions of service quality in higher education (Hill, Lomas, & MacGregor, 2003; Pozo-Munoz, Rebolloso-Pacheco, & Fernandez-Ramirez, 2000). The teacher has a strong impact on their students, so it is no wonder that researchers suggest that institutions or academic administrators should pay more attention to what students want and should focus on understanding the needs of students (Joseph, Yakhou, & Stone, 2005; Oldfield & Baron, 2000). The feedback from students regarding their perceptions of teaching may help teachers to improve their teaching.

There has been increasing attention on teaching quality and effectiveness internationally, and considerable emphasis on promoting reflection and self-assessment in teachers, unclear what teachers are supposed to reflect on when desiring to become better teachers. Thus, a better framework for assessment may be needed to answer “what is effective teaching?” Whether it is even possible to judge teacher effectiveness outside of direct observations of their teaching remains an open question (Goldhaber & Anthony, 2007). The problem is compounded by many variables, for instance, the recognition of teaching effectiveness, the type of course, class size, student abilities, and grading practices, and so on (Young & Shaw, 1999). Indeed, Young and Shaw (1999) suggest the construct of teacher effectiveness is not only multidimensional but may also require multiple definitions. In the present study, we focused on one key factor – “teaching charisma” – and attempted to identify characteristics manifested in the classrooms of effective teachers that are indicators of charisma.
It is common to hear that some teacher is so welcome or popular that students like to attend his/her class. There may be some quality that deeply attracts students that not all teachers possess. We call this teaching charisma. The meaning of charisma comes from the Greek word translated as “gift,” suggesting that charismatic teachers have special gifts to distribute. It is commonly thought that it is the pleasing personality of the charismatic person that is his/her greatest gift (Raelin, 2006). We define teaching charisma as the positive behaviours of teacher, in the college classroom, which can deeply attract students to learn. A charismatic teacher not only is good in students’ perception but has appeal for students.

What are the essential qualities of a charismatic teacher? Certainly, he/she should have the characteristics which are considered to be good or desired to students. Several researchers have attempted to describe a good teacher or students’ desires regarding an ideal teacher. For example, Greimel-Fuhrmann and Geyer (2003) pointed out that a good teacher should offer explanations, answer questions, vary their teaching methods, and should be interested in and express concern for their students and their learning progress. Brown (2004) indicated that competent teachers know their subject, are willing to answer questions, are approachable, and also have a sense of humor. Gregory (2011) suggested that students favor teachers who are demanding, yet helpful and attentive, and a class that is strict, fair, and informative, and thereby perceive quality teaching and learning to comprise the same. Most of students want teachers to be friendly, knowledgeable, well-organized, encouraging, helpful, sympathetic, concerning for students’ individual needs, and strongly interested in the students’ learning (Anderson, 2000; Hill et al., 2003; Lammers & Murphy, 2002; Voss & Gruber, 2006).

As revealed from the above studies, in the eyes of students, there are several commonly held characteristics of good teaching. We consider that these indicators are similarly essential for a charismatic teacher. First, he/she is knowledgeable (Hill et al., 2003; Lammers & Murphy, 2002). It is an important attribute of teachers. Teaching is a complicated practice that requires an interweaving of many kinds of specialized knowledge (Koehler & Mishra, 2009). Scholars unanimously recognized that both subject matter knowledge and pedagogical knowledge are crucial to good teaching (Shulman, 1986). Shulman (1986) suggested that teaching expertise should be described and evaluated in terms of pedagogical content knowledge (PCK), which concerns the manner in which teachers relate their subject matter knowledge to their pedagogical knowledge. Central to Shulman’s conceptualization of PCK is the notion of the transformation of the subject matter for teaching (Koehler & Mishra, 2009).

The study (Pozo-Munoz et al., 2000) concluded that competence is the by far most important characteristic of ideal teachers. Teachers should have knowledge of their subject and be able to communicate it clearly to their students (Voss & Gruber, 2006). Bain (2004) mentioned, in the book—“What the Best College Teachers Do”, the outstanding teachers know their subjects extremely well. More important, they know how to simplify and clarity complex subjects, to cut to the heart of the matter with provocative insights, and they can think about their own thinking in the discipline, analyzing its nature and evaluating its quality.

Second, he/she has positive character traits such as friendliness, approachability, patience and enthusiasm (Anderson, 2000; Greimel-Fuhrmann & Geyer, 2003; Hill et al., 2003; Voss & Gruber, 2006). Teachers are expected to be good role models for the students, thus are expected to possess high moral and ethical standards. Chou (1997) investigated the teacher-student relationship and suggested that teachers should pay attention to the performance of the individual self. Students expect teachers not to display unbecoming emotions in the classroom. Teacher’s behaviours are always regarded as a model for students to imitate, thus students expect teachers to display appropriate behaviours, especially in the aspect of morality. Hsiao (2009) pointed out, after interviewing several outstanding teachers that their behaviours, attitudes, appearance, and character may influence the image students perceive, and may even influence the interaction between teacher and student.

Third, he/she attaches importance to teaching methods (Greimel-Fuhrmann & Geyer, 2003; Voss & Gruber, 2006). Research indicates that expert teachers have a large repertoire of teaching skills and knowledge about when to use them (Stronge, 2002; Walls, Nardi, von Minden, & Hoffman, 2002). The ability of teachers to choose the most suitable teaching method from a variety of teaching tools is important to students as teachers can then offer interesting lessons, which results in students paying better attention. Bain (2004) concluded that the best teachers use their knowledge to develop techniques for grasping fundamental principles and organizing concepts that others can use to begin building their own understanding and abilities. Effective teachers are more successful in their use of classroom time, more organized, establish effective daily routines, and have successful classroom management skills (Stronge, 2002).
When we say a teacher is charismatic, he/she not only displays the above-mentioned characteristics but, in particular, employs teaching methods that are not boring. That is, the teacher is identified as good and the class must be interesting or attractive, so students like to attend the course. A “sense of humor from the teacher” is always identified as an important factor that seriously influences students’ learning, and might be considered as a strong indicator of a charismatic teacher.

There is some support for this assertion. Minchew (2001) and Neumann, Hood and Neumann (2009) pointed out that students do prefer listening to teachers who incorporate humor into the lecture. Humor has the power to make instructors more likable, approachable, facilitate comprehension, increase attentiveness, improve creativity, and promote social relationships if using it appropriately (Lei, Cohen, & Russler, 2010; Minchew, 2001). Humor was found to lighten the mood during lectures and helped reduce stress and anxiety in students (Neumann et al., 2009). According to Voss and Gruber (2006), students think that they can perform well if the atmosphere in class is supportive, thus can be positively influenced by the perceived humor and friendliness of the teacher. When college students were asked to identify what makes a good instructor, among the first characteristics noted is a sense of humor (Lei et al., 2010). Students often remember their favorite teachers as being those who created an interesting environment. No wonder researchers documented a positive relationship between instructors’ use of humor and student evaluations of instructors (Torok, McMorris, & Lin, 2004). Therefore, in addition to being knowledgeable, having positive character traits, and using great teaching techniques, a teacher has a good sense of humor is likely to be considered charismatic.

From the viewpoint of Vevere and Kozlinskis (2011), students' surveys, aiming at evaluating the teaching quality, have to consider the most valuable factors of the teaching quality and qualities of lecturers, which comprise knowledge transfer, knowledge evaluation, accessibility of a lecturer and his/her personality traits. Based on our review of the teaching literature, we identified these four key factors (knowledge, character traits, teaching techniques, and humor) that are the valuable factors for students’ perception and that should provide a suitable framework for further work on toward understanding teacher charisma.

In reviewing a number of literatures concern with teacher’s charisma, Li and Liu (2009) concluded with two comments. First, the previous studies primarily focused on the charisma about teacher’s personality, only a few studies paid attention to teaching charisma. Second, how to measure or assess the teaching charisma of a teacher is always an important issue. Additionally, although there were many studies concerning teacher effectiveness and teaching quality, the majority are theoretical (Bain, 2004; Brown, 2004; Greimel-Fuhrmann & Geyer, 2003; Hill et al., 2003; Murray, 1997; Pozo-Munoz et al., 2000; Voss & Gruber, 2006; Walls et al., 2002). They suggested some behaviours or characteristics a teacher should have to encourage students’ learning, but few provide specific instruments or methods to evaluate students’ perceptions of teachers’ teaching. Accordingly, in order to advance the effectiveness of higher education, we concentrated our attention on the teacher’s teaching and aimed to develop an inventory for evaluating teacher’s charisma in college classroom. The newly developed instrument can be used to help teachers improve teaching skills to enhance students’ learning.

**Study purpose**

What are the essential qualities of a charismatic teacher? How can we help people to become charismatic teachers? As Voss and Gruber (2006) mentioned, if lecturers know what their students expect, they may be able to adapt their behaviour to their students’ underlying expectations, which should have a positive impact on their perceived service quality and their levels of satisfaction. In order to gain further insight into students’ perspectives of teaching charisma, we decided to explore the issue more profoundly. The purpose of the study was to develop a teaching evaluation questionnaire, the Inventory of Teaching Charisma in College Classroom (ITCCC), based on students' learning experiences and selected teacher characteristics identified as indicators of teaching quality.

**Method**

The first phase of this study was to generate an initial pool of items to measure the teaching charisma of a teacher in college classroom. Then an exploratory factor analysis (EFA) was conducted to assess the factor structure of the scale items. The second phase of this study used the confirmatory factor analysis (CFA) to test the factor structure of
the scale obtained from the first stage through the other independent sample. The Cronbach’s alpha coefficient was calculated to examine the internal consistency of the constructed scale. Finally, the criterion-related validity was also examined.

Scale construction

The initial phase of the development and item generation of the ITCCC was informed by a comprehensive review of teaching theory (e.g., Bain, 2004; Hsiao, 2009; Koehler & Mishra, 2009; Lei et al., 2010). Based on a review of the literature, we conceptualized the teaching charisma is composed of four interrelated facets, including knowledge, character traits, teaching techniques, and humor. The "knowledge" facet concerns with the professional knowledge and pedagogical knowledge which the teacher possess. The "character traits" facet concerns with the teacher’s performance with respect to behaviours or qualities such as friendly, approachable, patient and responsible. The "teaching techniques" facet concerns with teacher’s teaching techniques such as creative method or unique style. The "humor" facet concerns with the teacher’s humorous style in the classroom such as interesting talk or funny joke.

Following facet definition, we constructed a pool of 36 items to reflect the four facets. Item generation began through a brainstorming session by a group consisting of 40 undergraduates. Throughout the session, participants offered suggestions by which teaching charisma could be assessed to aid in the construction of initial items. As a result, the initial version of the ITCCC administered to participants contained 36 items.

To help establish content validity, we identified 4 scholars in department of guidance and counseling to evaluate our preliminary work. These experts evaluated the clarity of the facet descriptions for character traits, humor, knowledge, and teaching techniques. The experts provided revision suggestions for facet descriptions they judged to be unclear and were asked to provide recommendations for additional items for any or all of the identified facets. Expert feedback resulted in the elimination of 6 items due to lack of clarity, conciseness, or relevance to the construct of teaching charisma and 2 items added to enhance the representativeness of the entire item pool. Overall, the revision process led to a final pool of 32 items.

Participants

In the current study, we set “Calculus,” a professionally required course for freshmen of college of business in Taiwan, as the case study curriculum and targeted on population of years 2-4. For the first stage, the convenience sample consisted of 283 students volunteers recruited from 8 undergraduate classes in 4 midland universities of Taiwan. Participants’ ages ranged from 21 to 26 (mean = 22.61, SD = 2.55), and 54% were female. For the second stage, a total of 1,300 questionnaires were administered from the same population (but different classes) as above, and of these a total of 1,078 were valid. Among the valid sample, 56% of the participants were female. Participants' age ranged from 21 to 25 years old (mean = 22.23, SD = 2.10). Table 1 shows characteristics of the participants.

<table>
<thead>
<tr>
<th></th>
<th>the first stage</th>
<th>the second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>129 (46%)</td>
<td>478 (44%)</td>
</tr>
<tr>
<td>female</td>
<td>154 (54%)</td>
<td>600 (56%)</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>second year (sophomore)</td>
<td>113(40%)</td>
<td>495(46%)</td>
</tr>
<tr>
<td>third year (junior)</td>
<td>93(33%)</td>
<td>313 (29%)</td>
</tr>
<tr>
<td>fourth year (senior)</td>
<td>77(27%)</td>
<td>270 (25%)</td>
</tr>
<tr>
<td>Major</td>
<td></td>
<td></td>
</tr>
<tr>
<td>international trade</td>
<td>85 (30%)</td>
<td>345 (32%)</td>
</tr>
<tr>
<td>business administration</td>
<td>68 (24%)</td>
<td>226 (21%)</td>
</tr>
<tr>
<td>computer engineering</td>
<td>71 (25%)</td>
<td>183 (17%)</td>
</tr>
<tr>
<td>insurance and finance</td>
<td>28 (10%)</td>
<td>162 (15%)</td>
</tr>
<tr>
<td>applied statistics</td>
<td>20 (7%)</td>
<td>140 (13%)</td>
</tr>
<tr>
<td>others</td>
<td>11 (4%)</td>
<td>22 (2%)</td>
</tr>
<tr>
<td>Total</td>
<td>283</td>
<td>1,078</td>
</tr>
</tbody>
</table>
By the way, we have 4 and 17 instructors in the two sampling stages respectively. Each instructors provided about 2 participating classes which each includes 35-45 students. After the questionnaires were collected, we checked the rating scores for each class in order to examine the consistency for the same instructor. As an expected result, each participating class rated their instructor similarly. It provides some information as the validity of the instrument.

Measures

*Inventory of Teaching Charisma in the College Classroom (ITCCC).* The formal version of the ITCCC consisted of 32 items and measured the students’ perceptive degree of teaching charisma from his/her teacher. Students rated the resulting 32-item on the extent to which they agreed with each statement using a 5-point Likert-type scale, ranging from 1 (never true) to 5 (always true). The higher the score, the better the degree of teaching charisma.

*Learning Engagement Scale for College Students.* The instrument applied for examination of the criterion-related validity about ITCCC is "Learning Engagement Scale for College Students" (Lin & Huang, 2012). The scale contains 5 facets: skills engagement, emotional engagement, performance engagement, interaction engagement, and attitude engagement. It is a 5-point Likert-type scale with ratings from 1 (extremely disagree) to 5 (extremely agree), so that the higher the score, the better the degree of learning engagement. According to their findings, the Cronbach’s alpha coefficient for the entire scale was .86, and the coefficients of internal consistency for 5 subscales ranged from .71 to .79. The results from EFA, CFA, and cross-validation examination suggested that the "Learning Engagement Scale for College Students" had good reliability and validity. The cross-validation for the model was also examined by Lin & Huang (2012).

*Learning Satisfaction Scale for College Students.* The other instrument applied for examination of the criterion-related validity about ITCCC is "Learning Satisfaction Scale for College Students" (Kuo, 2010). The original scale contains 6 facets: teacher’s teaching, content of curriculum, academic achievements, interpersonal relationship, administration, and learning environment. It is a 4-point Likert-type scale with ratings from 1 (extremely disagree) to 5 (extremely agree), so that the higher the score, the better the degree of learning satisfaction. According to Kuo (2010), the construct validity and internal consistency of the scale were examined, and the scale was applicable for college students. In the present study, in order to focus on teacher’s teaching behaviours in the classroom and also for avoiding too many questions, only some subscales obviously involving teacher’s teaching were carried out. Three subscales were extracted and adopted: teacher’s teaching (8 items), content of curriculum (5 items) and academic achievements (6 items).

Procedure and data analyses

After conceptualizing the ITCCC and developing an item pool, we sought and obtained approval to conduct the research investigation at the 4 universities in 2011 fall semester. We explained the purpose of the study to the target students and obtained consent to participate in the study. The entire data set was scrutinized to detect missing values, invalid values and outliers. The CFA was conducted using maximum likelihood estimation procedures of LISREL and all other analyses were executed using SPSS.

To examine the psychometric properties and the factorial validity of the ITCCC, an EFA was conducted by the first sample. We reviewed the results of Bartlett’s sphericity test to clarify the factorability of the data and Kaiser-Meyer-Olkin to measure the sampling adequacy, which was favorable at a level of .90. In the social sciences, the behaviour is rarely partitioned into neatly packaged units that function independently of one another, some correlation among factors are generally expected (Costello & Osborne, 2005). Thus, we selected principal component analysis and promax rotation to identify a few coherent constructs that best reflect the teaching charisma of a teacher in college classroom. To form the retention of the components and to avoid under or over factoring, we established an a priori criterion for the inclusion of items: only those with a loading of .4 or higher would be considered in the development of the ITCCC. Cronbach’s alpha coefficient was calculated to determine the internal consistency of the ITCCC’s underlying constructs.

A CFA was conducted by the second sample to examine the factor structure of the ITCCC based on the findings from the EFA. The fit was evaluated using the chi-squared test, which helps indicate appropriate data model fit;
Results

Exploratory factor analysis (EFA) (n=283)

Item analysis was carried out first. Item-total correlations of the scale were found upon .53 except for 3 items (r < .30), thus they were deleted. The differences between mean scores of the upper 27% and lower 27% were examined for remainder items. The results from t test showed significant differences between each item’s means of the upper 27% and lower 27% points. Hence 29 items were retained.

Since the Kaiser-Meyer-Olkin measure of sampling adequacy index was .95, and Bartlett's sphericity test was significant, $\chi^2 = 6630.32$, $p < .0001$, indicating that the sample and correlation matrix were appropriate for the factor analysis. We performed EFA on the 29-item ITCCC. The number of factors to retain was based on a combination of methods (e.g., parallel analysis, eigenvalue>1.0, scree plot) as well as conceptual clarity, interpretability and theoretical salience of the rotated factors, and simple structure. As a result, a four-factor model was established, accounting for 68% of the variation. However, if an item did not load at .4 or higher, it was not retained. Our preliminary analysis showed that 6 candidate items did not have significant loadings and were discarded. We then ran a second EFA with the remaining 23 items, which cohered into four unique factors. The resulting four-factor model accounted for 70% of the variation.

Based on the 23-item ITCCC, the EFA produced four factors, and they were labeled in the following way: Factor 1, character traits, consists of 6 items and concerns with the teacher’s performance with respect to behaviours and morals. Factor 2, knowledge, consists of 7 items and concerns with the professional knowledge and pedagogical knowledge which the teacher possess. Factor 3, humor, consists of 6 items and concerns with the teacher’s humorous style in the classroom. Factor 4, teaching techniques, consists of 4 items and concerns with teacher’s teaching techniques. The items of 4 factors and factor loadings are presented in Table 2.

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. character traits</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My teacher has a lot of patience.</td>
<td>.93</td>
<td>-.10</td>
<td>-.21</td>
<td>.19</td>
</tr>
<tr>
<td>My teacher is fair and objective in grading.</td>
<td>.81</td>
<td>.09</td>
<td>.03</td>
<td>-.23</td>
</tr>
<tr>
<td>My teacher is very responsible.</td>
<td>.81</td>
<td>.02</td>
<td>.14</td>
<td>-.09</td>
</tr>
<tr>
<td>My teacher has good moral characteristic.</td>
<td>.79</td>
<td>.02</td>
<td>.23</td>
<td>-.20</td>
</tr>
<tr>
<td>My teacher is very democratic and can accept students’ different opinions</td>
<td>.73</td>
<td>.07</td>
<td>-.09</td>
<td>.22</td>
</tr>
<tr>
<td>My teacher practices what he/she preaches, sets a good example for us</td>
<td>.72</td>
<td>.05</td>
<td>.23</td>
<td>-.15</td>
</tr>
<tr>
<td>2. knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My teacher can solve all of the course-related problems.</td>
<td>-.04</td>
<td>.94</td>
<td>-.11</td>
<td>-.06</td>
</tr>
<tr>
<td>My teacher is an expert in this field.</td>
<td>-.03</td>
<td>.91</td>
<td>-.12</td>
<td>.02</td>
</tr>
<tr>
<td>I admire teacher’s high level of proficiency in this field.</td>
<td>.07</td>
<td>.82</td>
<td>-.02</td>
<td>-.13</td>
</tr>
<tr>
<td>My teacher prepares rich materials for the lessons.</td>
<td>.09</td>
<td>.63</td>
<td>.04</td>
<td>.21</td>
</tr>
<tr>
<td>My teacher has a wealth of knowledge.</td>
<td>.11</td>
<td>.61</td>
<td>.19</td>
<td>-.04</td>
</tr>
<tr>
<td>My teacher applies simple teaching methods that help me to understand the curriculum.</td>
<td>.01</td>
<td>.59</td>
<td>.09</td>
<td>.24</td>
</tr>
<tr>
<td>My teacher has a wide range of knowledge covering many fields.</td>
<td>.10</td>
<td>.57</td>
<td>.11</td>
<td>.04</td>
</tr>
<tr>
<td>3. humor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>We are never bored in my teacher’s class.</td>
<td>-.01</td>
<td>-.03</td>
<td>.85</td>
<td>.01</td>
</tr>
</tbody>
</table>
My teacher is a humorous teacher. My teacher often shares funny stories with us. My teacher teaches has fun teaching methods. My teacher’s teaching is very exciting. My teacher creates a fun and relaxed learning environment.

4. teaching techniques
My teacher uses some teaching materials that are new and interesting. My teacher often uses some new and non-traditional methods in teaching. My teacher is able to use new and creative ideas to stimulate our learning. My teacher uses some creative teaching techniques.

Note. Pattern coefficients with values of .40 or greater are in bold type.

Internal consistency and scale inter-correlations

Table 3 contains the inter-correlations, means, standard deviations, and internal consistency reliability values for each ITCCC subscale and for the total scale score. The internal reliability estimates for the individual subscales ranged from .89 (teaching techniques) to .93 (humor), providing support for their internal consistency. The ITCCC total scale produced an alpha coefficient of .95. Inter-correlations among the individual ITCCC subscales were generally medium to large, ranging from .35 to .74. These results echo the factor analytic findings suggesting that the various subscales represent overlapping, yet somewhat distinct, aspects of teaching charisma. Not surprisingly, the ITCCC total scale was strongly related to each of its component scales (rs = .75 to .88).

Table 3. Correlations, means, standard deviations, and internal consistency estimates

<table>
<thead>
<tr>
<th>Scale</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>M</th>
<th>SD</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. character traits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.67</td>
<td>.73</td>
<td>.90</td>
</tr>
<tr>
<td>2. knowledge</td>
<td>.73</td>
<td></td>
<td></td>
<td></td>
<td>3.65</td>
<td>.70</td>
<td>.90</td>
</tr>
<tr>
<td>3. humor</td>
<td>.54</td>
<td>.65</td>
<td></td>
<td></td>
<td>3.29</td>
<td>.81</td>
<td>.93</td>
</tr>
<tr>
<td>4. teaching techniques</td>
<td>.35</td>
<td>.50</td>
<td>.74</td>
<td></td>
<td>2.99</td>
<td>.85</td>
<td>.89</td>
</tr>
<tr>
<td>ITCCC, total score</td>
<td>.80</td>
<td>.88</td>
<td>.88</td>
<td>.75</td>
<td>3.45</td>
<td>.64</td>
<td>.95</td>
</tr>
</tbody>
</table>

Note. n = 283. All correlations are significant (p < .01).

Confirmatory factor analysis (CFA) (n = 1,078)

Subsequently, a CFA is conducted to analyze at what level a pre-determined or designed structure is confirmed by the second sample. Table 4 summarizes the results of the LISREL program. Except the $\chi^2/df$ value and AGFI were less appropriate, the set of fit indices suggested the data were well fit by the model: $\chi^2 = 1388.41, df = 224, p < .00$; GFI = .90 (≥.90); AGFI = .88 (close to .90); SRMR = .05 (≤ .08); RMSEA = .07 (≤ .08); NNFI = .98 (≥ .90); CFI = .98 (≥ .90).

Composite reliability was calculated using the formula suggested by Fornell and Larcker (1981). As shown in Table 4, all figures are above .8, which indicates that measures adopted for each construct are reliable (Fornell & Larcker 1981). Moreover, in terms of the constructs, all of the $t$-values of items showed statistical significance at the .05 level, indicating that all of those items within each subscale were highly correlated with each other and, therefore, revealed convergent validity. Further evidence of convergent validity is shown in Table 5. None of the correlations between the latent constructs was too high to challenge the convergent validity of the construct. The variance extracted test shows the discriminant validity. Discriminant validity is indicated when the extracted variance is greater than the square of the correlation between the two constructs (Fornell & Larcker, 1981). We compared all pairs of factors and they showed acceptable variance extracted.

These showed that the model had a highly satisfactory fit. The results of the CFA provided further preliminary support for the factor structure of the ITCCC established in former study.

Note. n = 283. All correlations are significant (p < .01).
Table 4. Summary of the confirmatory factor analysis

<table>
<thead>
<tr>
<th></th>
<th>Standardized factor loadings</th>
<th>t-values</th>
<th>$R^2$</th>
<th>Composite reliability</th>
<th>Variance extracted</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1: character traits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 5</td>
<td>.74</td>
<td>27.85</td>
<td>.55</td>
<td>.91</td>
<td>.64</td>
</tr>
<tr>
<td>Item 3</td>
<td>.79</td>
<td>30.72</td>
<td>.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 4</td>
<td>.84</td>
<td>33.76</td>
<td>.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 2</td>
<td>.86</td>
<td>34.71</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 6</td>
<td>.76</td>
<td>28.84</td>
<td>.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>.80</td>
<td>31.33</td>
<td>.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2: knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.91</td>
</tr>
<tr>
<td>Item 10</td>
<td>.75</td>
<td>28.16</td>
<td>.56</td>
<td></td>
<td>.58</td>
</tr>
<tr>
<td>Item 12</td>
<td>.80</td>
<td>31.19</td>
<td>.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 8</td>
<td>.78</td>
<td>29.83</td>
<td>.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 13</td>
<td>.78</td>
<td>29.90</td>
<td>.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 7</td>
<td>.78</td>
<td>29.69</td>
<td>.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 11</td>
<td>.75</td>
<td>28.28</td>
<td>.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 9</td>
<td>.70</td>
<td>25.83</td>
<td>.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3: humor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.92</td>
</tr>
<tr>
<td>Item 15</td>
<td>.79</td>
<td>30.68</td>
<td>.62</td>
<td></td>
<td>.67</td>
</tr>
<tr>
<td>Item 16</td>
<td>.87</td>
<td>35.87</td>
<td>.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 14</td>
<td>.67</td>
<td>24.57</td>
<td>.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 18</td>
<td>.86</td>
<td>35.28</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 19</td>
<td>.82</td>
<td>32.47</td>
<td>.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 17</td>
<td>.87</td>
<td>35.91</td>
<td>.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4: teaching techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.88</td>
</tr>
<tr>
<td>Item 23</td>
<td>.78</td>
<td>29.75</td>
<td>.61</td>
<td></td>
<td>.66</td>
</tr>
<tr>
<td>Item 21</td>
<td>.85</td>
<td>33.59</td>
<td>.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 22</td>
<td>.86</td>
<td>34.02</td>
<td>.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 20</td>
<td>.75</td>
<td>27.95</td>
<td>.56</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. All t-values are significant ($p < .01$).

Table 5. Construct correlation matrix

<table>
<thead>
<tr>
<th>Scale</th>
<th>character traits</th>
<th>knowledge</th>
<th>humor</th>
</tr>
</thead>
<tbody>
<tr>
<td>knowledge</td>
<td>.51</td>
<td>.51</td>
<td></td>
</tr>
<tr>
<td>humor</td>
<td>.64</td>
<td>.68</td>
<td></td>
</tr>
<tr>
<td>teaching techniques</td>
<td>.44</td>
<td>.75</td>
<td>.80</td>
</tr>
</tbody>
</table>

Internal consistency

To establish that each factor has satisfactory internal consistency, Cronbach’s alpha coefficient was calculated based on the second sample. The internal consistency coefficient of the overall scale was .95, and .91, .91, .92 and .88 for the four factors respectively, thus suggesting high consistency in the scale.

Criterion-related validity

We examined evidence of the measure’s criterion-related validity by assessing its relationships with learning engagement and learning satisfaction. Correlations of the ITCCC subscales to the criteria were shown in Table 6. As expected, the four ITCCC subscales were significantly and positively correlated with students’ learning behaviours, including their engagement and satisfaction in specific aspects (ranges were .20 to .60, and .20 to .66, respectively), and so was the ITCCC total score.
Table 6. Correlations of the ITCCC to the criterion variables

<table>
<thead>
<tr>
<th></th>
<th>ITCCC scale</th>
<th>ITCCC scale</th>
<th>ITCCC scale</th>
<th>ITCCC scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>character</td>
<td>knowledge</td>
<td>humor</td>
<td>teaching</td>
</tr>
<tr>
<td>learning engagement</td>
<td>traits</td>
<td>.20</td>
<td>.43</td>
<td>.35</td>
</tr>
<tr>
<td>scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>emotion</td>
<td>.21</td>
<td>.33</td>
<td>.44</td>
<td>.32</td>
</tr>
<tr>
<td>performance</td>
<td>.22</td>
<td>.39</td>
<td>.41</td>
<td>.40</td>
</tr>
<tr>
<td>interaction</td>
<td>.39</td>
<td>.42</td>
<td>.56</td>
<td>.50</td>
</tr>
<tr>
<td>total scale</td>
<td>.43</td>
<td>.59</td>
<td>.60</td>
<td>.54</td>
</tr>
<tr>
<td>learning satisfaction</td>
<td>teacher’s</td>
<td>.23</td>
<td>.61</td>
<td>.59</td>
</tr>
<tr>
<td>scale</td>
<td>content of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>academic achievements</td>
<td>curriculum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.31</td>
<td>.56</td>
<td>.46</td>
<td>.62</td>
</tr>
<tr>
<td>total scale</td>
<td>.44</td>
<td>.60</td>
<td>.61</td>
<td>.66</td>
</tr>
</tbody>
</table>

Note. n = 1,078. All correlations are significant (p < .01).

Discussion

The primary goal of the study was to construct a valid evaluation questionnaire that allows students to express their thoughts and perceptions concerning the teacher after they have completed a course. After developing and piloting the original, 32-item ITCCC, we conducted EFA to examine the initial factor structure and reliability of the scores. The exploratory study gave a valuable first insight into the characterization of a charismatic teacher in college classroom. The results of EFA helped us to refine the measure, and the resulting revised 23-item ITCCC was examined in the second stage (CFA). We established that the final revised ITCCC was a reliable and valid measure in the current sample.

The ITCCC identified in the factor analysis showed notable conceptual similarity to the constructs, initially drawn from the results of teaching literature and students’ opinion, proposed for measurement. In particular, the study results indicated that students perceived charismatic teachers to be humorous, knowledgeable, having great teaching methods, and possessing positive personality traits (e.g., approachable, friendly and patient). Consistent with the long-held view of the nature of the good teacher construct (Brown, 2004; Greimel-Fuhrmann & Geyer, 2003; Hill et al., 2003; Lammers & Murphy, 2002; Neumann et al., 2009; Voss & Gruber, 2006), together with the essential factor “the sense of humor” (Lei et al., 2010; Minchew, 2001; Torok et al., 2004), the ITCCC data formed four factors in the current study. Furthermore, criterion-related validity of the ITCCC was established by its predicted relations with students’ learning behaviours. Teacher’s charisma was found statistically significant positive relationships with students’ engagement and satisfaction in learning. According to this result, the teaching charisma increases the engagement and satisfaction in student’s learning. It is the evident support that teacher’s teaching behaviours play an important role in students’ learning.

In fact, the objective of this study is not to present definitive definition to a charismatic teacher. We believe that the criterion may be different depending on the context, and perhaps it is even impossible or pedagogically undesirable to formulate a definitive description of “the charismatic teacher.” Nevertheless, we do intend to offer a general framework for an applicable starting point for helping teachers to perform well. The analyses have consistently provided evidence to support the validity and reliability of the ITCCC in order to measure the teacher’s charisma in teaching. The ITCCC can function as a tool of exploration for teachers to increase their awareness of personal strengths and weaknesses and to make changes that may improve the teaching effectiveness. So we suggest the instrument can be administered during or at the end of the course so the results can offer useful information for improving teaching for the remainder or following courses. An advantage of the instrument is that it has few items and is therefore easy to administer and to complete (it takes less than 10 minutes). It will result in good response rate from students and the analysis is straightforward, since the scores on the items are simply summed to get the subscale scores. Apart from that, this study has provided useful evidence to support usage of this inventory for future studies in this area.
The results imply some suggestions for teachers. First, what a teacher knows is one of the most important influences on what is done in classrooms and ultimately on what students learn. Students want teachers to be knowledgeable (Hill et al., 2003; Lammers & Murphy, 2002), thus teachers should pursue further knowledge constantly. Second, teachers are expected to be good role models for the students; teachers should perform what a teacher should have (Chou, 1997). Third, it is beneficial to teacher’s teaching to incorporate humor in the classroom. Students prefer listening to teachers who incorporate humor into the lecture (Minchew, 2001; Neumann et al., 2009). Finally, teachers should attach importance to teaching methods (Greimel-Fuhrmann & Geyer, 2003; Voss & Gruber, 2006). Teachers should possess teaching skills and to be able to choose the most suitable teaching method from a variety of teaching tools.

Although these conclusions are interesting and thus may have implications for further research, we cannot ignore the limitations of the present study. The most important limitation of the present research was that the respondents were only from 4 universities in Taiwan. To make a generalization about the results of this study, the research needs to be repeated at different regions or countries. A second limitation concerns the absence in our model both of the score and the prior subject interest factors about the course; they might influences students’ rating. It is our intention to incorporate questions concerning these factors in a further research project on this subject.

Further questions arise concerning the correlation between scores and student ratings. That is, students with higher scores in the course tend to give more positive evaluations of the teaching in the course. Therefore, we suggest exploring more profoundly in order to gain further insight into students’ perspectives of teaching and possible bias influencing student ratings. Besides, the measurement invariance of the current scale across different populations (e.g., different genders, different subjects) needs to be examined. These tests are necessary in order to infer that differences in observed variables are indeed a function of a true quantitative difference between groups on the construct of interest and not due to different conceptualizations of the construct or bias. In other words, differences should be tested only after the factorial equivalence of the measures’ internal structure has been established for the groups involved.

**Conclusion**

Evaluating students’ perceptions of the quality of a teacher’s teaching is an important issue for higher education. The inventory presented here is an initial effort toward developing a psychometrically valid instrument that can accurately assess these perceptions and that can be easily administered and quickly scored by instructors. The development and item generation of the ITCCC was informed by a comprehensive review of teaching literature and the feedback of interviewed students. The results of the present study indicated that the ITCCC is a valid and reliable tool, providing a comprehensive framework with which to measure the teacher’s charisma in college classroom. The ITCCC provides measurable indices for assessing four aspects of teaching charisma: knowledge, character traits, teaching techniques, and humor. Unlike student evaluations of overall quality or effectiveness, these indices offer teachers specific feedback about key aspects of teaching that engage and enhance student learning. By the measured results, interventions designed to enhance teaching charisma might aim to encourage or enhance teachers’ techniques or qualities, and so help teachers to become more charismatic, and thus promote students’ learning. A next logical step in this line of investigation might focus on attempts to modify charisma by changing these teaching behaviours and measuring the impact on student learning. Future researchers may also find other important factors that contribute to teacher’s charisma, such as enthusiasm and tolerance. Researchers are also encouraged to apply this scale in different countries and cultural backgrounds. Subsequent studies could investigate the relationships among teacher’s charisma, students’ academic achievement, satisfaction, engagement, etc. A prudent exploration of teacher’s behaviours regarding the effect upon students’ learning in these areas could provide additional insight into the problems of education.

**References**


An Investigation of the Effects of Different Types of Activities during Pauses in a Segmented Instructional Animation

Jongpil Cheon¹, Sungwon Chung¹, Steven M. Crooks¹, Jaeki Song²* and Jeakyeong Kim³

¹College of Education, Box 41071, Texas Tech University, Lubbock, Texas, U.S.A. // ²Rawls College of Business Administration, Texas Tech University, Lubbock, Texas, U.S.A. // ³The School of Management, Kyunghee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul, Korea // jongpil.cheon@ttu.edu // sungwon.chung@ttu.edu // steven.crooks@ttu.edu // jaeki.song@ttu.edu // jaek@khu.ac.kr

*Corresponding author

(Submitted February 21, 2013; Accepted June 28, 2013)

ABSTRACT
Since the complex and transient information in instructional animations requires more cognitive resources, the segmenting principle has been proposed to reduce cognitive overload by providing smaller chunks with pauses between segments. This study examined the effects of different types of activities during pauses in a segmented animation. Four groups were asked to do different tasks in system-controlled pauses after each segment of an instructional animation: passive pauses (i.e., no-reflection vs. reflection), and active pauses (i.e., free-recall vs. short-answer). The results showed that active pause with free-recall group outperformed the two passive pause groups on both recall and transfer tests. However, no significant differences in mental effort for the instruction or the tests were found. The findings of this study provide valuable implications for effective ways of using pauses between segments in instructional animations.

Keyword
Segmenting principle, Instructional animation, Pauses in segments, Active pauses

Introduction
Instructional animations are dynamic visualizations that display a series of pictures for educational purposes. They are often used to illustrate dynamic changes within complex processes by depicting the motion or trajectory of processes (Betrancourt & Tversky, 2000; Hoffler & Leutner, 2007), such as the formation of lightning (Schmidt-Weigand & Scheiter, 2011), or the motions of electrons (Yang, Andre, & Greenbowe, 2003). Animations have been shown to be more effective than static graphics in helping learners build mental models of processes involving change over time (Hoffler & Leutner, 2007; Wouters, Paas, & van Merriënboer, 2008)

However, animations are not always superior to static graphics, because they may impose additional cognitive load (Hegarty, Kriz, & Cate, 2003; Mayer, Hegarty, Mayer, & Campbell, 2005; Phan, 2011; Spanjers, Wouters, van Gog, & van Merriënboer, 2011; Tversky, Morrison, & Betrancourt, 2002). For example, the information portrayed in complex or fast-paced animations is frequently too transient to enable essential cognitive processing (Mayer & Moreno, 2003). In other words, the transiency of information may not provide learners with sufficient time to process all of the elements in an animation (Hegarty et al., 2003). This may hinder learning by inducing high cognitive load (Ayres & Paas, 2007; Spanjers, van Gog, & van Merriënboer, 2010).

However, animations can be designed to reduce cognitive overload (Ayres & Paas, 2007) by visually cueing important information (e.g., de Koning, Tabbers, Rikers, & Paas, 2007), presenting related information before animations (e.g., Mayer, Mathias, & Wetzell, 2002), or dividing animations into segmented pieces (e.g., Mayer, 2009; Spanjers et al., 2010). The segmenting principle proposes that animations depicting complex tasks should be divided into smaller parts in order to be processed effectively (Clark & Mayer, 2011; Mayer, 2009). In instructional animation research, segmentation has typically been operationalized by inserting a pause between segments of an animation. Pauses have been shown to provide learners with sufficient time to process information presented in the previous segment. In addition, pauses between segments appear to serve as event boundaries to enhance understanding of procedural information (Spanjers et al., 2010).

More recently, Cheon, Crooks, and Chung (in press) found an embedded question condition to be superior to a pause condition. In their study, university students who received embedded retention questions related to the previous animation segment (i.e., active pause) outperformed students who received only a pause in the animation segments
(i.e., passive pause). However, because the length of each pause was learner controlled, their findings may have resulted from participants spending more time to answer the embedded questions in the active pause condition than in the passive pause condition with no activity. The current study was designed to investigate the effects of active and passive pauses under system-paced conditions, thereby controlling for pause time.

### Instructional animation and cognitive load

Instructional animations have been found to provide both representational and aesthetic benefits. In terms of representational benefits, animations appear to facilitate the learners’ ability to transform the dynamic aspects of the animation into a robust mental model (Betancourt & Tversky, 2000; de Koning et al., 2007; Tversky et al., 2002; Phan, 2011; Wouters et al., 2008). This is likely because animations can depict continuous movement and trajectory as well as externalize abstract concepts, such as quantum mechanics and blood circulation. This visualization characteristic helps learners mentally simulate a process or procedure. Hoffler and Leutner’s meta-analysis (2007) showed that representational animations were superior to representational static pictures, whereas decorative animations were not. Animations appear to be especially beneficial for portraying procedural information that is difficult to describe in words (Wouters et al., 2008). The aesthetic benefits of animations pertain to helping learners to focus on the relevant parts of the animation because movement can hold learners’ attention. In addition, the authentic portrayals can increase learners’ motivation (Phan, 2011).

Even though instructional animations can facilitate learning by externalizing cognitive processes in specific areas, other studies have failed to find the superiority of the animation to static pictures. For example, Mayer et al. (2005) compared static illustration with printed text to narrated animation about four different topics (i.e., the process of lightning formation, how a toilet tank works, how ocean waves work, and how a car’s braking system works). They found that the paper groups outperformed the animation groups in four tests out of eight tests. In other studies, in a variety of learning domains, instructional animations did not yield better learning performance comparing static graphics in various learning domains (e.g., Byrne, Catrambone, & Stasko, 1999: computer algorithm; Hegarty et al., 2003: mechanical system; Hegarty, Narayanan, & Freitas, 2002: a flushing cistern; Morrison & Tversky, 2001: permissible social paths). Thus, Tversky et al. (2002) and Phan (2011) asserted that discrete steps with static graphics may be more beneficial than animations because learners mentally animate the key steps presented in static graphics.

As noted above, the limitations of instructional animations result from the transiency of information presented in animations. The effects of transiency on learning from an animation depend upon the amount, complexity, and speed of information presented in the animation. That is, the benefits of an animation may disappear as the amount, complexity, or speed of information exceeds the processing capacity of the learner (Hegarty et al., 2002; Hegarty et al., 2003; Mayer et al., 2005; Morrison & Tversky, 2001). This is primarily due to the fact that the learner “views one frame at a time, and once the animation has advanced beyond a given frame, it is no longer available …” (Hegarty, 2004, p. 346). In other words, the detrimental effects of transiency in an animation are experienced as new information keeps coming while learners try to process previously presented information (Ayres & Paas, 2007; Lowe, 2003). Under these conditions the transience of an instructional animation imposes excessive cognitive load on the learner (Spanjers, van Gog, Wouters, & van Merriënboer, 2012; Tversky et al., 2002). The cognitive overload imposed by animations has been explained for the perspective of cognitive load theory.

To learn specific information, working memory needs to maintain and process the information by transferring the information to and retrieving it from long-term memory (Baddeley, 2007). However, working memory is limited in its capacity to simultaneously process a large amount of information (Baddeley, 2007; Barrouillet & Camos, 2007; Lusk et al., 2009; Mayer & Moreno, 2003). Cognitive overload takes place when the demands on working memory exceed its capacity (Baddeley, 1992). In other words, cognitive overload negates the learner’s ability to process critical information in working memory (Baddeley, 2007; Clark & Mayer, 2008; Phan, 2011). When the transience of the information in animations exceeds the learners’ cognitive resources (Spanjers, van Gog, Wouters et al., 2012), cognitive overload will occur. Under these circumstances, instructional animations should be less effective than static graphics if the cognitive load associated with the static graphics goes beyond the learners’ capacity. Based on the assumption of limited cognitive capacity, cognitive load theory asserts that there are three different types of cognitive load: (a) intrinsic load, caused by the level of element interactivity in learning content, (b) extraneous load, caused by unnecessary cognitive processing, and (c) germane load, caused by cognitive processing that is essential for learning (Sweller, 2010; Sweller, van Merriënboer, & Paas, 1998; van Merriënboer & Sweller, 2005). Cognitive
overload occurs with excessive intrinsic or extraneous load, and effective instruction can be developed by optimizing the three types of cognitive load.

Previous multimedia learning research argued that animations could pose extraneous cognitive load because the information does not stay long enough to be processed (e.g., de Koning et al., 2007; Wouters et al., 2008). However, the extraneous load caused by animations can be reduced by a meaningful instructional medium. The multimedia learning principle suggests a number of ways to reduce extraneous cognitive load, such as the modality principle, the contiguity principle, and the segmenting principle (Clark & Mayer, 2011). In addition, Wouters et al. (2008) proposed that learner pacing and segmentation of dynamic visualization (i.e., animations) would reduce extraneous cognitive load and free learners’ cognitive capacity for learning.

**Pauses between segments**

The segmenting principle states that “people learn better when a multimedia message is presented in user-paced segments rather than as a continuous unit” (Mayer, 2009, p. 175). The segmentation strategy was first explored in the context of reading. Several studies found that segmented text, grouped in meaningful units, was more beneficial for learning (i.e., recall and comprehension) than continuous text (Florax & Ploetzner, 2010; Gaddy, van den Broek, & Sung, 2001; Glynn, Britton, & Tillman, 1985; Hartley, 1986; Weiss, 1983). Spatially segmented text appeared to provide learners with the opportunity to stop the flow of information (Lusk et al., 2009) and to process the information more deeply. In a similar manner, multimedia researchers have examined the effectiveness of presenting dynamic instructional animations in smaller pieces (i.e., segments) rather than in a continuous format (Clark & Mayer, 2011; Mayer, 2009; Spanjers et al., 2010; Spanjers, van Gog., & van Merriënboer, 2012; Spanjers et al., 2011). Several studies have demonstrated the positive effects of segmentation in instructional animations (e.g., Hasler, Kersten, & Sweller, 2007; Mayer & Chandler, 2001; Moreno, 2007; Spanjers et al., 2011; van Gog & Paas, 2008).

In previous studies, the segmenting principle was applied as pauses between segments of the animation. The benefits of pauses have been explained from two viewpoints: (a) extra time, and (b) event boundaries. First, pauses are purported to reduce cognitive load by giving learners extra time to transfer information from the previous segment to long term memory and to prepare to process information in the next segment (Mayer & Moreno, 2003; Spanjers et al., 2010). Second, pauses are purported to provide temporal cues for chunking information. When continuous animations are divided into meaningful chunks, the segments provide natural boundaries between procedural events (Schwan, Garsoffky, & Hesse, 2000; Spanjers et al., 2012; Spanjers, Van Gog, Wouters et al., 2012).

Animation pauses have been investigated in previous studies within the context of either learner or system controlled pauses. In system-controlled pauses, computers control the time between segments. In learner-controlled pauses, learners control the time between segments. In studies employing learner-controlled pauses, Mayer and colleagues compared college students receiving either segmented or non-segmented animations about the formation of lightning and the functioning of electric motors. Their results showed that students studying the segmented versions outperformed those studying the non-segmented versions on transfer tests (Mayer & Chandler, 2001; Mayer, Dow, & Mayer, 2003). In another study, Hasler et al. (2007) compared two different learner-controlled conditions (i.e., pauses between segments vs. providing learners with stop and play buttons throughout the animation) to two continuous conditions (i.e., text and narration vs. narration only animations). They found that the two learner-controlled conditions outperformed two continuous conditions on the more difficult test items. They also found that the condition with learner-controlled pauses between segments was especially beneficial for novice learners (Moreno, 2007) and learners with low working memory capacity (Lusk et al., 2009). Spanjers et al. (2011, 2012) conducted two studies comparing system-controlled pauses to a continuous animation (i.e., no segmentation). The results of their first study showed that students studying animations with system-controlled pauses between segments learned more efficiently (i.e., invested less mental effort with no loss in performance) than students studying a continuous animation (i.e., no segmentation). Their second study showed that the segmentation group with system-paced pauses outperformed the continuous animation group on a posttest.

Spanjers et al. (2012) conducted another segmentation study designed to increase learner engagement by having learners create their own segments within a series of written worked examples. However, the results showed that asking learners to create their own segments required more mental effort (and reduced learning efficiency) than
presenting learners with predefined segment pauses. Learner created segmentation appears to require extensive cognitive resources that may overload the cognitive system. Even though their study did not use animation, the results can be explained by the guidelines for designing effective video-based model prosed by Wouters, Tabbers, and Paas (2007). Their guidelines suggest that pacing with predefined segments would be beneficial for learners with low prior knowledge, whereas creating segments would be suitable for learners with high prior knowledge. However, no studies have compared the two pace control options with pre-defined segments in an instructional animation.

**Passive pauses and active pauses**

Other strategies, besides segmentation, have been studied in conjunction with animations in order to increase their effectiveness. For example, Paas, Van Gerven, and Wouters (2007) showed a sequence of key frames from an animation directly after the animation. In another study, Mayer et al. (2003, Experiment 3) gave learners a question before showing an animation about the working of an electric motor and told them that they would have to answer the question after the animation. Both of these methods successfully enhanced student learning by actively engaging students with the information presented in the animation. Interestingly, however, few studies have attempted to engage students in active processing activities between segments of an animation (i.e., during the pause period). This is an important area of research since pauses between segments of an animation are only effective to the degree that learners actually use the time during pauses to process relevant information (Wouters et al., 2008), and some researchers have questioned whether learners use pause time effectively (Spanjers, van Gog, Wouters et al., 2012). Therefore, the primary purpose of this study was to explore the effectiveness of providing different types of student engagement during pauses in an animation.

The conventional pause between animation segments can be referred to as a **passive pause** because no instructional activity is presented during the pause. On the other hand, an **active pause** occurs when an instructional activity is presented during the pause (e.g., answering a question) (Cheon et al., in press). Presenting learners with an engaging instructional activity during a pause in an animation is likely to make the pause period more meaningful, since the activity should enhance the cognitive processing of information from the previous segment. For example, Cheon et al. (in press) presented embedded questions that required learners to retrieve information from the previous animation segment. In their study, the active pause group outperformed a passive pause group on recall and transfer tests; the authors concluded that this kind of active pauses promote germane cognitive load by enhancing schema construction. However, the positive effects of the active pause in their experiment might have been caused by the extra time spent by learners who responded to embedded questions because the pauses in their study were learner-controlled. Little is known about the direct effects of embedded questions during an animation pause. Even less is known about how different types of instructional activities during pauses affect learning performance. In addition, the effects of active pausing on cognitive load (Paas, Tuovinen, Tabbers, & Van Gerven, 2003) should be taken into consideration. The instructional activity presented during a pause may increase learners' mental effort as they strive to understand the information depicted in the animation (Wouters et al., 2008). Alternatively, active pauses may decrease mental effort for a test because learners have already encoded and retrieved information by the pause activity.

The aim of the present study was to determine the relative advantage of active pauses over passive pauses on learning performance and mental effort in an instructional animation incorporating system-controlled pauses. Participants receiving passive pauses were instructed to either wait (i.e., no instruction to reflect) or to reflect on the information presented in the previous segment; participants receiving active pauses were asked to either write down everything they could remember from the previous segment (i.e., free recall) or answer short-answer test items during the pause period. The research questions were (a) How does pause type (active vs. passive) affect the learning performance of college students studying a segmented animation? and (b) How does pause type (active vs. passive) affect the mental effort of college students studying a segmented animation? It was hypothesized that active pauses would enhance learning outcomes (i.e., recall and transfer) more than passive pauses. Regarding mental effort, active pause groups were hypothesized to invest higher mental effort than the passive pause group during the instructional animation, but lower mental effort during the test phases.
Methodology

Participants and design

Ninety-nine undergraduate students from a large Southwestern university participated in this study (Female: 50, Male: 49; Freshman: 35, Sophomore: 38, Junior: 19, Senior: 7). The participants were enrolled in an undergraduate course on computer literacy, and the course had an instructional research module designed to provide students with an opportunity to participate in a study related to instructional technology. Participants were randomly assigned to one of the four experimental conditions (i.e., two passive pause groups and two active pause groups): (a) passive pause, no-reflection, n = 29, (b) passive pause, reflection, n = 25, (c) active pause, free-recall, n = 24, and (d) active pause, short-answer, n = 21.

Materials

Instructional animation

The instructional animation consisted of a 160-second instructional animation about the formation of lightning (see Figure 1) created with Adobe Flash software. The animation was based on the animation used by Mayer and Moreno (1998) and consisted of 16 frames, each depicting a step in the process of lightening formation. The animation was divided into four 40-second segments (four steps per segment). The animation was accompanied by a narration identical to the text used by Mayer and Moreno (1998). The narration explained each of the steps in the formation of lightning.

Differences between the conditions involved the type of activity required of students during the four pauses in the animation. The animation, including pauses, was system-controlled, and total time to complete the animation was 320 seconds (i.e., 40 seconds per segment and 10 seconds per pause). During pauses, the passive pause, no-reflection group was presented with a blank screen with the follow message typed in the middle of the screen: “Please wait. You will be moving to the next animation in 40 seconds.” The passive pause, reflection group was presented with the following message encouraging self-reflection on the material they had just studied: “Please try to remember what you saw in the previous animation. You have 40 seconds.” The active pause, free-recall group was presented with the following message: “Please type as much of the information from the previous section that you can remember. You have 40 seconds.” The active pause, short-answer group was presented with the following message followed by two short-answer questions pertaining to the previous segment, “Please answer the following questions. You have 40 seconds.”

Figure 1. A screenshot of instructional animation
Prior knowledge test

Prior meteorology knowledge was collected from the participants using items developed by Mayer and Moreno (1998) and consisted of a seven-item meteorology-knowledge checklist (e.g., “I know what a cold front is ___,” “I know what a low pressure system is ____”) and a self-rating that asked participants to rate their knowledge of meteorology on a 5-point scale (1 = very little; 5 = very much).

Learning performance tests

Learning performance was measured with two tests (i.e., recall test and transfer test). The recall test instructed the participants to type everything they could remember from the instructional animation they had just studied. Following Mayer and Moreno (1998), free recall performance was scored by awarding participants one point for each of the 19 idea units contained in the instructional text, resulting in a total of 19 points possible. The transfer test, also developed by Mayer and Moreno (1998), consisted of four open-ended questions that required participants to transfer their knowledge of the instructional animation to correctly answer the questions (e.g., “What could you do to decrease the intensity of lightning?”). The transfer test was scored by awarding from zero to 3 points per item, resulting in a total of 12 points possible.

Mental effort rating

A mental effort rating scale was used to measure mental effort invested in both the instruction and the test phases. The nine-point rating scale was developed by Paas (1992) and consisted of the following statement: “Please indicate how much mental effort you invested in this test ranging from 1 (Extremely low) to 9 (Extremely high).” This subjective self-rating measure is non-intrusive and has proven to be a reliable indication of the mental effort experienced by participants (Paas et al., 2003).

Procedure

The study was conducted in a computer lab with 15 to 20 participants in each experimental session. A research proctor began the study by instructing the participants to turn on their computer monitors and silently read the instructions pertaining to their randomly assigned condition. The participants were also informed that they would be given a series of brief assessments after their study of the material; however, they were not informed about the nature of the assessments. They were next asked to complete the participant questionnaire including demographic information and prior knowledge. The participants then studied their assigned instructional animation. Since the animation including pauses was system-controlled, all groups completed the animation at the same time. After studying the instructional animation, the participants completed the recall test and then the transfer test.

Results

Regarding prior knowledge of the learning material, there were no significant differences among groups in either of the prior knowledge measurements (Seven-item meteorology-knowledge checklist, $F(3, 98) = 1.616, \ p = .191$; self-rating question $F(3, 98) = 1.257, \ p = .294$).

The research questions in this study were investigated through one-way ANOVAs. The analyses showed that test scores were influenced by the types of pauses (research question #1). However, mental effort scores were not significantly different (research question #2). The means for each measure by group are presented in Table 1.

Regarding recall test scores, the results revealed significant differences among the groups, $F(3, 98) = 4.861, \eta^2 = .365, \ p = .003$. Post-hoc comparisons indicated that the mean scores for the active pause, free-recall group ($M = 10.54, \ SD = 3.13$) was significantly higher than both passive conditions: no reflection ($M = 7.31, \ SD = 3.24$) and reflection ($M = 8.08, \ SD = 3.51$). However, the active pause, short-answer group ($M = 8.86, \ SD = 2.65$) was not significantly different from any of the other groups.
Table 1. Means and standard deviations for recall, transfer, and mental effort by treatment group

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Passive pause, no-reflection&lt;br&gt;n = 29</th>
<th>Passive pause, reflection&lt;br&gt;n = 25</th>
<th>Active pause, free-recall&lt;br&gt;n = 24</th>
<th>Active pause, short-answer&lt;br&gt;n = 21</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall test scores</td>
<td>( M = 7.31 ) (SD = 3.24)</td>
<td>( M = 8.08 ) (SD = 3.51)</td>
<td>( M = 10.54 ) (SD = 3.13)</td>
<td>( M = 8.86 ) (SD = 2.65)</td>
<td>.003*</td>
</tr>
<tr>
<td>Transfer test scores</td>
<td>( M = 3.93 ) (SD = 2.25)</td>
<td>( M = 4.52 ) (SD = 2.14)</td>
<td>( M = 5.58 ) (SD = 1.31)</td>
<td>( M = 4.90 ) (SD = 1.61)</td>
<td>.019*</td>
</tr>
<tr>
<td>Mental effort for instruction</td>
<td>( M = 5.79 ) (SD = 1.42)</td>
<td>( M = 5.64 ) (SD = 1.63)</td>
<td>( M = 6.21 ) (SD = 1.50)</td>
<td>( M = 6.00 ) (SD = 1.70)</td>
<td>.604</td>
</tr>
<tr>
<td>Mental effort for recall test</td>
<td>( M = 6.59 ) (SD = 1.30)</td>
<td>( M = 6.00 ) (SD = 1.47)</td>
<td>( M = 6.50 ) (SD = 1.14)</td>
<td>( M = 6.67 ) (SD = 1.32)</td>
<td>.285</td>
</tr>
<tr>
<td>Mental effort for transfer test</td>
<td>( M = 6.38 ) (SD = 1.18)</td>
<td>( M = 5.68 ) (SD = 1.28)</td>
<td>( M = 6.08 ) (SD = 1.44)</td>
<td>( M = 6.52 ) (SD = 1.47)</td>
<td>.138</td>
</tr>
</tbody>
</table>

There was also a significant transfer effect, \( F(3, 98) = 3.459, \eta^2 = .314, p = .019 \). Post-hoc comparisons indicated that the transfer test scores in the active pause, free-recall group (\( M = 5.58, SD = 1.31 \)) were significantly higher than the scores in the passive pause, no reflection group (\( M = 3.93, SD = 2.25 \)). There were no significant differences among the other groups.

There was no effect for mental effort for instruction, \( F(3, 98) = .620, p = .604 \), recall test, \( F(3, 98) = 1.281, p = .285 \), and transfer test, \( F(3, 98) = 1.884, p = .138 \). Taken together, these results suggest that the active pause, free-recall condition positively affected learning over the passive pause groups without increasing the mental effort of the participants.

**Discussion and conclusion**

This study investigated whether active pauses in an instructional animation would increase learning outcomes and mental effort more than passive pauses. The results partially supported this hypothesis in that the only one active pause group (i.e., free-recall condition) outperformed both passive pause groups (i.e., no-reflection and reflection) on a free recall test and a transfer test. Although the effect of the active pause, short-answer condition was not significant, the means for both recall and transfer were higher than the two passive pause groups (i.e., no reflection and reflection). However, neither mental effort in the instruction nor the tests was influenced by active pauses. In the following paragraphs we present implications relating to the advancement of the literature on the segmenting principle and the practice of instructional design.

The positive effects of the active pause, free-recall condition (i.e., free recall test: \( p = .003 \), transfer test: \( p = .019 \)) can be explained as resulting from an emphasis on germane cognitive load as opposed to extraneous cognitive load. Based on the limitations of working memory when processing large amounts of information (Baddeley, 2007), the segmenting principle asserts that presenting a continuous animation in small chunks can reduce learners’ cognitive load (Mayer & Moreno, 2003; Moreno & Mayer, 2007). Thus, learners can process previously presented information during pauses between segments. The positive effect of pauses has been found in the previous studies, but the effect is clearly dependent upon the metacognitive skill and/or motivation of the learners, as they may not use the pause time effectively (Spanjers, van Gog, Wouters et al., 2012). The conventional pause was called as a passive pause in this study because there is no required task for learners. On the other hand, an active pause was proposed using embedded assessments between segments. We hypothesized that while the passive pause condition may reduce extraneous cognitive load, the active pause conditions should promote generative cognitive load by facilitating schema construction. The findings are consistent with the results obtained by Cheon et al. (in press) and more clearly illustrate the effects of an active pause condition. The previous study found that a short-answer question group...
outperformed a no-reflection pause group, but these results might have been caused by the additional time spent by the active pause group. Therefore, the present study controlled the pause time to investigate whether learning performance differences were caused solely by the additional time afforded by the embedded questions. The results of the current study show that the active pause likely facilitates germane cognitive load by requiring learners to actively process material from the prior segment. Interestingly, it was the group asked to free recall rather than respond to short-answer questions that was responsible for the learning differences. We conjecture that because the short-answer questions covered only two steps, out of four steps, in each segment they may have focused on only half of the material in each segment. Also, learners may not have had enough time to fully take advantage of the stimulus, because they had to read and understand two questions first in order to type their responses. Thus, with a system-controlled pace, we contend that free recall during pauses is superior to short answer.

Contrary to our expectations, the current study did not find a mental effort effect. For the instruction with animation, we hypothesized that mental effort in the active pause groups would be higher than the other groups because both passive and active pauses already reduce extraneous cognitive load, but the questions in the active pause groups required more mental effort to answer the questions. This assumption was partially supported in that the mental effort ratings in both active pause groups were higher than the passive pause groups, but not statistically significant. Mental effort ratings during the tests were even less disparate, even though we predicted that mental effort would decrease in active pause groups since they could easily retrieve information from their long-term memory. We speculate that the inconsistent mental effort level may be caused by the complexity of the learning contents. The lower mean scores across the tests is an indication that the tests were complicated in general.

From a practical point of view, this study provides instructional designers with practical implications regarding a potential role of active pauses between segments in instructional animations. The general goal for instructional design is to minimize extraneous load and maximize germane load to a level that remains within working memory limits (Wouters et al., 2007). The current findings expand the knowledge base established from previous studies, such as the effects of asking questions before an animation (Mayer et al., 2003), or showing key frames after the whole animation (Paas et al., 2007). In line with Spanjers, van Gog, Wouters et al.’s study (2012) in which additional time after each segment was more beneficial than additional time after whole animation, active pauses (e.g., free recall test) may be more useful for learners to process transient information. However, the active pauses should not be too complex to interrupt the flow of learning from the animation.

Our results and interpretations also yield a number of suggestions for further research. First, the questions used in the active pause conditions prompt the retrieval of information from the previous segment. However, a prompt to predict the next step could be used in active pauses. For example, Hegarty et al. (2003) asked learners to predict the behavior of a machine from a static diagram in order to enhance mental animation. The prediction activity may trigger cognitive processes to select and organize information from the previous segment in order to expect the next event. Moreover, learners’ prior knowledge can be another consideration of using active pauses. Depending on the level of prior knowledge, the active pause may affect learning differently. Further studies could examine the differences between reflection and prediction prompts during pauses in terms of prior knowledge. Second, this study predefined the length of the segments, but further studies may explore the effect of active pause depending on the length of the segments. Third, the topic of the instructional animation used in this study was meteorology. It seems that this subject was somewhat difficult for the participants. Future studies may use a different topic to examine the effects of active pauses with animations. Last, the primary benefit of using segmented animation is to reduce extraneous cognitive load. On the other hand, this study proposes another potential role of pause that is to increase germane cognitive load. The mental effort level gauges the overall load, but it would be useful to have an instrument that measure different types of cognitive load (Mayer & Moreno, 2003; Spanjers, van Gog, Wouters et al., 2012).

Instructional animations can be a promising tool to deliver complex and procedural information, but the transient nature of animated information may strain learners’ cognitive system and hinder their learning. In other words, instructional animations may require significant cognitive resources that exceed the limitations of the learners’ working memory. Thus, instructional designers should design multimedia instruction in a way that minimizes unnecessary cognitive load (Mayer & Moreno, 2003). The segmenting principle is proposed to be one of the ways to optimize instructional animations. Our approach of using active pauses between segments is based on the idea that learner’s cognitive processing could be stimulated during pauses. The potential role of active pauses can add more value to the segmenting principle when using animated instruction.
References


A Review of Research on Technology-Assisted School Science Laboratories

Chia-Yu Wang1*, Hsin-Kai Wu2, Silvia Wen-Yu Lee3, Fu-Kwun Hwang4, Hsin-Yi Chang5, Ying-Tien Wu6, Guo-Li Chiou1, Sufen Chen7, Jyh-Chong Liang8, Jing-Wen Lin9, Hao-Chang Lo10 and Chin-Chung Tsai7

1Graduate Institute of Education, National Chiao Tung University, Hsinchu 30010, Taiwan // 2Graduate Institute of Science Education, National Taiwan Normal University, Taipei 11677, Taiwan // 3Graduate Institute of Science Education, National Changhua University of Education, Changhua 500, Taiwan // 4Department of Physics, National Taiwan Normal University, Taipei 11677, Taiwan // 5Graduate Institute of Science Education, National Kaohsiung Normal University, Kaohsiung 82446, Taiwan // 6Graduate Institute of Network Learning Technology, National Central University, Taoyuan 32001, Taiwan // 7Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei 10607, Taiwan // 8Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taipei 10607, Taiwan // 9Department of Curriculum Design and Human Potentials Development, National Dong Hwa University, Hualien 97401, Taiwan // 10Department of Digital Content & Technology, National Taichung University of Education, Taichung 40306, Taiwan // cgw25@mail.nctu.edu.tw // hkwu@ntnu.edu.tw // silviawyl@cc.nctu.edu.tw // hwang@phy.ntnu.edu.tw // hsinyichang@nkucc.nknu.edu.tw // ytwu@cl.nctu.edu.tw // glehiou@mail.nctu.edu.tw // sufchen@mail.ntust.edu.tw // aljc@mail.ntust.edu.tw // jwlin@mail.ndhu.edu.tw // simonlo@mail.ntcu.edu.tw // cctsai@mail.ntust.edu.tw

*Corresponding author

(Submitted February 22, 2013; Revised May 7, 2013; Accepted June 21, 2013)

ABSTRACT

Studies that incorporate technologies into school science laboratories have proliferated in the recent two decades. A total of 42 studies published from 1990 to 2011 that incorporated technologies to support school science laboratories are reviewed here. Simulations, microcomputer-based laboratories (MBLs), and virtual laboratories are commonly used to assist students in engaging in laboratory activities, followed by remote laboratories, databases, and other miscellaneous technologies. We report the demographics and characteristics of these technology-assisted laboratory studies and provide examples to illustrate how technologies have facilitated science learning in laboratories for different subjects and science domains, with various levels of student involvement and components of investigations. Major findings of the reviewed articles are summarized to understand the effects of applying technologies in school laboratories. Incorporation of technologies in school science laboratories has changed students’ learning experiences in terms of the phenomena to be explored, their interactions with the natural phenomena or materials, and approaches to handling and making sense of data.

Based on our findings, possible directions for future research and emerging instructional issues regarding technology-assisted laboratories are discussed.

Keywords

Laboratory, Technology, Science learning, K-16

Introduction

Laboratory learning constitutes an indispensable part of science education. Laboratories can motivate students, provoke active learning, and convey practice of science (Linn, 1997). Lunetta, Hofstein, and Clough (2007) defined school science laboratories (SSLs) as “learning experiences in which students interact with materials or with secondary sources of data to observe and understand the natural world” (p.394). For decades, SSLs have played a unique role in science curricula, providing opportunities for students to engage in investigation processes. Through experiencing scientific investigations, students might develop the abilities necessary to do scientific inquiry as well as an understanding of scientific inquiry (National Research Council [NRC], 2011).

The recent reform documents for K-12 and for post secondary education have stressed the importance of technology in SSLs. In these documents, the learning goals of having students gain knowledge and adequate skills for using technologies while practicing authentic science are explicitly stated (NRC, 2011; College Board, 2009). SSLs are expected to gradually equip students with the abilities to use technologies to gather, analyze, and transform data, as well as to create models, to communicate and to collaborate with others. Activities that reflect features of authentic inquiry are relatively complex and cognitively demanding (Chinn & Malhotra, 2002). Supports from technologies
have been used to reduce cognitive load and overcome obstacles of learning in realistic experiments due to time, space, scale, and resources. A large-scale international survey, the Programme for International Student Assessment (PISA), administered by the Organization for Economic Co-operation and Development (OECD), has also begun to explore the possibilities of assessing students’ skills in science using simulated laboratory experiments (OECD, 2010). Technology is playing an ever-increasing role in laboratory learning.

The incorporation of technologies in laboratory instruction may change not only the layout and supplementary resources, but also the nature of learning and teaching. Given the rapidity with which these technologies are being implemented in SSLs, a review of empirical studies is timely to understand how features of technologies can be matched with learners’ needs and task demands as well as what the impacts on learning science in a technology-assisted laboratory are. Several reviews have been carried out to analyze and discuss the purposes and styles of SSLs, as well as the key elements that affect laboratory activities (e.g., Domin, 1999; Hofstein & Lunetta, 2004; Lunetta et al., 2007). Other reviews have in part examined the effects of technology on learning in science education or inquiry (e.g., Lee et al., 2011; Ma & Nickerson, 2006; Nelson & Ketelhut, 2007; Scalise, Timms, Moorjani, Clark, Holtermann, & Irvin, 2011). However, the combination of technology and SSLs has not been well documented in previous reviews. Thus, the purpose of this study was to conduct a systematic analysis of the general trend of applying technology to assist SSLs, including types of technology and their effectiveness. Our findings can be used as a springboard for the design and implementation of technology-assisted laboratories that support various aspects of scientific inquiry.

In the subsequent sections, we define the crucial concepts used in this study. Empirical studies that incorporate technologies to assist laboratory learning are analyzed using the narrative review method. Through the analyses, we hope to pinpoint some less explored areas in the current trends of technology-assisted laboratory studies. Alignments among characteristics of technologies, student involvement, and elements of scientific investigation as well as technology affordances for important learning outcomes are then discussed. Although technologies have shown some advantages in assisting students’ learning of science in laboratories, which are presented in our findings, the incorporation of technologies may have changed students’ experiences of the real-world phenomena and of scientific investigation in comparison to their experiences in traditional laboratories. For instance, observing a simulated representation of an unobservable phenomenon (e.g., the particulate nature of matter for phase changes) may have changed the nature of observations. For this reason, we conduct a critical reflection to redefine laboratory experiences in technology-assisted laboratories. Through an in-depth review, we hope to generate some new insights for the future design of research and instruction on using technologies in SSLs.

**Method**

**Formation of an analytical framework**

We formed a panel of ten experts in different areas of science and science education to construct the analytical framework (Chen et al., 2012). To address our first aim regarding the types of technologies used in SSLs, our first category included demographics of the studies and two characteristics of the technology-assisted laboratories: types of technologies such as simulations, MBLs, databases, virtual laboratories, remote laboratories, and others, and hardware devices involved such as desktop computers, handheld devices, sensors (e.g., motion detectors or thermometers), and others (Table 1). For types of technologies, free computer software, interactive videodiscs, or simple digital instruments such as pH meters were categorized as ‘others’. For hardware devices, ‘others’ referred to equipment such as videodisc players, TV sets, or video cameras.

Next, we coded the length of intervention, levels of students’ involvement in the technology-assisted laboratories, and the roles of technology in supporting components of scientific investigations. The length of intervention depicted the duration of the laboratory instruction in a study. The level of student involvement varied from teacher-centered to student-centered (Millar, Le Maréchal, & Tiberghien, 1999). In the light of the investigation web (Krajcik, Blumenfeld, Marx, Bass, Fredricks, & Soloway, 1998), seven components of scientific investigation were focused on (Table 1).

Finally, to explore what learning outcomes were assessed, we took the three major domains of the revised Bloom’s Taxonomy of educational objectives (Anderson et al., 2001) including cognitive process, attitude, and psychomotor
skills. Also, student learning in laboratories often involves gaining or exhibiting a combination of knowledge and skills such as problem-solving, scientific reasoning, or inquiry abilities. Therefore, we created the fourth domain, integrated skills. Other measured outcomes which did not belong to the aforementioned four domains were categorized as ‘others’ (Table 1). More details of the framework can be found in Chen et al. (2012).

Table 1. Coding scheme of the study

<table>
<thead>
<tr>
<th>Category</th>
<th>Sub-category</th>
<th>Associated codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic information</td>
<td>Demographics</td>
<td>Science domains, subjects, sample sizes</td>
</tr>
<tr>
<td></td>
<td>Characteristics of technology-assisted laboratory--Types of technologies</td>
<td>Simulations, MBLs, databases, virtual laboratories, remote laboratories, and others</td>
</tr>
<tr>
<td></td>
<td>Characteristics of technology-assisted laboratory-- Hardware devices</td>
<td>Desktop computer, handheld devices, sensors, other equipment</td>
</tr>
<tr>
<td>Length of intervention</td>
<td>One shot, 1-2 weeks, 3-8 weeks, one semester, unclear</td>
<td>Teacher demonstration (demonstrated by the teacher; students observe)</td>
</tr>
<tr>
<td>Student involvement</td>
<td>Teacher demonstrates with assistance from students/student demonstrates under teacher’s instruction</td>
<td>Student group work (practical work was carried out by students in small groups)</td>
</tr>
<tr>
<td></td>
<td>Student individual work (practical work was carried out by individual students)</td>
<td></td>
</tr>
<tr>
<td>Roles of technology in supporting components of investigations</td>
<td>Messing about with phenomena</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Finding information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asking and refining questions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Planning and designing an investigation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Carrying out procedures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Making sense of data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sharing findings</td>
<td></td>
</tr>
<tr>
<td>Measured outcomes</td>
<td>Cognitive process, attitude, psychomotor skills, integrated, and others</td>
<td></td>
</tr>
</tbody>
</table>

Identifying papers for this review

Journal articles published from 1990 to 2011 in Science Citation Index and Social Science Citation Index journals were searched for using three sets of keywords: (1) Internet, Internet-based, computer, technology, web, web-based, online, or on-line, (2) laboratory or laboratories, and (3) science, physics, chemistry, biology, or geosciences. This process yielded 312 articles. Each abstract and title was read and screened individually by two panel members. An article was selected to the potential pool when it involved laboratory instruction in a science domain, incorporated technology in assisting learning, and provided empirical evidence. Articles that collected data of user satisfaction or involved exploratory or case studies were included as long as empirical data were reported. Articles in science domains such as physics, chemistry, biology, earth science and environmental science (including urban science), medical science (including nursing education), or which were interdisciplinary were included. Laboratories in engineering education for mastering software such as AUTOCAD, programming, or mechanistic design were excluded (n = 3), because the purpose of such laboratories deviates from our focus on science learning. In this stage, studies with vague descriptions regarding the selection criteria were kept to avoid missing important studies. The above screen process yielded 89 studies, including three that contained two sub-studies.

Using the same criteria, the panel members conducted the second screening by reading the entire articles. Discrepancies between the two panel members regarding the above screening process were discussed and resolved by consulting the opinions of a third panel member. Review articles (n = 12) or ones that did not contain empirical data were further excluded from the pool. Studies that did not involve utilization of technology (n = 12) or which did
not relate to SSLs (n = 19) were also removed from the review pool. These articles were mainly in the fields of gene
technology, computer laboratories, psychology laboratories, etc. Their context deviated from our view of technology-assisted SSLs. Finally, 39 articles (42 studies) entered the final review.

A narrative review rather than meta-analysis was selected as the review method because there were not many empirical studies applying experimental design for each technology except for simulations. Among the studies that used a control group, students’ performances were assessed with various self-developed instruments including types of questions that students asked, score of concept maps or score of students’ laboratory reports. It is difficult to compare effects of learning with a meta-analysis when various self-developed instruments are used. A narrative review seeks to comprehend the diversities and pluralities of understanding around scholarly research topics and is best suited for comprehensive topics (Collins & Fauser, 2005). The topic of our study is therefore appropriate for a narrative review. A narrative review is carried out by critically analyzing literature of a specific theme and drawing results and conclusions in a systematic way from a theoretical and contextual point of view. This is a rather qualitative approach that synthesizes in-depth information for what has been observed among the selected studies, contributed by the researchers’ own experience and the existing theories (van Dinther, Dochy, & Segers, 2011).

Analyzing procedure

Each of the 42 studies was independently analyzed and coded by two members of the panel based on the coding scheme in Table 1. Disagreements between the two examiners were resolved by revisiting and discussing specific segments of the article or by adding opinions of a third member. For studies that contained multiple sub-studies, we treated and coded each as a separate work. Information regarding experimental and control groups of a quasi-experimental design were also coded as different settings. Since one study may involve participants across different levels, address topics across subject areas, incorporate more than one type of technology, or compare effects of different types or designs of technology-assisted laboratories, the frequency counts reported in the result tables may exceed the total number of reviewed articles.

Results and discussion

Demographics and characteristics of the selected studies

We first analyzed subjects, science domains, and use of hardware devices to illustrate the contextual information on types of technology-assisted laboratory activities.

Subjects

Most studies recruited undergraduate students (n = 25) or K-12 students (n = 20). Only one involved graduate students, and four incorporated preservice or in-service teachers. Waight and Abd-El-Khalick’s (2011) study provided an interesting case by including five scientists, a computer programmer, and two science educators in their analysis and, therefore, was categorized as ‘others’. The upper section of Table 2 shows cross-tabulation between types of technologies and subjects. Simulations were applicable to students across K-16; whereas MBLs, virtual laboratories, and remote laboratories have addressed specific participant groups. MBLs were exclusively used at the K-12 level, while virtual laboratories and remote laboratories were mostly implemented with undergraduate students. Application of remote laboratories at the undergraduate level may reflect a growing need to foster collaborations and to share sophisticated instruments among universities (Scanlon, Colwell, Cooper, & di Paolo, 2004).

It should be noted that teacher education is an area that draws less attention in the selected studies. Teachers require both pedagogical content knowledge (PCK) and technological PCK (e.g., structuring students’ interactions with a simulation, Eskrootchi & Oskrochi, 2010) in order to successfully implement laboratories infused with technology. Thus, more attention should be paid to teacher training to develop related pedagogies and increase practical skills in implementing various types of technology-assisted laboratories. More innovative designs for incorporating MBLs and virtual laboratories in teacher education and professional development may significantly enhance the quality of laboratory learning.
Table 2. Types of technologies applied for different subjects, science domains, and hardware devices

<table>
<thead>
<tr>
<th>Demographics and characteristics</th>
<th>Type of technologies</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simulation</td>
<td>MBL</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K-12</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Undergraduate</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Graduate</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Teachers</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Science domains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biology</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Chemistry</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Physics</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Earth science and environmental science</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Medical science</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Interdisciplinary</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hardware devices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desktop computers</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Handheld devices (PDA, handheld computers)</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Sensor</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Other equipment</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Science domains

The majority of the selected studies addressed areas of physics (n = 18), chemistry (n = 15), and biology (n = 10). Relatively few applications were in environmental and earth science or medical science (n = 2). We would like to point out that there are technology-supported curricula in environmental and earth science. These examples include projects built upon geographic information systems (GIS) (e.g., Bodzin, 2008) and curricula that enable students to perform an inquiry in a 3D immersive virtual environment such as “River City” (Ketelhut & Nelson, 2010). These projects allow students to gain experience of scientific practices and yet were not considered as laboratories in the traditional sense. This leads us to reconsider the question of what counts as a laboratory in the areas of earth science, environmental, and medical science, and whether there can be a definition of laboratory commonly shared across science domains (see further discussion of this issue in the Conclusions and a Critical Reflection section).

Table 2 also shows that biology laboratories often utilized simulations, MBLs, and databases, while chemistry laboratories often involved simulations and virtual laboratories. In physics, about half of the selected studies incorporated simulations, followed by MBLs. Overall, it seemed that a particular technology was more often used to serve a specific student group or to fit better with the nature of laboratory learning in some science domains than in others.

In general, the studies provided positive evidence for integrating technology into laboratories. Among the selected studies, a few factors that affect learning in technology-assisted laboratories were identified. Chang, Chen, Lin, and Sung (2008) indicated that students who are better at higher abstract reasoning benefit more from simulation-based physics learning. Also, students’ prior knowledge and attitudes toward learning were found to affect the outcomes of a simulated undergraduate chemistry pre-laboratory (Winberg & Berg, 2007). Nevertheless, relatively few studies took into account participants’ needs and characteristics (e.g., level of reasoning ability, proficiency of scientific process skills or technology skills). The lack of such information makes a systematic analysis impossible. Future studies can seek new applications of technologies in less explored science domains and teacher training. Meanwhile, learners’ needs and characteristics as well as the nature of the content and laboratory learning are important factors that need to be taken into consideration when designing technology-assisted laboratories.
**Types of hardware devices**

Hardware devices could be desktop computers, handheld devices, and sensors (e.g., thermo sensors, or motion detectors as a part of the MBL tool kits) (Table 2). Other equipment included video cameras (Fiore & Ratti, 2007) and pre-made video-clips which were delivered with a videodisc player and/or a TV monitor (Brungardt & Zollman, 1995; Leonard, 1992) or were incorporated with a virtual laboratory (Kozma, 2003; Swan & O'Donnell, 2009).

Desktop computers are the most common hardware device since simulations, virtual laboratories, or databases were installed on desktop computers. MBLs, to some extent, reflect a feature of mobility by using a combination of a sensor (n = 8) and handheld devices (n = 2) or by collecting data with sensors and further processing it on a desktop computer (n = 8). Although mobile phones, tablets, instant response systems, or interactive whiteboards are becoming increasingly popular in classrooms, we found no studies in our review that applied them to science laboratories. However, this might change in the near future since such devices have impacted different aspects of everyday life. Exploring innovative applications of technology and examining its effectiveness on learning in laboratories is a potential area for future research.

**Length of intervention, student involvement, and support for scientific investigation**

Most studies designed a relatively short-term instruction of one shot (n = 5), one to two weeks (n = 14), or three to eight weeks (n = 15). Six studies implemented the laboratories for the entire semester, leaving two which did not specify their length of intervention. In the following section, we examine how the remaining two aspects were incorporated with different types of technologies.

**Student involvement**

Among the four levels of student involvement described by Millar et al. (1999), technologies were more frequently used in settings of students working in small groups (n = 23) or working individually (n = 18) than for teacher demonstration (n = 6) (Table 3). Although teacher demonstration with assistance from students or student demonstration under teachers’ instruction was common in traditional laboratory activities, no technology was specifically utilized in this setting. Among the 19 studies using simulations, nine were carried out by students in small groups, six by individual students, and three by teacher demonstrations. Teacher demonstration often appears as a tutorial or as a preparation section with the purpose of familiarizing students with the program or system and to prepare them to work independently in the following experiments (e.g., Harris, Peck, Colton, Morris, Neto, & Kallio, 2009). The application of MBLs shows a similar trend of being used with small groups (n = 4) or individuals (n = 2). Half of the virtual laboratories were implemented for undergraduates to work individually. Remote laboratories were also often carried out by individual students. In sum, technology-integrated laboratories promote group work, whereas laboratories for undergraduates and in distance education settings are more likely to be designed for individuals.

<table>
<thead>
<tr>
<th>Table 3. Cross tabulation of technologies and student involvement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level of student involvement</strong></td>
</tr>
<tr>
<td>Teacher demo</td>
</tr>
<tr>
<td>Teacher demo with assistance from students/Student demo under teacher’s instruction</td>
</tr>
<tr>
<td>Student group work</td>
</tr>
<tr>
<td>Student individual work</td>
</tr>
<tr>
<td>Unclear</td>
</tr>
</tbody>
</table>

Questions remaining to be solved include how the various types of technologies facilitate or interfere with student-student, student-content, and/or student-technology interactions, and under what social conditions technology-assisted laboratories could have better learning outcomes. For example, Kelly and Crawford (1996) revealed how the
use of MBLs influences learners’ conversations about physics concepts. Manlove, Lazonder, and de Jong (2009) also indicated that, during a simulated inquiry activity, students who worked in a collaborative setting outperformed those who worked individually on their model quality and laboratory report, whereas task duration and specific tool use (e.g., notetaking, hints, and help seeking) did not differ between the two conditions. The above two studies were the few attempts to tackle the abovementioned issues that we observed among the selected studies. We urge that more studies investigate the social and conceptual interactions between learners and technologies.

Support of scientific investigation

Table 4 reveals variations among types of technologies in terms of their roles in assisting different components of scientific investigations. In general, technologies were most frequently used to support investigations associated with messing about with phenomena (n = 29), carrying out procedures (n = 31), and making sense of data (n = 33), whereas sharing findings was the least emphasized component (n = 2). We also found that these technologies seemed to be applied for different functions. For instance, simulation was most likely used to aid students to mess about with phenomena (n = 12) and make sense of data (n = 12). MBLs were often used to assist in carrying out procedures (n = 8) and making sense of data (n = 8). Virtual and remote laboratories were designed to facilitate messing about with phenomena and carrying out procedures. Moreover, simulations, virtual laboratories, and remote laboratories are more applicable to supporting a broader range of scientific investigation processes, whereas MBLs were specifically used to aid students to mess about with phenomena, to carry out procedures, and to interpret findings. For example, Barros, Read, and Verdejo (2008) created a chemistry virtual laboratory composed of virtual laboratory scenarios, tools for remote experimentation, and electronic-interactive note-booking to support university students in all phases of an online inquiry about chemical substance identification.

<table>
<thead>
<tr>
<th>Components of scientific investigations</th>
<th>Type of technologies</th>
<th>Simula</th>
<th>MBL</th>
<th>Data base</th>
<th>Virtual laboratory</th>
<th>Remote laboratory</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messing about with phenomena</td>
<td></td>
<td>12</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>Finding information</td>
<td></td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Asking and refining questions</td>
<td></td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>Planning and designing an investigation</td>
<td></td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>Carrying out procedures</td>
<td></td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>31</td>
</tr>
<tr>
<td>Making sense of data</td>
<td></td>
<td>12</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>33</td>
</tr>
<tr>
<td>Sharing findings</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

The empirical results of the studies reveal distinct features of each technology in assisting laboratory learning. For instance, simulations and virtual laboratories have the advantage of representing a model, system, or process that cannot be easily observed in real life, such as phenomena in slow-motion or the motions of molecules which cannot be seen with the naked eye (Scalise et al., 2011). With this advantage, researchers with the purpose of helping students to gain in-depth qualitative understanding of the observed phenomena would favor simulations and virtual laboratories over other types of technology. Simulation-based laboratories make explicit the underpinning scientific theory, for instance, showing the moving electrons as a visual clue of the current flow when students build an electric circuit. Moreover, simulations could serve at the phase of carrying out procedures by simulating behaviors in which students can manipulate resistors, light bulbs, wires, and batteries as they would in the real world (Finkelstein et al., 2005; Jaakkola & Nurmi, 2008). Simulations can also be a powerful tool in helping students to develop a stronger sense of the data by giving them access to building a model of scientific phenomena and to realizing causal relationships among variables using visualization tools (e.g., Eskrootchi & Oskrochi, 2010). By the same token, MBLs will be an adequate resource if the main purpose of a laboratory activity is to aid students by simplifying procedures of data reading and logging and by processing data and displaying graphs real-time. By doing so, students would have spare time and cognitive resources to engage in conversations related to making meaning of graphs or to constructing alternative explanations (Russell, Lucas, & McRobbie, 2004).

We have observed a few examples starting to incorporate experiences of the simulated and real world phenomena or mixing remote and local resources to enhance conceptual learning and collaboration. Future research may pursue in-
depth understanding of the nature of the technologies and of learning interactions during laboratory activities to design an effective technology-assisted laboratory and to explore innovative applications of the technologies in supporting students’ laboratory learning. Furthermore, we found an inadequate alignment among the purposes and nature of scientific investigation processes and the actual learning outcomes in the reviewed articles. Nor did they address the degree of openness and freedom given to students in their design. These are crucial issues in laboratory learning and should have received more attention (Chen et al., 2012).

Exploring technology affordances for important learning outcomes

In this section, we examine the effects and learning outcomes reported by the selected studies. Five types of learning outcomes were identified: cognitive processes, attitudes, psychomotor skills, integrated skills, and others.

Most studies reported students’ cognitive processes as their only or major outcomes (Table 5). They revealed that technologies enhance students’ cognitive processes, such as manipulating variables (Russell et al., 2004), visualizing data or scientific phenomena (e.g., Brungardt & Zollman, 1995; Jaakkola & Nurmi, 2008), and improving conceptual understandings (Cao & Bengu, 2000; Nakhleh & Krajcik, 1994). Among these studies, about half were supplemented with students’ reports on their attitudes, specifically their satisfaction, interests, preferences regarding the technology used (e.g., Nickerson, Corter, Esche, & Chassapis, 2007), or their experience with the technology-assisted learning environment (e.g., Johnston & McAllister, 2008; Swan & O’Donnell, 2009). It should be noted that the focus was on students’ feedback or perceptions, rather than their attitudes or motivations toward science. Finally, very few studies actually assessed psychomotor skills or integrated skills. The psychomotor skills that did come under investigation in these studies were visualization skills (Harris et al., 2009), graphing skills (Adams & Shrum, 1990) and operational skills (Finkelstein et al., 2005). With regard to integrated skills, a few studies were designed to support students in deploying or gaining a set of joint knowledge and skills to accomplish a complex task, such as acquiring reasoning skills (Barros et al., 2008) or inquiry abilities (Manlove et al., 2009). Rogers and Wild’s (1996) study, which measured changes in the nature of classroom lessons and teaching style, was categorized as “others”.

Table 5. Learning outcomes measured by studies with different types of technologies

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Simulation</th>
<th>MBL Database</th>
<th>Virtual laboratory</th>
<th>Remote laboratory</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive processes</td>
<td>17</td>
<td>7</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Attitudes</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Psychomotor skills</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Integrated skills</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

A consistent conclusion of various studies was that combining simulations or virtual manipulations with physical laboratory activities creates a greater learning effect, such as gaining more conceptual understanding (e.g., Eskrootchi & Oskrochi, 2010; Jaakkola & Nurmi, 2008), posing more theoretical questions (Winberg & Berg, 2007), or making acceptable predictions and explanations (Zacharia & Anderson, 2003; Zacharia, Olympiou, & Papaevripidou, 2008) than learning in the simulation only and the laboratory only conditions. Simulations or virtual laboratories were often used in a pre-laboratory session to equip students with essential concepts and principles; the gained cognitive process in the preparatory session would then enhance learning and performance in the subsequent laboratory activities. However, there were counter findings regarding whether using simulations or virtual laboratories alone may (Finkelstein et al., 2005; Sun, Lin, & Yu, 2008) or may not (Eskrootchi & Oskrochi, 2010; Jaakkola & Nurmi, 2008) lead to significantly better performance than learning in the physical experiment-only condition. In addition to the cognitive outcomes, students reported preferences for the simulated or virtual laboratory learning environment (e.g., Eskrootchi & Oskrochi, 2010; Sun et al., 2008) and perceived the technologies as helpful (Swan & O’Donnell, 2009).

Considering the flexibility of simulations and virtual laboratories in supporting all aspects of scientific investigation discussed in the previous section, it is regrettable to find that only two studies have indeed examined the effects of the technologies on improving inquiry skills (Barros et al., 2008; Manlove et al., 2009). In terms of improvement in psychomotor skills, only two studies suggested that simulated equipment was effective in enhancing students’ skills.
Concerning effects of laboratories employing MBLs, all studies focused on investigating their benefits in terms of the students’ cognitive processes, but none on the other objectives. MBLs improve students’ conceptual understanding and the ability to interpret graphs (Adams & Shrum, 1990) because this technology provides multiple methods of graph display (Russell et al., 2004), affects the focus of students’ observations (Nakhleh & Krajcik, 1993), and saves time for thinking, reflecting, and discussing (Rogers & Wild, 1996).

Remote laboratories and database studies were more diverse in terms of what learning outcomes were actually measured. Among the four studies using remote laboratories, findings have shown that students learned the content information (Barros et al., 2008; Nickerson et al., 2007) and rated the remote laboratories as equally effective as physical laboratories (Nickerson et al., 2007), with additional advantages in flexibility in class delivery, convenience in scheduling, and reliability of setup (Lemckert, 2003). Remote laboratories also show their potential in improving students’ skills for observation and analysis of animal behaviors (Fiore & Ratti, 2007) and obtaining better integrated skills (Barros et al., 2008). Similarly, in the two database studies, Harris et al. (2009) used a variety of measures to provide evidence for the effectiveness of the learning environment. These measures included asking students to conduct investigations using the protein database and simulations, to visualize the protein structure, and to rate their preferences for the environment. However, the learning effect of database studies is inconclusive due to the small number of related studies. The results of the above analysis indicate that the majority of the simulation, virtual laboratory and MBL studies have provided evidence for whether and how these technologies promote students’ cognitive processes and conceptual understanding. In comparison, how remote laboratories and databases can benefit learning outcomes and laboratory performance are less clear. More empirical research on using these technologies is encouraged.

When examining outcomes reported among the studies, most studies centered on the cognitive domain. Some studies looked at students’ attitudes, mostly perceptions and reactions, toward the technology-assisted environment. Very few studies have actually measured psychomotor skills or integrated skills. Objectives such as changes in students’ conversation or reasoning patterns, attitudes toward science, or skills of higher-order thinking (e.g., problem-solving ability), albeit crucial to laboratory learning, have nevertheless been neglected in the field. One possible reason could be the need for valid metrics to assess psychomotor or integrated skills such as visualization skills, graphing skills, operational skills and higher-order reasoning or inquiry skills in technology-assisted SSLs. As pointed out by Johnson, Levine, Smith, and Haywood (2010), new forms of technologies continue to emerge, bringing innovative interventions which aim to improve education and benefit student learning, but appropriate metrics to evaluate the educational value or impact often lag behind. There is a need for more studies on developing assessments focusing on psychomotor or integrated skills and establishing common criteria to evaluate the effectiveness of laboratory activities (Ma & Nickerson, 2006) to help reveal the variety of possible benefits that technology can bring to improving students’ learning in science laboratories. Furthermore, more analyses of studies which match the nature and value of technology with the measured outcomes of the technology-assisted laboratories may help address this gap.

Conclusions and a critical reflection

In the present study, we have reviewed empirical studies from the past two decades that incorporated technology-assisted laboratories. The characteristics of these studies are reported, supplemented with examples to illustrate how these technologies were used to aid laboratory learning at different science domains, subjects, student involvement, and phases of inquiry. Our results reveal that each technology has its distinct features and has different advantages or limitations in terms of supporting student involvement and learning of inquiry. We observed that most researchers employed technologies to support student group work or individual work and emphasized mainly cognitive outcomes. Developing assessments and criteria addressing affective aspects, psychomotor and integrated skills may generate more convincing and promising results beyond cognitive processes.

Articulating underlying interactions among technologies, the to-be-learned content, and the needs of learners may help advance our understanding and synthesize useful guidelines for designing effective SSLs. Based on our analytical framework and findings, we portray the key components, possible factors, and dynamic interactions occurring in the context of a technology-assisted laboratory in Figure 1. Learning and teaching in a technology-assisted laboratory may help address this gap.

315
assisted laboratory is a complex phenomenon that involves features of technology, nature of content and laboratory instruction, as well as the characteristics of learners. The analytical framework allows us to examine this phenomenon from a holistic viewpoint. Resulting from our analyses, we have identified factors and interactions that have failed to be studied and which lack verification. These can be summarized as (1) how to match and/or infuse the features and value of technologies with the nature of content and laboratory learning, (2) how the technology facilitates or interferes with student-student, student-content, and/or student-technology interactions, and (3) how participants’ needs and characteristics are assisted as well as under what conditions technology-assisted laboratories produce better learning outcomes.

One should understand that the components depicted in the above framework are interrelated, and that change in one component or interaction may lead to changes in the other parts of the framework. For this reason, when researchers and curriculum developers seek opportunities to match and integrate technologies with laboratory activities to improve learning, they need to be aware that the incorporation of technologies in SSLs may change students’ experiences of the real-world phenomena and of scientific investigations. Physical and conceptual interactions between learners and the technology-mediated laboratory may differ in vital ways from those in a traditional laboratory. With this notion in mind, we conduct a critical reflection to reinterpret the definition of SSLs in the context of technology-assisted laboratories. We intend to discuss: (1) how the phenomena provided for students to explore in the technology-assisted laboratories vary from the “natural phenomena” demonstrated in a traditional laboratory, (2) how the way students observe or interact with the natural phenomena or materials by using tools or equipment changes, and (3) how their approaches to handling and making sense of data differ from those in traditional laboratories.

Figure 1. Factors and interactions for research on learning in technology-assisted SSLs

What counts as phenomena

The availability of technologies in a SSL affords students the chance to investigate augmented, theorized, or inaccessible phenomena in comparison to the real world phenomena they explore in traditional laboratories. Augmented features such as virtual objects or simulated visual representations (e.g., moving electrons) may be added
to a real world phenomenon to cue or enhance students’ discussion and understanding during a scientific investigation (e.g., Finkelstein et al., 2005; Jaakkola & Nurmi, 2008). Likewise, students are able to simultaneously explore a real world phenomenon with augmented phenomena at the submicroscopic and symbolic levels (e.g., Kozma, 2003). Sometimes, learners may have difficulties in accessing core concepts when building theories or models of a target phenomenon. Engaging students in investigating a simplified version of theorized phenomena (e.g., Zacharia & Anderson, 2003), such as a friction-free situation, may help overcome these difficulties. When students investigate with these technology-mediated artifacts, some features, such as smells, might be unintentionally omitted from the real-world phenomena when they are delivered with remote or virtual laboratories.

Technology-enhanced phenomena may enrich and stretch students’ experiences in SSLs. On balance, it may create new difficulties, arising from authenticity, fidelity, and credibility issues. For authenticity, we must consider whether activities and practices of inquiry embedded in the technology-enhanced environment align with those in the real world and science community (Klopfer, 2008) and how the mis-alignment influences students’ ability of solving real world problems in the long run. In terms of fidelity, we are concerned with whether the technology-mediated representation accurately and clearly portrays experts’ mental models of the to-be-explored phenomena (Roschelle, 1996) and also what kinds of mental models students obtain from the technology-embedded learning experiences. Finally, we view credibility as the degree to which students believe something, such as observations or evidence, based on their experiences with the technology-enhanced phenomena (Tseng & Fogg, 1999). For instance, investigating theorized phenomena may leave students with no experience of dealing with errors in experiments (Chen, 2010). If students lose these senses in science laboratories, how may this influence their views regarding what counts as evidence? Few of the studies we reviewed discussed these new areas of student difficulties. It seems interesting for future studies to investigate students’ notions of technology-mediated phenomena from the authenticity, fidelity and credibility aspects, as well as how their views affect their learning.

Changes in interactions with the natural phenomena or materials

With the use of technologies, students nowadays can engage in scientific experiments without direct interactions with physical equipment or even without physical presence in the laboratory. Technology-mediated laboratories allow students to run repeatable operations (Sun et al., 2008) and accelerate the experiment with limited or no requirements of setting up the equipment by themselves (Zacharia et al., 2008) in comparison to a physical laboratory experiment. Finkelstein et al. (2005) argue that such laboratories may stress discovery rather than verification of science by allowing students to make mistakes without worrying about the consequences, to make changes in the setups and receive immediate feedback, and to repeat the experiments as many times as they desire. As science educators shift away from physical laboratory settings toward various technologies to explore new opportunities, future studies need to investigate what aspects of the physical experiences can or cannot be replaced with virtual or remote experiences, as well as to what extent this replacement or supplementation of physical manipulations influences students’ attitudes toward science and laboratory activities and also their development of psychomotor skills.

Difference in approaches to handling and making sense of data

The initial meaning of “making sense of data” proposed by Krajcik and colleagues (Krajcik et al., 1998) included processes of analyzing, transforming, and visualizing data, creating models, constructing scientific explanations, and also making conclusions. When technologies such as simulations or MBLs take the work of data-handling (e.g., data analysis, transformation, and visualization) from students, they can devote greater mental efforts to the remaining steps. The meaning of “making sense of data” may have to be redefined as model construction and development of explanations and conclusions since students now work with the “processed data” rather than starting from the directly-observed raw data.

In sum, the availability of technologies in a SSL has changed several aspects of learning during scientific investigation. Incorporation of technologies enriches and stretches the features of the to-be-explored phenomena; in the meantime, this change may have altered the way in which students conceptualize phenomena and their development of conceptual understanding. Students’ sensory experiences and incidents about operating instruments or manipulating variables are physically changed when working in a technology-mediated laboratory. Working with
processed data also alters the procedure of data handling and students’ reallocation of time and mental efforts. Jointly, these intended or unintended changes in experiences of the technology-assisted SSL may, in turn, change the nature of learning with and about scientific inquiry. Effects of the aforementioned changes may show in students’ learning outcomes in terms of conceptual understanding, attitudes, psychomotor and integrated skills, or views of nature of science. Based on our findings, we have suggested some possible directions for future research and have pinpointed some issues which have emerged when the incorporation of technologies has indeed changed the nature and process of a science laboratory as well as the nature of learning in it. Since technologies continuously and increasingly change and improve learning and activities in SSLs, we encourage more studies to explore opportunities for innovative integration of technologies with laboratories as well as to deepen our understanding regarding how students, technologies, and laboratory activities interact.

Acknowledgements

The funding of this review paper is supported by the National Science Council, Taiwan, under grant numbers NSC 100-2631-S-011-001, NSC 101-2631-S-011-002.

References


SEEK-AT-WD: A Social-Semantic Infrastructure to Sustain Educational ICT Tool Descriptions in the Web of Data

Adolfo Ruiz-Calleja*, Guillermo Vega-Gorgojo, Juan I. Asensio-Pérez, Eduardo Gómez-Sánchez, Miguel L. Bote-Lorenzo and Carlos Alario-Hoyos

School of Telecommunication Engineering, University of Valladolid, Paseo Belén 15, Valladolid, Spain // adolfo@gsic.uva.es // guiveg@tel.uva.es // juase@tel.uva.es // educog@tel.uva.es // migbot@tel.uva.es // calario@it.uc3m.es

* Corresponding author

(Submitted February 28, 2013; Revised May 14, 2013; Accepted July 3, 2013)

ABSTRACT

There are several Information and Communication Technology (ICT) tool registries that support educators when searching ICT tools for their classrooms. A common problem in these registries is how their data is sustained, since educational descriptions of ICT tools are hard to create and maintain updated. This paper proposes SEEK-AT-WD, an infrastructure that aims at sustaining an educational ICT tool registry in the Web of Data following a social-semantic approach. Its key idea is to take advantage of the data already published in the Web to sustain a collection of ICT tool descriptions, as well as to enable the community of educators to enrich this collection sharing their experience using ICT tools. Following this approach, 6760 descriptions of educational ICT tools have been retrieved from the Web of Data to build an initial dataset. Moreover, the descriptions obtained from the Web are automatically updated without human intervention while more than a hundred tool descriptions have been enriched by educators. Finally, a search system and an annotation tool are presented to illustrate that educational applications can take advantage of SEEK-AT-WD.

Keywords

Educational ICT tools, infrastructure, Web of Data, Linked Data, Web 2.0

Introduction

Information and Communication Technologies (ICTs) are massively employed in learning situations (i.e. scenarios designed for students to learn through the realization of learning tasks (Osuna-Gómez, 1999)). As learning situations may differ in many aspects, there are several characteristics of ICT tools that determine whether they are suitable for each specific learning situation (Kurti, Spikol, Milrad, & Flensburg, 2006). On the one hand, the tool functionality should support the tasks established for the educational situation (Vega-Gorgojo et al., 2008); e.g., a Wiki can be adequate to support students when writing and sharing documents. On the other hand, the tool should be compatible with the rest of the technology employed in the situation, as well as appropriate for the people who is going to use it (Gómez-Sánchez et al., 2009); e.g., a suitable tool for higher education may not fit in primary school. This diversity of purposes and situations makes the selection of ICT tools a critical step when designing learning situations (Gómez-Sánchez et al., 2009; Vega-Gorgojo et al., 2010; Vignollet Ferraris, Martel, & Burgos, 2008).

This technology selection requires educators to be aware of the capabilities of different software tools that can potentially be employed for educational purposes (Vega-Gorgojo et al., 2008) (this paper calls such tools “educational ICT tools”). In order to inform educators several educational organizations maintain ICT tool registries. Some examples are Sisoft (Universidad Complutense de Madrid, 2013) or Ontoolsearch (Vega-Gorgojo et al., 2010). These registries contain data about ICT tools and some of them structure it using semantic technologies, thus providing higher precision and recall in their results when educators search for tools (Vega-Gorgojo et al., 2010). A common issue to these registries is how to create and sustain the data they contain, which is a well-known problem in the educational domain (Bateman, Brooks, & McCalla, 2006). Some of these registries, such as Ontoolsearch, follow a traditional approach where an administrator is responsible for sustaining the dataset. However, this approach requires the organization to assume the whole effort of creating and updating the data. Others adopt the Web 2.0 principles (O’Reilly, 2005) to sustain their respective datasets, involving users in the creation and maintenance of the content. For example, tool providers publish descriptions of their ICT tools in the ROLE Widget Store (Govaerts et al., 2011). Another example is (CoolToolsForSchools, 2013), where educators not only retrieve tool descriptions, but they also publish and update them. Nonetheless, these specific search facilities are still limited because their communities are isolated: the data published in a registry cannot be retrieved from another. In addition, they suffer
from the cold start problem (Maltz & Ehrlich, 1995): at the beginning the utility of the registry is limited since it contains very few tool descriptions; hence, educators are not motivated to enrich the registry since they do not perceive its utility.

In our research, we address this data sustainability problem by taking advantage of the open data already published in the Web of Data (Heath & Bizer, 2011). The Web of Data is a recent proposal that foresees a Web-scale federation of datasets that follow a common methodology for publishing their information: the Linked Data principles (Berners-Lee, 2006). Currently, thousands of data providers, including some educational institutions as (University of Southampton, 2013), publish their information on the Web of Data, which is growing very quickly (Heath & Bizer, 2011, chap. 3). Our key idea is to automatically import ICT tool descriptions from third-party updated repositories of the Web of Data and relate them to a vocabulary understandable by educators; thus, the effort of creating and updating the dataset would be significantly reduced. Our previous work (Ruiz-Calleja, Vega-Gorgojo, Asensio-Pérez, et al., 2012; Ruiz-Calleja, Vega-Gorgojo, Gómez-Sánchez, et al., 2012) shows that retrieving descriptions of educational ICT tools from the Web of Data is feasible, since several thousands of them were obtained, and can well be employed to satisfy educators’ information needs; nonetheless, the educational information that can be inferred from the Web of Data is limited and its support to educational applications can be improved (Ruiz-Calleja, Vega-Gorgojo, Asensio-Pérez, et al., 2012). Our approach to increase the quality of this data is to combine the Web 2.0 and the Linked Data principles in a social-semantic approach (Mikroyannidis, 2007). Thus, an initial dataset of educational ICT tools is generated and periodically updated from the Web of Data, while the community of educators can enrich it by publishing their experience using tools in the classroom. Also note that this dataset will count on a big collection of tool descriptions from the very beginning and thus the cold-start problem can be overcome.

In order to reach this aim, this paper proposes SEEK-AT-WD (Support for Educational External Knowledge About Tools in the Web of Data), a social-semantic infrastructure that collects educational descriptions of ICT tools. SEEK-AT-WD is used to sustain a dataset publicly available in the Web of Data, so its tool descriptions can be freely exported by third parties or directly used to build applications that leverage the benefits of SEEK-AT-WD. By means of an envisioned scenario and two prototype applications, this paper illustrates such benefits and how this infrastructure enables educators to interact with the information about ICT tools available in the Web of Data.

The rest of the paper is structured as follows: first, the state of the art of Linked Data in education is overviewed, with especial attention to the vocabularies and datasets of ICT tools. Next, SEEK-AT-WD is described, including its data model and its software architecture. Then, SEEK-AT-WD is evaluated analyzing the data obtained from the Web of Data and illustrating its consumption by educational applications. Finally, the conclusions of this research work are summed up.

Related work: Educational linked data sources of ICT tools

During the last decade, significant efforts have been made to facilitate the interoperability of educational datasets. Specifically, several educational vocabularies (i.e. Learning Object Metadata (LOM) (Hodgins, 2002)) have been standardized and their use is now widespread in this domain. Additionally, the Semantic Web technologies (Berners-Lee, Hendler, & Laslila, 2001) were exploited as a way to reduce the interoperability effort of the educational data published in the Web (Devedžić, 2006, chap. 1). However, the impact of Linked Data on education is still limited (Dietze et al., 2013) although it has become the de facto standard to publish data on the Web. Current Linked Data efforts in the educational domain mainly focus on the publication of already-existing data on the Web. For example, (University of Southampton, 2013) and (Open University, 2013) are two pioneer projects carried out by educational institutions for the publication of Linked Open Data. However, most of the data currently published by educational institutions focuses on administrative information and not on information that can directly support educators in their classrooms. Other institutions that currently publish Linked Open Data are public libraries, such as the German National Library of Economics (Neubert, 2009) or the Open Library (Open Library, 2012), a social-semantic book registry in the Web of Data that contains over a million books.

Despite these efforts, there are very few proposals that consume Linked Open Data for educational purposes, although some examples can be seen at (Dietze et al., 2013) and (Open Library, 2012). In this respect, (Zablith d’Aquin, Brown, & Green-Hughes, 2011) detects a threefold benefit for educational applications: It reduces the
effort of sustaining data; it decreases the user effort to find relevant information; and it connects educational communities in a global data space. Our proposal looks for educational ICT tool registries to reach these benefits. Thus, current datasets of educational ICT tools available on the Web of Data are presented next.

**Educational ICT tool descriptions in the Web of Data**

To the best of the authors’ knowledge, two educational datasets of ICT tools can be found in the Web of Data: Ontoolsearch and ROLE Widget Store. Ontoolsearch is a specialized search facility that describes about 100 tools using an educational-specific vocabulary defined by Ontoolcole ontology (Vega-Gorgojo et al., 2008). This ontology considers a taxonomy of ICT tools based on the educational tasks they support. Each tool is described by stating its tool type (i.e., “drawing tool”), the tasks it supports (i.e., “collaborative edition of multimedia documents”) and the artifacts it manipulates (i.e., “stores .mpg video documents”). Thus, using Ontoolcole it is possible to formally express complex descriptions, such as “whiteboard that allows the collaborative edition of images, as well as chatting.” Nonetheless, Ontoolcole cannot express the educational experience using ICT tools in the classroom.

ROLE Widget Store (Govaerts et al., 2011) is a recommender system of widgets developed under the ROLE (Responsive Open Learning Environment) Project which trusts in a Web 2.0 approach to sustain its dataset. It recently moved part of its data to the Web of Data using common vocabularies (e.g., DublinCore), as well as the Role Vocabulary (Govaerts et al., 2011). The core of the Role Vocabulary is somehow similar to Ontoolcole since it defines some tool categories and some supported educational activities. However, the Role Vocabulary does not classify tool categories or educational activities in a taxonomy, nor it defines any relationship between different concepts. Additionally, it does not allow to formally describe the experience of using these tools in the classrooms.

There are some other educational registries of ICT tools in the Web of Data, but they only contain a small subset of the tools that are currently employed in the classrooms. Besides, other educational ICT tool descriptions can be found in cross-domain datasets (Heath & Bizer, 2011, chap. 3), although they are not specifically published for educational purposes. Among them, the most popular is DBpedia (Auer et al., 2007), which mirrors part of Wikipedia to the Web of Data. Others are Freebase (Google, 2011), OpenCyc (Cycorp Inc., 2013) or Factforge (Ontotext AD, 2012), which is especially interesting since it collects the data from the previous ones and offers it from a single endpoint. These cross-domain data sources provide descriptions of thousands of ICT tools potentially useful in educational scenarios. They describe tools specifying their genre (e.g., “Word processor”), several administrative data, a textual description and links to other sources where more information about the tool can be found. However, they are not related to an educational vocabulary, so further processing is required to make them useful for applications that manage educational abstractions.

**SEEK-AT-WD**

This section presents SEEK-AT-WD, a social-semantic infrastructure that sustains a registry of educational ICT tools in the Web of Data to overcome the sustainability problems that have been previously detected. First, an envisioned usage scenario that SEEK-AT-WD should support is presented. This scenario will be used to collect the requirements that will guide the design and development of SEEK-AT-WD information model and its software architecture, as this section later presents.

**Envisioned usage scenario**

Marie is a teacher in engineering who is designing a peer-review activity: as homework her students will write a document in pairs and, after that, each pair will review the documents written by three other pairs. Marie does not know which ICT tools could support this activity, so she uses an ICT tool search system to discover tools that allow a group to write text, and that have been used by other to support peer-review activities in blended learning scenarios. Several months before, someone published in Wikipedia a description of Google Docs, stating that it is a collaborative text editor and a file hosting service, and a description of Wikispaces, stating that it is a proprietary wiki software. Furthermore, some educators used an annotation tool to publish in the Web of Data their experiences using these tools, reporting that they used them satisfactorily to support a peer-review activity in blended scenarios.
These two tools can support Marie’s students in the peer-review activity and their descriptions are publicly available on the Web of Data; for this reason, when Marie submits the abovementioned request, she obtains Wikispaces and Google Docs. Marie reads these tool descriptions and she chooses Google Docs to support her activity.

Since Marie’s experience using Google Docs to support her peer-review activity is very positive, she thinks that it may be interesting to use this same tool to support written debates: she publishes some documents and she writes a question in each one, expecting her students to discuss on-line about this question. Once the activity is finished she employs an annotation tool to describe her experience with Google Docs, and share it so that it can help others in selecting the right technology for their educational activities.

This scenario could be achieved by a search facility that allows educators to make queries using educational abstractions (Vega-Gorgojo et al., 2008) while relying on an infrastructure that gathers descriptions of educational ICT tools from different sources of the Web of Data (functionality F1 in Figure 1). As the information retrieved from the Web of Data does not describe tools using educational abstractions, the descriptions obtained should be converted (F2) relating them to an educational vocabulary. In order to allow the data consumption by external applications, the infrastructure should publish its information on the Web of Data with an open license (F3). Further, the infrastructure should also allow educational applications to submit data (F4), thus enabling the community of educators to enrich its dataset out of their educational experience using ICT tools. Finally, as educators are supposed to search and annotate tools through interactive applications, the response time of the infrastructure should not be too high. Figure 1 graphically sums up the data flow of this scenario.

**Figure 1. Educational consumption and publication of information about ICT in the Web of Data**

Note that this approach has several benefits regarding the sustainability of the data. First, the cold-start problem would be solved since the Web of Data already contains information that can be used to support educators when discovering educational ICT tools, as evaluated in a previous research work by the authors (Ruiz-Calleja, Vega-Gorgojo, Asensio-Pérez, et al., 2012). Second, this approach takes advantage of the data already published by the Web community, reusing it for educational purposes. Finally, the educators’ community can share more information obtained from their experience using ICT tools. The infrastructure that supports this scenario is called SEEK-AT-WD. The rest of this section describes its data model and its software architecture.

**SEEK-AT-WD data model**

In order to support the scenario described in the previous subsection, SEEK-AT-WD needs an information model whose objective is twofold: to describe educational ICT tools and to allow the community of educators to review them including educational information of their use. As previously seen, ICT tools can be described by stating their tool type, the tasks they support and the artifacts they manage. In addition, several authors (Jorrín-Abellán, & Stake, 2009; Kurti, Spikol, & Millard, 2008) agree that the educational context where technology is employed has a huge influence on its selection by educators. For this reason, the contexts where tools have already been used are important when educators discover and select tools (Gómez-Sánchez et al., 2009; Kurti, Spikol, & Millard, 2008).
SEEK-AT-WD information model is defined by an ontology called SEEK Ontology, which is designed and implemented following the methodology On-To-Knowledge (Staab, Studer, Schnurr, & Sure, 2001) and the best practices to publish data on the Web (Heath & Bizer, 2011, chap. 3-4). These methodologies and best practices highly recommend reusing vocabularies already employed by third parties since it reduces the development effort and facilitates the federation of datasets. As Ontoolcole is the only ontology that defines a taxonomy of tool types, educational tasks and artifacts, these taxonomies are taken for the description of ICT tools in SEEK Ontology (see (Ruiz-Calleja, Vega-Gorgojo, Alowisheq, Asensio-Pérez, & Tiropanis, 2012) for more details). On the other hand, the Review Vocabulary (Heath & Motta, 2007) is used as a base to describe reviews since it is commonly employed for this purpose in the Web of Data. The Review Vocabulary is a very simple ontology that defines reviews related to any reviewed entity and to the person who published the review. Nonetheless, SEEK Ontology also relates these reviews to the educational contexts where ICT tools were employed, thus collecting the domain-specific characteristics of the use of ICT tools. Figure 2 represents the relationships of the main concepts defined by SEEK Ontology.

SEEK Ontology describes the concepts reused from other vocabularies with the same parameters as they are described in Ontoolcole and Review Vocabulary. However, no ontology that defines the characteristics of an educational context was found. Therefore, several information sources were analyzed, including educational literature (e.g., Kurti, Spikol, & Millard, 2008), other related ontologies (e.g., LOM), other educational applications currently in use (e.g., ROLE Project) and the analysis of educational context descriptions written by educators in real situations. After this analysis, the concepts that were shared among different source were included in SEEK ontology. Table 1 shows and exemplifies the most important concepts defined by SEEK Ontology to describe Google Docs, and Marie’s review and educational context.

### Table 1. Part of the description of a tool, an educational review and an educational context with SEEK Ontology

<table>
<thead>
<tr>
<th>Tool: Google Docs</th>
<th>Review: 0000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Publisher</td>
</tr>
<tr>
<td>“Google Docs is a free office suite…”</td>
<td>Marie</td>
</tr>
<tr>
<td>Description</td>
<td>Text</td>
</tr>
<tr>
<td>Asynchronous text editor</td>
<td>“A very useful tool…”</td>
</tr>
<tr>
<td>Type</td>
<td>Rating</td>
</tr>
<tr>
<td>Writing</td>
<td>4</td>
</tr>
<tr>
<td>Supports task</td>
<td>Has educational context</td>
</tr>
<tr>
<td>Manages artifact</td>
<td>Subject</td>
</tr>
<tr>
<td>Text document</td>
<td>Computers Architecture</td>
</tr>
<tr>
<td>Operating System</td>
<td>Area of knowledge</td>
</tr>
<tr>
<td>Web application</td>
<td>Engineering</td>
</tr>
<tr>
<td>License</td>
<td>Learning goal</td>
</tr>
<tr>
<td>Proprietary software</td>
<td>Requirement analysis</td>
</tr>
<tr>
<td>Developer</td>
<td>Teaching technique</td>
</tr>
<tr>
<td>Google</td>
<td>Peer review</td>
</tr>
<tr>
<td>Has review</td>
<td>Delivery mode</td>
</tr>
<tr>
<td>Review: 0000</td>
<td>Blended</td>
</tr>
<tr>
<td>Educational context: ComAr</td>
<td></td>
</tr>
<tr>
<td>Educational context</td>
<td></td>
</tr>
</tbody>
</table>

SEEK-AT-WD software architecture and prototype implementation

The aim of SEEK-AT-WD software architecture is to sustain a registry of educational ICT tools (F3) retrieving information from the Web of Data (F1) and educational annotation tools (F4). In order to design such infrastructure, the crawling pattern (Hartig & Langegger, 2010) is followed. This pattern is recommended for applications that integrate data from different sources while allowing complex queries with low response time (Heath & Bizer, 2011, chap. 6). It considers a crawler that traverse links in the Web of Data to obtain useful information, which is
automatically cleaned and stored in a data cache. Thus, the tasks of retrieving the data (F1) and aligning it to a common vocabulary (F2) are separated from offering this data (F3). As a disadvantage, the cache may contain stale data, so the Web of Data needs to be periodically crawled.

Figure 3 shows SEEK-AT-WD architecture. It consists of a crawler and a data repository (SEEK-KB) with its corresponding interfaces to add and retrieve data. The crawler plays a central role since it is responsible for gathering ICT tool descriptions from cross-domain linked data sources (F1) and aligning them to SEEK Ontology (F2) using techniques of ontology mapping (Choi, Song & Han, 2006). The current version of the crawler retrieves data from DBpedia and Factforge repositories, as well as any other dataset linked by these two. However, its design is extensible to retrieve information from other linked datasets if needed.

The ontology mapping techniques require relating several DBpedia genre concepts to SEEK Ontology. For example, the DBpedia category “Collaborative real-time editor” is related to the tool categories “Synchronous Text Editor” and “Text Viewer” of SEEK Ontology, as well as the tasks of “Writing” and “Synchronous Communication” and the artifact “Text Document.” This way, the crawler infers the relationships to SEEK Ontology for any tool description classified by DBpedia as “Collaborative real-time editor,” such as Google Docs. For further details see (Ruiz-Calleja, Vega-Gorgojo, Gómez-Sánchez et al., 2012).

The formalization of these mappings requires several iterations and the definition of 114 relationships between DBpedia and SEEK Ontology concepts. However, once the mappings are formalized, SEEK Ontology is related to the ICT tools that are published in DBpedia and those that will be published in the future, either by DBpedia or by any other Linked Data source that uses the same vocabulary. In addition, this approach allows crawling the Web of Data: Once DBpedia mappings are defined, the relationships between SEEK Ontology and other linked vocabularies can be automatically inferred following the links that are published on the Web of Data. In fact, 599 additional tool categories that are also used by DBpedia and Factforge were automatically retrieved from different sources of the Web of Data. These additional categories are enough to formalize the relationships between DBpedia and Factforge, and consequently, between SEEK Ontology and Factforge. It can be seen that this architecture is extensible and can easily import data from other interesting sources – including educational-specific ones - that may appear in the future.

The DBpedia and Factforge mappings, as well as all data extracted by the crawlers, are published at http://www.gsic.uva.es/seek/dataset/. This data has also been stored in SEEK-KB, which can be reached at: http://seek.rkbexplorer.com/. SEEK-KB offers a SPARQL endpoint (World Wide Web Consortium, 2008) and a Linked Data interface, as shown in Figure 4; thus, the data can be browsed and queried in a standard way. Moreover, it provides a data publication interface that is used by the crawler and other applications to enrich or update the data it contains; thus, the community of educators can publish more information about ICT tools following a social approach.
As an example of use, Marie’s information request “tools that allow a group to write text, and that have been used by others to support peer-review activities in blended learning scenarios” can be coded with the SPARQL query shown on Figure 4 (left); “Google Docs” and “Wikispaces” are obtained if this query is launched.

**SEEK-AT-WD evaluation**

Once SEEK-AT-WD infrastructure is developed, it is needed to evaluate whether it provides the functionality required to support the envisioned usage scenario. This evaluation first focuses on SEEK-AT-WD crawler, analyzing the tool descriptions it gathers from the Web of Data. Their quantity, their relationship with SEEK Ontology concepts and their similarity to the descriptions that a human would create are taken into account. Then, SEEK-AT-WD data interfaces are assessed by means of two example applications that illustrate how end-user applications interact with the data stored by SEEK-AT-WD. This analysis will allow discussing -before making extensive use of it with a collection of educational applications- if SEEK-AT-WD supports the envisioned usage scenario overcoming the cold-start problem that other datasets suffer.

**Obtaining educational ICT tool descriptions from the Web of Data**

The first evaluation task analyzes the tool descriptions inferred by SEEK-AT-WD crawler out of the information retrieved from the Web of Data. First, the descriptions obtained are quantified regarding their distribution in relation to SEEK Ontology concepts. After that, the quality of these descriptions is discussed analyzing how an expert on SEEK Ontology (i.e. someone who deeply understands the ontology) would classify 100 of the descriptions inferred by the crawler.

As on February 2013, 6760 tool descriptions were retrieved and almost all the range of tool types was covered. Specifically, SEEK Ontology defines 46 tool types, 11 of which obtained more than 350 results; 12 categories obtained between 130 and 350 tool descriptions, while another 12 categories obtained between 27 and 130 tool descriptions. Finally, 11 categories were related to less than 27 tool descriptions. It is also important to mention that three concepts of SEEK Ontology, such as Questionnaire Manager, could not be related to any concept of an external linked dataset vocabulary, since no synonyms were found in other vocabularies used in the Web of Data; hence, no tool descriptions were imported for these concepts. Table 2 shows some of the tool types defined by SEEK Ontology and the number of tool descriptions extracted by the crawlers related to them.
Table 2. Number of descriptions retrieved from the Web related to some categories of SEEK Ontology

<table>
<thead>
<tr>
<th>Tool type</th>
<th>#tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representation tool</td>
<td>3565</td>
</tr>
<tr>
<td>Audio player</td>
<td>181</td>
</tr>
<tr>
<td>Video player</td>
<td>154</td>
</tr>
<tr>
<td>Group tool</td>
<td>1956</td>
</tr>
<tr>
<td>Chat</td>
<td>128</td>
</tr>
<tr>
<td>Audio conference tool</td>
<td>27</td>
</tr>
<tr>
<td>Information management tool</td>
<td>1904</td>
</tr>
<tr>
<td>Blog</td>
<td>229</td>
</tr>
<tr>
<td>Document repository</td>
<td>776</td>
</tr>
<tr>
<td>Processing tool</td>
<td>1268</td>
</tr>
<tr>
<td>Compiler</td>
<td>1005</td>
</tr>
<tr>
<td>Simulator</td>
<td>237</td>
</tr>
<tr>
<td>Construction tool</td>
<td>1479</td>
</tr>
<tr>
<td>Text editor</td>
<td>473</td>
</tr>
<tr>
<td>Wiki server</td>
<td>314</td>
</tr>
<tr>
<td>Electronic calendar</td>
<td>2</td>
</tr>
<tr>
<td>Concept map tool</td>
<td>54</td>
</tr>
<tr>
<td>Slide composer</td>
<td>44</td>
</tr>
</tbody>
</table>

As a general rule, the more specific a category is, the less amount of tool descriptions are obtained from the Web of Data; for example, the number of concept map tools retrieved is much lower than the number of document repositories. It is also important to note that not all of the data imported refers to educational ICT tools. For example, out of the 314 descriptions that the crawler relates to Wiki server there are many that do not refer to ICT tools, but to other concepts related to Wiki servers. Despite this noisy data, which is a well-known problem when reusing data from the Web (Heath & Bizer, 2011, chap. 6), our previous work showed that these descriptions can well be employed to support educators when discovering ICT tools (Ruiz-Calleja, Vega-Gorgojo, Asensio-Pérez, et al., 2012). Additionally, this noise can be expected to reduce over time since the community of educators can clean SEEK-AT-WD dataset.

Regarding the quality of the data, 100 descriptions were randomly selected and their classification was completed by an expert on SEEK Ontology. Thus, the descriptions automatically inferred by SEEK-AT-WD are compared to the ones that a human would publish. Table 3 quantifies how these 100 descriptions were related to the three taxonomies defined by SEEK Ontology: tool types, educational tasks supported and artifacts managed.

Table 3. Classification of 100 tool descriptions retrieved from the Web of Data by SEEK-AT-WD

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>#descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact categorization</td>
<td>66</td>
</tr>
<tr>
<td>Lack of some tool types</td>
<td>23</td>
</tr>
<tr>
<td>Lack of some tasks supported</td>
<td>26</td>
</tr>
<tr>
<td>Lack of some artifacts managed</td>
<td>14</td>
</tr>
<tr>
<td>Erroneous data</td>
<td>0</td>
</tr>
</tbody>
</table>

66 out of 100 descriptions retrieved by the crawlers are classified in the same categories as the expert would. Typically, these are descriptions of tools that only support the generic tasks defined by the mappings (e.g. Google Mail, which allows to communicate and to send e-mail messages). On the other hand, the crawler could not infer all the tasks supported by 26 tools. An example is Windows Live Messenger, which is described as a chat client, but the crawlers did not infer some functionalities that were described by the expert, such as video communication or collaborative drawing because this tool is not classified by DBpedia as a videoconference tool nor as an image editor. Further, 23 descriptions include all the educational tasks that the tools support, although some information about the artifacts they manage is missing. An example is FreeMind, which is properly classified as a concept map tool that manages models, but the crawler could not infer that it also manages images. Finally, it is very positive that the crawlers did not infer any erroneous data in the 100 tool descriptions analyzed. Again, these results are very satisfactory, but the tool descriptions retrieved from the Web of Data can be improved by the community of educators if a social annotation tool is available.

Current applications that interact with SEEK-AT-WD data

The data gathered by SEEK-AT-WD can be retrieved and enriched through its data access and data publication interfaces. However, as previously discussed, educators cannot be expected to manipulate these interfaces. Instead, end-user applications are required to support educators the interaction with the data contained in SEEK-AT-WD. In order to evaluate the feasibility of such applications, U-Seek and We-Share are presented as two example applications that allow educators to retrieve and publish information about educational ICT tools in SEEK-AT-WD.

U-Seek (Vega-Gorgojo, Ruiz-Calleja, Asensio-Pérez, & Jorrín-Abellán, 2012) is a web application that provides a graphical user interface to obtain information from SEEK-KB. It supports the creation of queries adding restrictions...
about the functionality of the tools obtained and about the contexts where they have been used. When an educator launches a query, U-Seek creates a SPARQL query accordingly, submits it to SEEK-AT-WD and presents the results obtained. As an example, Figure 5 shows the formulation of the query that Marie would submit in the envisioned usage scenario. U-Seek is now available at: http://www.gsic.uva.es/seek/socialuseek/

U-Seek is an interactive application that is directly used by educators. For this reason, the response time of SEEK-AT-WD to submit the results of a query should be lower than a few seconds. In this regard, 1637 queries were launched with U-Seek before February 2013; the arithmetic mean of SEEK-KB response time when answering these queries was 608 milliseconds, while its standard deviation is 518 milliseconds. It can be seen that SEEK-KB response time is low enough as to satisfy the requirements of interactive applications that take advantage of its data. In fact, none of its users complained about it when being asked.

We-Share is a social annotation application. It is a web application that is currently employed by a community of educators to submit data to SEEK-AT-WD. It provides an interface based on forms that can be easily manipulated by educators to publish descriptions of ICT tool, educational contexts or educational reviews. When an educator adds or modifies data using We-Share, this application automatically relates that data to SEEK Ontology and submits it to SEEK-AT-WD. We-Share is available at http://seek.cloud.gsic.tel.uva.es/weshare/.

Between January and February 2013, We-Share users published 27 educational contexts and 94 educational reviews; moreover, they published or updated 116 ICT tool descriptions. All the data created by We-Share is reachable at http://seek.cloud.gsic.tel.uva.es/upload/ and is currently published by SEEK-AT-WD (and hence queried by U-Seek) in conjunction with the data retrieved from the Web of Data. For example, the Google Docs description presented in Figure 4 (http://seek.rkbexplorer.com/id/tool/Google_Docs) contains data obtained by SEEK-AT-WD crawler, as well as data created with We-Share. It can be seen that all the information provided is coherently combined.

Discussion

Evaluation results show that a big collection of ICT tool descriptions can be automatically obtained from the Web of Data and related to SEEK Ontology. Thus, a registry of educational ICT tools can be automatically created and kept updated, since the data sources where information is retrieved are continuously updated (Auer et al., 2007). As the
tool descriptions gathered from the Web of Data cover almost the whole range of concepts defined by SEEK Ontology, they can well be used to overcome the cold start problem that social registries suffer. Further, as tool descriptions are semantically structured, the advantages of semantic searches for this domain (Vega-Gorgojo et al., 2010) – such as more accurate results - are still present.

It was also shown that the tool descriptions inferred by SEEK-AT-WD crawler are classified in a similar way as a human expert would do, although a third of the descriptions analyzed lack of some relationships to SEEK Ontology concepts. In this regard, the social facet of SEEK-AT-WD is remarkable because it enables the educational community to complete these descriptions, to add new ones and to share their experience about the use of ICT tools in specific educational contexts. As these contexts are formalized with an ontology, semantic searches can use them to filter results. This way, Marie not only can ask for tools that allow a group to write, but also she can restrict the results to those that were previously employed in blended scenarios to support peer-review activities. Thus, she obtains less, but more relevant results.

Finally, by mean of two example applications it was shown that SEEK-AT-WD interfaces allow interactive end-user applications to consume and enrich the data it contains. Hence, the envisioned usage scenario is supported by SEEK-AT-WD: educational queries can be submitted to obtain descriptions of ICT tools that someone published in the Web of Data in conjunction with educational reviews of these same tools. Moreover, this same infrastructure allows educators to publish in the Web of Data information about their experience using ICT tools in the classroom. Finally, only a few tens of educational reviews have already been collected by We-Share yet, but currently there is a community of educators using it, so we expect more reviews to be published in the near future. Therefore, SEEK-AT-WD utility will increase since more educational-specific knowledge will be available.

Conclusions and future work

This paper presented SEEK-AT-WD, a social-semantic infrastructure that sustains an educational ICT tool registry in the Web of Data. SEEK-AT-WD currently retrieves the educational ICT tool descriptions linked by DBpedia and Factforge, but it can be easily extended to obtain information from other datasets of the Web of Data. In February 2013 SEEK-AT-WD gathered 6760 updated descriptions of ICT tools potentially useful for educational purposes, covering almost all the range of concepts defined by SEEK Ontology. Moreover, as the Web of Data is periodically crawled, these tool descriptions are automatically updated and new ones can be discovered. Further, the quality of the descriptions obtained is satisfactory since two thirds of the tools were classified as an expert would, and the others did not contain erroneous data, although some of the tool characteristics were missing. This information retrieved from the Web of Data can well be used to overcome the cold-start problem that social registries suffer while it can be further enriched by the educational community, exploiting the social facet of SEEK-AT-WD.

Two educational applications illustrate the feasibility of allowing educators to interact with SEEK-AT-WD. They enable the educational community to publish and consume information about the use of ICT tools in the classroom in an interactive way. Moreover, as this information is published as Linked Open Data, third parties can make use of it to develop their own applications. Thus, the potential impact of SEEK-AT-WD data is much higher than other isolated data sources since other educational communities can take advantage of it.

Near future work will focus on an evaluation of the data created by We-Share. Specifically, it will be very interesting to discuss how the data created by a community of educators can complement the information about ICT tools automatically obtained from non-educational data sources of the Web of Data. This same evaluation will also be used to gather information about the use of SEEK Ontology by its end-users. Then, the efforts will focus on giving a better support to the educators through the development of a collection of applications that take advantage of SEEK-AT-WD data. For example, semantic technologies can be exploited to give multilingual support or to facilitate the evolution of SEEK Ontology through the use of folksonomies; on the other hand, social information can be further exploited to recommend educators combinations of tools that can support a given context.
Acknowledgments

This work has been partially funded by the Spanish Ministry of Economy and Competitiveness (Project TIN2011-28308-C03-02), by the Government of Castilla y Leon, Spain (Project VA293A11-2) and the postdoctoral fellowship Alianza 4 Universidades.

References


Heath, T., & Bizer, C. (2011). Linked data: Evolving the web into a Global Data Space. doi:10.2200/S00334ED1V01Y201102WBE001


Automatic Generation and Ranking of Questions for Critical Review

Ming Liu\textsuperscript{1,2*}, Rafael A. Calvo\textsuperscript{2} and Vasile Rus\textsuperscript{3}

\textsuperscript{1}Faculty of Computer and Information Science, Southwest University, Chongqing, China // \textsuperscript{2}School of Electrical and Information Engineering, University of Sydney, Australia // \textsuperscript{3}Department of Computer Science, University of Memphis, USA // liuming9902@gmail.com // Rafael.Calvo@sydney.edu.au // vrus@memphis.edu

\textsuperscript{*}Corresponding author

(Submitted April 8, 2013; Revised December 2, 2013; Accepted January 8, 2014)

ABSTRACT

Critical review skill is one important aspect of academic writing. Generic trigger questions have been widely used to support this activity. When students have a concrete topic in mind, trigger questions are less effective if they are too general. This article presents a learning-to-rank based system which automatically generates specific trigger questions from citations for critical review support. The performance of the proposed question ranking models was evaluated and the quality of generated questions is reported. Experimental results showed an accuracy of 75.8\% on the top 25\% ranked questions. These top ranked questions are as useful for self-reflection as questions generated by human tutors and supervisors. A qualitative analysis was also conducted using an information seeking question taxonomy in order to further analyze the questions generated by humans. The analysis revealed that explanation and association questions are the most frequent question types and that the explanation questions are considered the most valuable by student writers.

Keywords

Automatic question generation, Writing support, Natural language processing, Learning to rank

Introduction

Academic Writing is one of the most challenging academic activities for most university students. Typical writing skills are not sufficient for academic writing which involves intense argumentation for supporting one’s work addressing a major research question. Students must identify relevant information, critically read and analyze research literature, and finally synthesize previous and related work in order to build the case for their answer the research question (Graswell, 2008). Thus, critically evaluating the literature and understanding its key concepts is an important aspect of academic writing. However, many students experience difficulties with critically analyzing the literature. Afolabi (1992), for example, identified some of the most common problems that students have when writing a literature review, including not being sufficiently critical, lacking synthesis, and not discriminating between relevant and irrelevant materials.

Trigger questions can be an effective tool to help with writing a critical review. A limited number of generic trigger questions are commonly used to guide students during academic writing (Taylor & Procter, 2008). For example, a research supervisor usually asks trigger questions such as: (1) Have you critically analyzed the literature you use? Instead of just listing and summarizing items, have you assessed them, discussing strengths and weaknesses? (2) Have you discussed how your project will contribute to that discipline or field? However, such questions are too general and not likely to provide meaningful feedback for many students writing about a specific topic. This paper focuses on investigating the performance of an automated system to generate specific trigger questions that offer contextualized feedback relevant to the target topic.

Automated Question Generation (AQG) is a challenging task which involves natural language understanding and generation (Rus & Graesser, 2009). It typically involves three major steps: Target content selection (what to ask about), question type selection (the type of question, e.g., Who, Why, Yes/No), and question construction (how to ask question). Current AQG approaches (Mostow & Chen, 2009) focus on generating mostly factual questions for supporting reading comprehension or vocabulary assessment. In contrast, the present work focuses on generating trigger questions for supporting critical thinking in academic writing.

In order to generate such trigger questions, two main AQG approaches were proposed: (1) question generation based on key phrases (M. Liu, Calvo, Aditomo, & Pizzato, 2012), and (2) question generation based on citations (M. Liu, Calvo, & Rus, 2012). The focus of this paper is on improving the system performance of the second approach. In this
approach, sentences containing citations are chosen as target sources for question generation because such sentences are more relevant to the task of writing research papers. They normally express an author’s Opinion or the description of an Application. For example, the following citation sentence expresses an Opinion of the cited author Cannon: Cannon (1927) challenged this view mentioning that physiological changes were not sufficient to discriminate emotions. Some examples of trigger questions that could be generated from the above citation sentence are: Why did Cannon challenge the view mentioning that physiological changes were not sufficient to discriminate emotions? (What evidence is provided by Cannon to prove the opinion?) Does any other scholar agree or disagree with Cannon?

Trigger questions are important during writing as they draw student writers’ attention and reflection to other researchers’ views and evidences. In our previous approach (Ming Liu, Calvo, & Rus, 2010; M. Liu, et al., 2012), rule-based methods were used for all major steps: Citation sentence extraction, sentence simplification, classification of simplified sentences, and syntactic transformation. However, in this processing pipeline errors could occur at any stage, such as citation sentence extraction, voice transformation, and classification. To overcome these challenges, an overgeneration-and-ranking approach was adopted. Compared with the previous study (M. Liu, et al., 2012), the present study describes the improved question generation system based on the overgeneration-and-ranking approach and its evaluation in an educational context. Specifically, our contributions are as in the followings:

- A novel AQG system is proposed for academic writing support based on an overgeneration-and-ranking approach. Two different ranking algorithms were evaluated: RankSVM (Pairwise Approach) and Logistic Regression (Pointwise Approach).
- An evaluation method for the AQG system is described: the performance of a question ranking model and the quality of system-generated trigger questions.
- A deep analysis of human supervisor and tutor trigger questions based on a question classification proposed by Trabasso and Magliano (1996).

The remainder of the article is organized as follows. Background section provides a brief literature review related to question generation and ranking. The System Design and Architecture section describes the proposed AQG framework and the ranking model. The Evaluation and Result section describes the experiments and results.

Background

This section first reviews the literature of how the questions are used for critical review. Then, it presents current AQG and the learning-to-ranking approaches.

Trigger questions and critical review

In higher education, students need to be able to critically review literature related to their research topics. According to Steward (2004), a critical review should be “Comprehensive, Fully referenced, Relevant, A synthesis of key themes and ideas, Balanced between different ideas and opinion, Critical in its appraisal of the literature, and Analytically developing new ideas from the evidence” (p. 496). However, critical review is not an easy task for most college students.

Questions are central aspects in the theories of learning, cognition, and education (Graesser & Person, 1994). It can engage students in learning activities, such as critical review, by helping them to recognize their knowledge deficiencies and reflect on their writing activity. But, many studies have shown that students have problems recognizing their own knowledge deficits (Hacker, Dunlosky, & Graesser, 1998) and ask very few questions (Graesser & Person, 1994). Thus, support in the form of generic trigger questions (asked by a computer or human tutor or supervisor) are often used to help self-reflection. Reynolds and Bonk (1996) showed that a group of students given generic trigger questions performed better than those students who received no trigger questions to support revision in a writing activity. As mentioned in Introduction, the aim of the AQG system is to support students’ writing activities by generating trigger questions that help them self-reflect to important aspects of the critical review writing task.
Question generation

Recently proposed question generation approaches (Heilman & Smith, 2010; Yao, 2010) focused on generating factual questions for reading compression and vocabulary assessment. These approaches are more related to our work since the question generation objectives are similar: Specific shallow/deep questions generation from natural text.

Yao (2010) classified question generation systems into three categories: Template-based, Syntax-based, and Semantic-based. The Syntax-based approach is the most popular approach to automatically generating factual questions. The key idea of this approach is to transform the declarative target sentence into an interrogative one by manipulating the derived syntactic tree typically parsed by using a Context Free Grammar parser. Tregex (Levy & Andrew, 2006) is a powerful syntactic tree search language for identifying syntactic elements (e.g., main verbs of sentences) and which has been used by researchers (Heilman & Smith, 2010) to define wh-movement rules. Compared to the syntax-based approach, the semantic approaches (Yao, 2010) rely on a semantic parse—a deeper level of linguistic analysis than the syntactic parse—aimed at transforming the declarative semantic parse into questions. Both syntax-based and semantic-based approaches can generate questions having high specificity because these questions contain more information about the answer. But, the generated questions are factual questions and generally not very deep.

Unlike syntax-based and semantic-based approaches, the template-based approach (Mostow & Chen, 2009) does not require complex question transformation rules that convert the parser output into questions. Instead, it focuses on generating deep questions by extracting knowledge from the text, make inferences when possible, and filling empty slots in question templates. This approach can generate deep questions. But, these questions are less specific.

In order to generate specific deeper questions, which are generated from original sentences with predefined deep question templates, a combined syntax-based and template-based question generation approach, as the one described in this article, is recommended.

Learning-to-rank

Learning-to-rank received increasing attention in both Information Retrieval and Machine Learning research during the past decade. Most of approaches to learning-to-rank are designed as supervised machine learning approaches. All instances are given a (binary or ordinal) score or label indicating their relevance as decided by an independent judgment process. In the training phase, a ranking function is learned based on a set of feature vectors together with their true labels. In the testing phase, the ranking function is used to rank a new set of instances and generate a ranked order of these instances.

Based on how they treat sets of ratings and loss functions, Cao, Qin, Liu, Tsai, and Li (2007) classified learning-to-rank approaches into 3 categories: (1) Pointwise Approach: Learning to classify the question or instance according to their label individually (e.g., positive or negative category), (2) Pairwise Approach: Classifying pairs of rated questions into two categories (correctly ranked or incorrectly ranked), and (3) Listwise Approach: Optimizing the loss function for ordering the questions related to a single academic paper instance. In the information retrieval literature, the Pointwise approach is viewed as the weakest of the three learning-to-rank approaches because it ignores the cluster of answer instances per query. Machine learning techniques that can be used in conjunction with the Pointwise approach are classifiers (e.g., Naïve Bayes) and regression (e.g., Logistic Regression). The Pairwise approaches are considered more effective than Pointwise approaches because pairs of answer instances are considered. The algorithms used in the Pairwise approaches include RankSVM (Joachims, 2006). Listwise approaches are more recent developments. Liu (2009) shows that the Listwise techniques reached scores similar to or better than Pairwise techniques.

However, Listwise approaches are less suitable than Pointwise and Pairwise approaches in computational linguistics because the labels used in the training set normally have very few categories, such as relevancy or irrelevancy. These few categories would cause Listwise approaches to have problems of accurately obtaining an ordered ranking sequence in the training set. Collins and Koo (2005) trained a logistic regression model, a Pointwise approach for ranking syntactic parses. Duh (2008) proposed an automatic machine translation evaluation based on learning-to-
rank. He compared RankSVM to RankBoost and found that RankSVM performed best when ranking-specific features are considered.

System design and architecture

This section presents an overview of the system's pipeline architecture (see Figure 1), briefly describing each step and emphasizing the question ranker.

Stage 1: Pre-processing

In stage 1, citation sentences from academic papers are extracted, simplified, and then transformed in active voice. The citation sentence extraction was based on the citation styles defined by Powley and Dale (2007). After the citations are extracted, sentence simplification is performed. This involves splitting compound and complex sentences and also removing phrase types such as appositives, non-restrictive relative clauses, and participial modifiers. Tregex (Levy & Andrew, 2006) were used for this sentence simplification and passive-to-active voice transformation.

Passive-to-active voice transformation is frequently applied since Hyland (1994) identified that the passive construction is one of the common grammatical styles used in citation sentences. The transformation algorithm is to set the new sentence in active form, which is needed for next steps in our processing pipeline, according to the passive construction type. Table 1 shows three common passive construction types used in academic writing (Hyland, 1994). In the examples shown in the table, the parenthesis indicates a syntactic element given by the term in the bracket. The Tregex Expression rules are used to detect the passive construction type and extract these syntactic elements. Finally, we reconstruct the passive sentences into their active form by using these elements.

<table>
<thead>
<tr>
<th>Passive construction</th>
<th>Tregex rules</th>
<th>Active form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1: It+is+[main verb]+ [that..]. e.g., It is {argued [Main Verb]} {that… [Clause]}</td>
<td>SBAR=Clause $· (VBN &lt; &quot;+Main Verb+&quot;)</td>
<td>[Subject]+[Main Verb]+ [That Clause] e.g., Peter argued that…</td>
</tr>
<tr>
<td>Type 2: The main verb followed by a particle.</td>
<td>PRT=Particle $· (VBN &lt; &quot; +Main Verb + &quot;)</td>
<td>[Subject]+[Main verb]+[Particle] + [Object]</td>
</tr>
</tbody>
</table>

Figure 1. System architecture: Multiple stages question generation process
Stage 2: Citation classification

The goal of this stage is to identify the citation category for each citation sentence extracted in Stage 1. The citation categories shown in the first column in Table 2 were used: ‘Opinion’, ‘Result’, ‘Aim of Study’, ‘System’, ‘Method’, and ‘Application’. These were chosen based on a taxonomy of conceptual citation categories proposed by Lehnert et al. (1990). To automatically detect the citation category, a Naive Bayes classifier is trained, and 17 generic features were proposed, including Cue Phrases, Sentiment Features, Negation Feature, and Syntactic Feature (M. Liu, et al., 2012).

Table 2. Citation types and examples

<table>
<thead>
<tr>
<th>Citation Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aim</td>
<td>present the aim of an author’s study, e.g., Bunescu et al. focused on extracting named entities from natural language documents.</td>
</tr>
<tr>
<td>Opinion</td>
<td>express the opinion of an author, e.g., Reiter and Dale (1997) state that template-based systems are more difficult to maintain and update.</td>
</tr>
<tr>
<td>Result</td>
<td>report the result of an author’s study, e.g., McCallum and Nigam (1998) show that the multivariate Bernoulli model performs well with small vocabularies.</td>
</tr>
<tr>
<td>Method</td>
<td>describe a method, algorithm, technique, model, or framework proposed by an author, e.g., Bi-Normal Separation is a relatively new feature selection method introduced by Forman (2003).</td>
</tr>
<tr>
<td>System</td>
<td>describe a system, e.g., AutoSlog (Riloff, 1996) is a dictionary construction system that creates extraction patterns automatically using heuristic rules.</td>
</tr>
<tr>
<td>Application</td>
<td>apply a method/system to a field, e.g., Kappa statistics K will be used to measure the reliability (Siegel and Castellan, 1988).</td>
</tr>
</tbody>
</table>

Stage 3: Question generation

This is the third stage in our approach to automatically generating trigger questions. Once the citation category and syntactic features were extracted from a citation sentence, a set of predefined patterns are used to generate the corresponding questions. Table 3 shows six rules and their questions templates defined in our Repository of Templates. For example, the following citation sentence is assumed to be extracted in Stage 1:

While it is shown that AES has inter-assessor correlations comparable to that of human assessors (Dikli, 2006), many scholars are still highly critical of the validity and robustness of the approach (Britt et al., 2004).

The sentence simplifier will split the complex sentence into two simpler sentences: It is shown that AES has inter-assessor correlations comparable to that of human assessors (Dikli, 2006), and Britt (2007) criticized the validity and robustness of the approach. Because the first sentence is passive, it is transformed into an active sentence: Dikli showed that AES has inter-assessor correlations comparable to that of human assessors.

In Stage 2, the classifier categorizes the first sentence as a ‘Result’ citation and the second sentence as belonging to the “Opinion” category. Stage 3 applies rule 1 to generate the following trigger questions, which are intended to prompt the student’s reflection by asking for evidence or other authors’ opinions:

Why did Britt criticize the validity...? Does any other scholar agree or disagree with Britt? Similarly, rule 3 is applied to generate the trigger questions below for evaluating the evidence. Did Dikli objectively show that AES has...? Is the
analysis of the data accurate and relevant to the research question? Each of these questions is then scored according to features of the source sentence. The following section will describe the scoring process in more detail.

### Table 3. Six rules and their question templates

<table>
<thead>
<tr>
<th>Rule</th>
<th>Category</th>
<th>Question Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Opinion</td>
<td>Why +subject_auxiliary_inversion ()? What evidence is provided by +subject+ to prove the opinion? Does any other scholars agree or disagree with +subject+?</td>
</tr>
<tr>
<td>2</td>
<td>Aim</td>
<td>Why does +subject+ conduct this study to +predicate+? What is the research question formulated by +subject+? What is +subject+s contribution to our understanding of the problem?</td>
</tr>
<tr>
<td>3</td>
<td>Result</td>
<td>Subject_auxiliary_inversion ()? Is the analysis of the data accurate and relevant to the research question? How does it relate to your research question?</td>
</tr>
<tr>
<td>4</td>
<td>Method</td>
<td>In the study of +subject+, why +subject_auxiliary_inversion ()? Which dataset does +subject+ use for this experiment? What are the strengths and limitations of this approach?</td>
</tr>
<tr>
<td>5</td>
<td>System</td>
<td>In the study of +subject+, why +subject_auxiliary_inversion ()? What are the strength and limitations of the system? Does it relate to your research question?</td>
</tr>
<tr>
<td>6</td>
<td>Application</td>
<td>Why+Subject_Verb_Inversion()? Could the problem have been approached more effectively from another perspective? Does it relate to your research question?</td>
</tr>
</tbody>
</table>

### Question ranker

The previous stages generate questions that vary in their quality, from syntactic to semantics to importance. This is unavoidable and happens for different reasons, such as errors in sentence parsing, voice transformation, and citation classification. To address this problem, ranking the large pool of questions according to their quality is needed. Stage 4 implements a learning-to-rank algorithm to meet this challenge.

### Ranking model

In the ranking model, two common learning-to-rank approaches were used: Pointwise approach (Logistic Regression; (Collins & Koo, 2005) and Pairwise approach (RankSVM; (Joachims, 2006) ). The logistic regression model is learned by fitting training data to a logit function by using the predictor binary variable which indicates whether a question is acceptable or not. The RankSVM model is learned using a Pairwise approach which can naturally specify questions that are of an equivalent rank. The Support Vector Machines (SVM) algorithm has been used previously for preference ranking in the context of Information Retrieval. The same framework was adopted for ranking questions in this case. In this model, given a collection of questions ranked according to preferences between two questions represented by feature vector q, and q, respectively, and a linear learning function f,  

\[ q_i > q_j \Leftrightarrow f(q_i) > f(q_j) \]  

(1)

Where \( \succ \) indicates that q, is preferred over q,. The function f is defined as f(q)=w*q, where  

\[ f(q_i) > f(q_j) \Leftrightarrow w * q_i > w * q_j \]  

(2)

In the context of SVM, these weight vectors or support vectors (w) are identified by minimizing the function using slack variables  \( \xi_{ij} \) :  

\[ \min_{w,\xi_{ij}} \frac{1}{2} \|w\|^2 + C \sum_{ij} \xi_{ij} \]  

(3)

Subject to the constraints:  

\[ \forall (q_i, q_j): w * q_i > w * q_j + 1 - \xi_{ij} \]  

\[ \forall (i, j): \xi_{ij} \geq 0 \]
Finding the support vectors and the generalization of the Ranking SVM is done differently (Joachims, 2006). If the data are linearly separable, the $\xi_y$ are all equal to 0. In this case, the ranking function is considered as projecting the data points onto the separating hyperplane and the support vectors as the two points $q_i$ and $q_j$ whose projections are nearest each other on the hyperplane. The generation is accomplished by calculating $w$ to maximize the distance $\frac{w(q_i - q_j)}{||w||}$. Like the classification SVM algorithm, the margin is maximized when $||w||$ is minimized.

Feature Definition

The features used in the ranking models were developed by an in-depth analysis of questions generated manually by human experts for the training set which contained 504 citations. This is the same set used for training the citation classifier. Because our current ranking models focus on acceptability of a question in terms of grammatical and semantic correctness, these features should indicate the likelihood of generating an acceptable or unacceptable question in terms of the complexity of source sentences (Num.of NamedEntities, NameAppearInBoth, Num.ofClauses and Length), the transformation performed during the processing (IsPronounResolved, PredicationConfidence and IsPassiveVoice), and a citation sentence in reporting form (ReportingVerb and NameAppearInSubject). For example, if the source sentence is very long and complex (many clauses), the parser is more likely to generate errors. Moreover, if some syntactic transformation was performed, like passive-to-active, the transformation stage might generate errors. However, if the source sentence is in reporting form using reporting verbs, it most likely generates a question without syntactic errors because this type of sentence is normally simple and transformation rules for this type of source sentence are well defined.

Some NLP tools are used or developed to extract these features. A state of art Named Entity Tagger, LBJ (Ratinov & Roth, 2009), was used to extract name entities in a sentence. A simple Pronoun Resolver, finding the nearest Name Entity appearing before the pronoun, was implemented to identify citations with pronominals. Tregex rules are developed to detect number of clauses and passive voice detection. The reporting verb list was obtained from academic writing tutorial websites (Centre for Academic Success, 2011).

The source sentence refers to the citation sentence in the descriptions of the 11 features defined below:

**Num. of Named Entities:** This numeric feature describes the number of author names in the source sentence.

**IsPronounResolved:** This Boolean feature detects whether the pronoun resolution has been resolved.

**IsPassiveVoice:** This Boolean feature detects whether the source sentence is passive voice.

**ContainsAnswer:** This is a Boolean feature that detects the presence of answer cue phrases, such as due to, in order to, considering, etc.

**Negation:** This Boolean feature detects whether the source sentence contains negation cue phrases (not, no, never), restrictive adverbs (few, rarely, seldom), negative verbs (fail, deny), or negative adjectives (insufficient).

**PredicationConfidence:** This numeric feature shows the predication confidence from the citation classifier.

**ReportingVerb:** This is a Boolean feature that detects whether the source sentence contains reporting verbs, such as show, argue, discuss, and explain.

**NameAppearInBoth:** This is a Boolean features indicating the presence of name entities in both subject and predicate.

**NameAppearInSubject:** This is a boolean feature to detect the presence of name entities in subject.
**Num. of Clauses:** This numeric feature shows the number of clauses in the source sentence. Each clause consists of a noun phrase (NP) followed by a verb phrase (VP).

**Length:** This feature judges whether the source sentence is too short or too large.

### An application scenario in a writing environment

![Image of a writing support environment](Figure 2. AQG in a writing support environment)

This section briefly describes the potential application of the AQG tool which will be integrated into iWrite (Calvo, O'Rourke, Jones, Yacef, & Reimann, 2011). iWrite is a web-based writing support environment which allows students to write and submit their assignments and provides them with a complete solution for supporting the write-review-feedback cycle of a writing activity. iWrite uses automatic feedback tools that help students conduct revision, such as Glosser (Villalon, Kearney, Calvo, & Reimann, 2008). AQG will be integrated into iWrite as one of the feedback tools.

In Figure 2, the input to the iWrite is an academic writing paper, written in natural text by a student using Google Docs. When the assignment deadline is due, the iWrite assignment manager downloads the student’s document and passes it to Intelligent Feedback Tools for processing. The output is the textual feedback including trigger questions, generated by human or intelligent feedback tools, such as AQG, which is delivered to the student author.

### Evaluation and result

#### Data collection

In order to train our question ranking model, 504 questions were collected and manually annotated as a training set. Those questions were generated from citations included in 45 academic papers from ACL conference and International Joint Conferences on Artificial Intelligence (IJCAI). Our testing set includes 489 questions generated from 33 literature review papers separately written by 33 engineering research students enrolled in a Research Methods course. Two human experts, who also annotated the citations for the classification task, were asked to label each question as acceptable or unacceptable according to two major criteria, (1) grammaticality and (2) semantic correctness. Grammatical correctness refers to the presence or absence of grammar errors. Semantic correctness refers to the overall meaning of the generated question within the context of the citation sentence that triggered the question, e.g., a generated question may not make any sense due to wrong author name entity extraction or wrong question template used because of errors in the citation classification step. For example, in the following generated question: Why does MANET conduct this study to consider the effect of velocity of the nodes?, MANET was wrongly identified as an author. The Cohen’ Kappa inter-agreement between human annotators was 0.65 (substantial agreement).
Besides collecting system generated questions from each literature review paper, questions were collected including those from the student’s supervisors, peers, and five generic questions. Table 4 shows the five generic questions used by the research method course intended to help students write literature review.

Table 4. Five generic questions used in literature review support

<table>
<thead>
<tr>
<th>Generic Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did your literature review cover the most important relevant works in your research field?</td>
</tr>
<tr>
<td>Did you clearly identify the contributions of the literature reviewed?</td>
</tr>
<tr>
<td>Did you identify the research methods used in the literature reviewed?</td>
</tr>
<tr>
<td>Did you connect the literature with the research topic by identifying its relevance?</td>
</tr>
<tr>
<td>What were the author's credentials? Were the author's arguments supported by evidence?</td>
</tr>
</tbody>
</table>

Like the Bystander Turing Test conducted by Graesser et al. (2004), our judges, the student writers, rated the quality of each generated question using a Likert scale (1 was “strongly disagree” and 5 “strongly agree”):
1. QM1: This question is correctly written.
2. QM2: This question is clear.
3. QM3: This question is appropriate to the context.
4. QM4: This question makes me reflect about what I have written.
5. QM5: This is a useful question.

Ranking model performance

The following ranking models were derived in order to compare algorithms and feature sets.

*RankSVM*: This model is implemented by RankSVM algorithm. The RankSVM (ALL) includes all the features described in previous section. According to the characteristic of each feature, the feature set is grouped into two clusters: syntactic group and semantic group. RankSVM(Basic) only contains the basic or syntactic features including isPassive Voice, Negation, Number of Clauses, isPronResolved, and Length. RankSVM(Semantic) includes only the semantic features: Number of Named Entities, Predication Confidence, Reporting Verb, NameAppearInBoth, NameAppearInSubject, ContainAnswer.

*Logistic Regression*: The model is trained using all the features

*Baseline*: The expected performance if questions were ranked randomly

*Gold Standard*: The expected performance if all the acceptable questions were ranked higher than unacceptable

The baseline was 0.657 (65.7% of all test set questions were labeled as acceptable by the human annotator). Figure 3 shows that most of ranking models were unstable in the top 100 questions. The RankSVM (All) and Logistic Regression had very similarly sharp curve, which are higher than baseline overall. Table 5 shows the ranking results of these models for the top 25% and 50% of the ranked questions. In the top 120 questions (25%), RankSVM (All) got better accuracy (0.758) than Logistic Regression (0.742). RankSVM (Basic) got the lowest accuracy (0.708). Basic features helped as the performance improves from 0.733 for RankSVM (Semantic) to 0.758 for RankSVM (All) with all the features. A one-way ANOVA, at a 95% confidence level, was conducted to examine whether there are statistical differences among these models. The ANOVA indicated a significant difference, F(3,476)= 3.08, p<0.05. Fishers’ least significant difference (LSD) tests at the 95% confidence level were performed to determine whether significant differences occurred between the mean scores for each pair of treatments. Results indicated no statistical difference between RankSVM (All) and Logistic Regression while RankSVM (ALL) significantly outperformed RankSVM (Basic). In the top 224 ranked questions (50%), the RankSVM (All), RankSVM (Semantic), Logistic Regression got equivalent scores, all of which were better than the RankSVM (Basic).
Figure 3. A graph of the percentage of acceptable questions in the top ranked 489 questions using various models.

Table 5. The percentage of the top 25% and 50% of ranked questions labeled acceptable with various models

<table>
<thead>
<tr>
<th>Model</th>
<th>Top 25%</th>
<th>Top 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>RankSVM (All Feature)</td>
<td>0.758</td>
<td>0.713</td>
</tr>
<tr>
<td>RankSVM (Semantic Feature)</td>
<td>0.733</td>
<td>0.713</td>
</tr>
<tr>
<td>RankSVM (Basic Feature)</td>
<td>0.708</td>
<td>0.692</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>0.742</td>
<td>0.713</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.657</td>
<td>0.656</td>
</tr>
</tbody>
</table>

Evaluation of the top ranked questions

The ability of our AQG system to generate quality and effective questions is explored by comparing questions generated by the system to those produced by human supervisor, peer, and five generic questions commonly used in writing literature review. For the evaluation, supervisors generated 107 questions, peers 133, our AQG system top ranked 126, and we also used 161 generic questions (five generic questions have been used repeatedly for each document).

Table 6. Comparisons of Normalized Mean Scores

<table>
<thead>
<tr>
<th>Question Producer</th>
<th>QM1</th>
<th>QM2</th>
<th>QM3</th>
<th>QM4</th>
<th>QM5</th>
<th>Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQG</td>
<td>4.26</td>
<td>4.18</td>
<td>4.06</td>
<td>3.99</td>
<td>4.02</td>
<td>4.10</td>
</tr>
<tr>
<td>Supervisor</td>
<td>4.57</td>
<td>4.53</td>
<td>4.45</td>
<td>4.26</td>
<td>4.21</td>
<td>4.40</td>
</tr>
<tr>
<td>Peer</td>
<td>3.93</td>
<td>3.96</td>
<td>3.77</td>
<td>3.62</td>
<td>3.56</td>
<td>3.77</td>
</tr>
<tr>
<td>Generic</td>
<td>4.04</td>
<td>3.92</td>
<td>3.65</td>
<td>3.49</td>
<td>3.57</td>
<td>3.73</td>
</tr>
</tbody>
</table>

A one-way ANOVA and post-hoc analysis using Fisher’s least significant difference (LSD) tests were performed. Table 6 and Table 7 show that questions from the AQG system significantly outscored generic questions in each quality measure (QM) while outperforming peers’ questions in quality measures 1, 3, 4, 5 and Average. The difference between supervisor’s questions and the system for QM5 is not significant. This result indicates that questions produced by the system are perceived to be as useful as human supervisors. This positive result might be explained by two factors. First, the system’s questions are useful because such questions are content-related and have...
semantic meaning. Second, students may have intended to give high scores to questions which they thought were from supervisors, where actually they were from the system.

Table 7. Fisher’s Least Significant Difference (LSD) tests

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Supervisor Vs AQG</th>
<th>AQG Vs Peer</th>
<th>AQG Vs GQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>QM1</td>
<td>MD = 0.308</td>
<td>MD = 0.330</td>
<td>MD = 0.225</td>
</tr>
<tr>
<td></td>
<td>LSD = 0.242*</td>
<td>LSD = 0.230*</td>
<td>LSD = 0.219*</td>
</tr>
<tr>
<td>QM2</td>
<td>MD = 0.350</td>
<td>MD = 0.220</td>
<td>MD = 0.263</td>
</tr>
<tr>
<td></td>
<td>LSD = 0.248*</td>
<td>LSD = 0.236</td>
<td>LSD = 0.225*</td>
</tr>
<tr>
<td>QM3</td>
<td>MD = 0.385</td>
<td>MD = 0.297</td>
<td>MD = 0.418</td>
</tr>
<tr>
<td></td>
<td>LSD = 0.258*</td>
<td>LSD = 0.246*</td>
<td>LSD = 0.235*</td>
</tr>
<tr>
<td>QM4</td>
<td>MD = 0.269</td>
<td>MD = 0.368</td>
<td>MD = 0.501</td>
</tr>
<tr>
<td></td>
<td>LSD = 0.257*</td>
<td>LSD = 0.245*</td>
<td>LSD = 0.234*</td>
</tr>
<tr>
<td>QM5</td>
<td>MD = 0.182</td>
<td>MD = 0.460</td>
<td>MD = 0.452</td>
</tr>
<tr>
<td></td>
<td>LSD = 0.269</td>
<td>LSD = 0.256*</td>
<td>LSD = 0.245*</td>
</tr>
<tr>
<td>Average</td>
<td>MD = 0.299</td>
<td>MD = 0.335</td>
<td>MD = 0.372</td>
</tr>
<tr>
<td></td>
<td>LSD = 0.226*</td>
<td>LSD = 0.215*</td>
<td>LSD = 0.206*</td>
</tr>
</tbody>
</table>

Note. 95% confidence interval and the * represents a significant difference.

Human question types evaluation and result

In order to design an efficient question template, it is very important to investigate the kinds of questions that human reviewers (supervisor and peer in this case) typically generate and how valuable each question type is. The citation questions were classified into 3 categories:

- Association questions, which address the entities of the system and their properties, such as who, what, when, or where questions. E.g., What is a biodegradable polymer?
- Explanation questions, which focus on justifications or explanation for these entities, such as why or how questions. E.g., Why is it difficult to control the driving forces of natural ventilation?
- Prediction questions, which address the need to foresee consequences, such as what happens next or if-then. E.g., Do you believe that future progress in microphysics modelling lies with bin or bulk microphysics modelling approaches?

These categories were proposed by Trabasso and Magliano (1996) as a classification of “Information Seeking Questions” (ISQs), i.e. questions that address knowledge deficits when subjects process scientific texts with a particular goal in mind. Since the questions generated by AQG system are “why” related questions, they are mapped to the Explanation category. Table 8 shows that Association and Explanation questions form the majority of human-generated questions, while there are only few Prediction questions.

Table 8. The number of questions generated according to question categories

<table>
<thead>
<tr>
<th>Source</th>
<th>Type (for citation–related questions only)</th>
<th>Example</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supervisor</td>
<td>Association</td>
<td>Which type of the wind turbine (A to D) has the highest potential electrical conversion efficiency?</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Explanation</td>
<td>Why coatings and linings are needed to apply to form a monolithic layer in the sewer pipes to inhibit further deterioration in Najafi’s paper?</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Prediction</td>
<td>Do you believe that future progress in microphysics modelling lies with bin or bulk microphysics modelling approaches?</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Association &amp; Explanation</td>
<td>What are the strength and ductility of materials and how are they related to the energy absorption ability of the materials?</td>
<td>3</td>
</tr>
<tr>
<td>Peer</td>
<td>Association</td>
<td>What is the beginning process of MIC that shown by</td>
<td>89</td>
</tr>
</tbody>
</table>
Why Carrasco et al. regarded wind energy as a relatively mature technology?

In your opinion, how NGMN can play a major role in future of Internet?

What is Emergence? How does it examine agencies interaction?

In the study of Mulligan, why are organic acids including citric, oxalic, malic, etc extensively used in mineral leaching?

Table 9 shows that Explanation questions are perceived to be the most valuable while Generic questions are the least valuable. A one-way ANOVA test indicated that there are significant differences between Association, Explanation, Generic, and Non-citation related questions in terms of their perceived quality ($F(4,398)=3.296, P<0.05$). Non-citation questions address presentation issues, such as referencing, formatting, and numbering. A posthoc test found that Explanation questions including both system and human generated questions significantly outscore Association and Generic questions (see Table 10).

**Table 9. The average score of different question types**

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Average Perceived Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association (human)</td>
<td>3.94</td>
</tr>
<tr>
<td>Explanation (human)</td>
<td>4.24</td>
</tr>
<tr>
<td>Non-Citation related (human)</td>
<td>3.93</td>
</tr>
<tr>
<td>e.g., Reference citing should be checked</td>
<td></td>
</tr>
<tr>
<td>Generic (human)</td>
<td>3.51</td>
</tr>
<tr>
<td>e.g., What is the relationship between your literature and your research?</td>
<td></td>
</tr>
<tr>
<td>Explanation (AQG)</td>
<td>4.10</td>
</tr>
</tbody>
</table>

**Table 10. Fisher's least significant difference (LSD) tests**

<table>
<thead>
<tr>
<th>Comparison in Pair</th>
<th>Association</th>
<th>Explanation</th>
<th>Non-Citation Related</th>
<th>Generic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanation</td>
<td>MD=0.30*</td>
<td>LSD=0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Citation Related</td>
<td>MD=0.01</td>
<td>LSD=0.38</td>
<td>MD=0.31</td>
<td></td>
</tr>
<tr>
<td>Generic</td>
<td>MD=0.44</td>
<td>LSD=0.45</td>
<td>MD=0.74*</td>
<td>MD=0.43</td>
</tr>
<tr>
<td>Explanation(AQG)</td>
<td>MD=0.16</td>
<td>LSD=0.23</td>
<td>LSD=0.25</td>
<td>LSD=0.55</td>
</tr>
<tr>
<td></td>
<td>MD=0.14</td>
<td>LSD=0.38</td>
<td>LSD=0.38</td>
<td>LSD=0.45</td>
</tr>
</tbody>
</table>

*Note. 95% confidence interval and the * represents a significant difference.

**Discussion and conclusion**

Automatic generation of natural questions is a challenging task. This article addressed this challenge and presented a novel AQG system for supporting academic writing by applying overgeneration-and-ranking approaches. The results indicate that this approach is effective since the ranking model improved the acceptability of the top 25% questions by 10%. In particular, RankSVM slightly outperformed Logistic Regression and the experiments revealed that the semantic features are important. However, if the document is poorly written and citation sentences contain grammatical errors, it would cause the system to generate no questions or poor questions with grammatical errors. It is because the question generation is a pipeline process and any error occurs at one stage will influence the following stages. Citations sentence containing grammatical errors could influence the parser to correctly extract predicate verb for sentence classification and transformation. This could cause to misclassify the sentence or incorrectly transform the sentence into a question.
As expected, the top ranked questions generated by the system outperformed the generic questions and are as useful as human generated one if excluding some questions with surface errors. One reason the system generated questions are as good as the human-generated ones is because the system questions are specific and addressed critical thinking aspects.

Explanation and Association questions are mostly common used by human. Particularly, Explanation questions are more useful than other questions types because they normally trigger deep reflection and invoke critical thinking. This would inspire us to design an effective question template for the AQG system. However, it has been found that association questions are also frequently used by human supervisors and peers. These questions are still valuable to help students to understand key concepts described in the document. Our future work will focus on generating association questions from the key concepts by using information extraction techniques.

However, the question ranking model may not be applied to other question generation approaches since some of the defined features are only related to citations. To generalize this question ranking model, more generic question generation approaches and fine-grained generic features are needed. Despite these shortcomings, we believe that this AQG approach is effective and the evaluation meaningful because real academic writings were used.

Our future work will investigate how the system generated feedback questions would influence student’s behaviors and learning performance. For example, a similar writing activity can be conducted, which consists of a draft session and a revise session. In draft session, a student writes a draft proposal and submits to his/her supervisor. After submitting the draft proposal, the student received a score and feedback based on the draft. The feedback could be either human feedback or system feedback. After receiving the feedback, the student revised the proposal in the revise session and finally gets a score for the final proposal. By this way, the system can track how many changes the students would make in the writing after receiving the feedback questions. The system can also found out the score difference between draft proposal and final proposal.

Acknowledgements
Ming is supported by the Young and Well CRC. This work is partially supported by Fundamental Research Funds for the Central Universities under Grant No. XDJK2014A002 and No. XDJK2014C141 and No. SWU114005.

References


Implementation of a Model-Tracing-Based Learning Diagnosis System to Promote Elementary Students’ Learning in Mathematics

Yian-Shu Chu1, Haw-Ching Yang2, Shian-Shyong Tseng3* and Che-Ching Yang1

1Department of Computer Science, National Chiao Tung University, Hsinchu, 30010, Taiwan // 2Institute of System Information and Control, National Kaohsiung First University of Science and Technology, Kaohsiung, 811, Taiwan // 3Department of Applied Informatics and Multimedia, Asia University, Taichung, 41354, Taiwan // trash@cis.nctu.edu.tw // hao@nkfust.edu.tw // sstseng@asia.edu.tw // jerome@cis.nctu.edu.tw

*Corresponding author

(Submitted July 12, 2012; Revised December 13, 2012; Accepted March 16, 2013)

ABSTRACT

Of all teaching methods, one-to-one human tutoring is the most powerful method for promoting learning. To achieve this aim and reduce teaching load, researchers developed intelligent tutoring systems (ITSs) to employ one-to-one tutoring (Aleven, McLaren, & Sewall, 2009; Aleven, McLaren, Sewall, & Koedinger, 2009; Anderson, Corbett, Koedinger, & Pelletier, 1995; Anderson & Reiser, 1985; Blessing, Gilbert, Ourada, & Ritter, 2009; Mitrovic et al., 2009; Mitrovic & Ohlsson, 1999; Suraweera, Mitrovic, & Martin, 2007; VanLehn et al., 2005). However, most ITSs have restricted user interfaces, which confine reasoning strategies of students during problem solving, thus ignoring the fact that students could use dissimilar strategies to solve a given question. Furthermore, student learning problems could be diagnosed from the derivation of their answers. In order to interpret students’ mathematical problem-solving behaviors, this study developed a Model-tracing Intelligent Tutor (MIT) to diagnose students’ learning problems and provide learning feedback for individual students. A quasi-experiment was conducted in an elementary school to evaluate the effectiveness of the proposed approach, in which 124 fifth graders participated. The experimental results show that the model-tracing-based learning diagnosis system is significantly more helpful to the students in learning mathematics than the conventional web-based test in terms of learning achievements.

Keywords

Learning diagnosis, Computer-assisted learning, Computer-assisted testing, Model tracing

Introduction

One-to-one human tutoring, which has been shown to be much more effective than conventional classroom instruction, enables higher achievements in most students (Bloom, 1984; Chou, Huang, & Lin, 2011; VanLehn, 2006). However, one-to-one human tutoring is extremely costly, and one-to-many classroom instruction leaves little time for teachers to take care of individual student needs (Zinn & Scheuer, 2006). Intelligent tutoring systems (ITSs) have been developed to cope with the problems stated above and to employ one-to-one tutoring (Aleven, McLaren, & Sewall, 2009; Aleven, McLaren, Sewall, et al., 2009; Anderson et al., 1995; Anderson & Reiser, 1985; Blessing et al., 2009; Mitrovic et al., 2009; Mitrovic & Ohlsson, 1999; Suraweera et al., 2007; VanLehn et al., 2005).

Although researchers have demonstrated the benefits of ITSs in a range of domains (Chou et al., 2011; Chu, Hwang, & Huang, 2010; Lee & Bull, 2008; VanLehn et al., 2005), some issues still need to be further discussed (VanLehn et al., 2005). One of the issues is many ITSs have user interfaces that guide students’ reasoning strategies by restricting the intermediate steps that students should follow. In other words, those ITSs adopted the restricted user interfaces to confine reasoning by offering a fixed type-in box for entering the intermediate steps. Because a given question can be solved by different strategies, students’ reasoning process should not be confined.

Singley and Anderson (1989) pointed out that one gets higher transfer from training to testing when the user interfaces are similar. Therefore, students who use a tutoring system with a less constrained user interface to learn might have similar learning gains to those who use pencil and paper (VanLehn et al., 2005). Moreover, keeping the user interface less constrained makes the tutoring system less invasive.

In addition, in a paper-and-pencil calculation test, learning problems of students could be diagnosed from answering derivations (in particular, the step-by-step reasoning processes). The answering derivations denote the necessary expressions, equations, or functions while solving a mathematical problem. Because students may have different
levels of misconceptions or mistakes during the problem-solving process, development of a less invasive tutoring system that can friendly interpret mathematical problem-solving behaviors is an important issue.

In general, there are three types of ITSs (cognitive tutors, constraint-based tutors, and example-tracing tutors), which possess different degrees of machine intelligence and provide different interaction mechanisms and different implementation complexities. Cognitive tutors apply a model-tracing approach to interpret and assess student behavior with reference to a cognitive model (Anderson et al., 1995; Anderson & Reiser, 1985; Blessing et al., 2009; VanLehn et al., 2005). Constraint-based tutors apply a constraint-based modeling approach to interpret and assess student work with respect to a set of constraints (Mitrovic, Martin, Suraweera, Zakharov, Milić, & Holland, 2009; Mitrovic, & Ohlsson, 1999; Suraweera, Mitrovic, & Martin, 2007). Example-tracing tutors apply the example-tracing approach to interpret and assess student behavior with reference to generalized examples of problem-solving behavior (Aleven, McLaren, & Sewall, 2009; Aleven, McLaren, Sewall, et al., 2009).

Since the example-tracing tutor’s knowledge is not generalizable to other similar problems, it fails to reason multiple problem scenarios by itself. This research is to discover the strategies and misconceptions students may acquire; therefore, a model-tracing-based approach is used to compare students’ activities with a cognitive model of student problem solving to achieve the goal of interpretation of students’ behaviors. Previous research, such as LISP Tutor (Anderson & Reiser, 1985), algebra cognitive tutor (Anderson et al., 1995), and Andes (VanLehn et al., 2005) showed that the model-tracing approach not only can analyze cognitive behaviors of students but can also evaluate their knowledge.

In addition, as a minimum requirement an ITS must have an “inner loop,” that is, provide minimal feedback within problem-solving activities (VanLehn, 2006). In this study, minimal feedback is to provide a timely diagnosis report on the final result (whether the whole solution is correct) and the working steps (which step is incorrect and what is the cause of error).

In order to cope with previous problems, this study has developed a testing and diagnostic system based on tutoring behavior identified by VanLehn (2006). The proposed system, Model-tracing Intelligent Tutor (MIT), includes four components: (1) lexical analyzer (scanner); (2) syntax analyzer (parser); (3) semantic analyzer; and (4) report generator. MIT is implemented with the aim of conducting a one-to-one tutoring mechanism with instant feedback to improve learning in mathematics of students. Therefore, the research question is “what are the learning achievements of students after using MIT.” Finally, an experiment on a fraction lesson in a mathematics course was conducted to demonstrate the effectiveness of the proposed system.

The remainder of this paper is organized as follows. Section 2 describes student problem-solving behavior in fractions. Section 3 presents how we built MIT and how MIT can be used as a web-based test system. Section 4 describes the methodology to evaluate the presented MIT and experimental results. Section 5 presents discussions and future research directions, and Section 6 draws conclusions.

Student problem-solving behavior in fractions

Following the work of Aleven, McLaren, Sewall, et al. (2009), solutions for a mathematical problem could be represented as a behavior graph (directed and acyclic), which may contain multiple paths corresponding to different ways of solving the problem. For example, a problem-solving behavior to solve a fraction question such as “adding mixed fractions” could be as follows: converting a mixed number to an improper fraction, reducing fractions to a common denominator, adding fractions with common denominators, and converting an improper fraction to a mixed number. Another way of solving the problem is by adding integers, reducing fractions to a common denominator, adding fractions with common denominators, and converting an improper fraction to a mixed number.

Also, a behavior graph may contain incorrect behavior links. In the graph, the links represent problem-solving actions, and the nodes show problem-solving states. In terms of incorrect behavior, when solving fractions, students may have common misconceptions that lead to errors in computation (Idris & Narayanan, 2011; Lee & Bull, 2008; Stead, 2012; Tatsuoka, 1984; Tirosh, 2000). Because of rote memorization and insufficient knowledge, students may overuse generalize rules and procedures when following the steps in worked-out examples (Idris & Narayanan,
Therefore, errors of students are often systematic and rule-based rather than random (Idris & Narayanan, 2011). Table 1 illustrates the typical misconceptions students may have.

**Table 1. Typical misconceptions**

<table>
<thead>
<tr>
<th>Operation</th>
<th>Behavior description</th>
<th>Example</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed to improper fractions</td>
<td>Move the whole number to the numerator</td>
<td>( \frac{3}{5} = \frac{13}{5} )</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Add the whole number to the numerator</td>
<td>( \frac{3}{5} = \frac{4}{5} )</td>
<td>B</td>
</tr>
<tr>
<td>Raising a fraction to higher terms</td>
<td>Multiply the numerator and the denominator by two different numbers, respectively</td>
<td>( \frac{2}{5} = \frac{10}{20} )</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>A random number is chosen and that number is added to both the numerator and the denominator</td>
<td>( \frac{2}{5} = \frac{4}{7} )</td>
<td>D</td>
</tr>
<tr>
<td>Reducing a fraction to lower terms</td>
<td>Divide the numerator and the denominator by two different numbers, respectively</td>
<td>( \frac{4}{8} = \frac{1}{4} )</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>A random number is chosen and that number is subtracted from both the numerator and the denominator</td>
<td>( \frac{2}{10} = \frac{1}{9} )</td>
<td>F</td>
</tr>
<tr>
<td>Adding fractions with uncommon denominators</td>
<td>Multiply the numerators together and multiply the denominators together</td>
<td>( \frac{2}{5} + \frac{1}{2} = \frac{2}{10} )</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>Add the numerators and multiply the denominators</td>
<td>( \frac{2}{5} + \frac{1}{2} = \frac{3}{10} )</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>Add the numerators together and add the denominators together</td>
<td>( \frac{2}{5} + \frac{1}{2} = \frac{3}{7} )</td>
<td>I</td>
</tr>
</tbody>
</table>

**Model-tracing Intelligent Tutor (MIT)**

A fraction question may comprise several sub-questions which may be solved in a variety of orders or ways. Furthermore, students may have misconceptions that lead to errors in computation. To interpret and assess student behavior, a learning diagnosis system, “Model-tracing Intelligent Tutor (MIT)” was developed to trace mathematical step-by-step operations of students after answering a fraction question. The running phase of MIT is as shown in Figure 1, which includes four components: lexical analyzer (scanner), syntax analyzer (parser), semantic analyzer, and report generator. The lexical analyzer reads the input stream (e.g., mathematical equation) and passes tokens to the parser. The syntax analyzer reads the tokens and identifies the syntactic structure. The semantic analyzer drives semantic processing. Finally, the report generator generates diagnosis results for students.
In this study, Lex, the lexical analyzer generator, is used to generate a scanner for dealing with lexical analysis. Yet Another Compiler Compiler (YACC) (Johnson, 1975) is used to generate a set of parse tables for analyzing syntax and semantic. To specify the effectiveness of the model-tracing mechanism, we address the use of YACC specifications to generate this learning diagnosis system. Also, at the end of this section, we describe the development of MIT.

The specification

YACC, a context-free language parser generator, is a look-ahead left-to-right, rightmost-derivation (LALR) parser generator developed by AT&T Bell Laboratory in C language, which is used to generate parsers with a given input file. YACC also supports a general mechanism for semantic analysis. The input file consists of three sections: declarations, productions, and subroutines. To develop the MIT system, the declarations section defines mathematical symbols used during operations, the productions section specifies rules for mathematical equations, and the subroutines section denotes rules for correctness determination of solving fractions. Below are the related definitions and examples.

Definition 1 (declarations). Mathematical symbols (e.g., operands and operators) used in an equation are defined in the declarations section.

An expression is a set of terms with mathematical operations combining them, and an equation consists of two expressions with a relational symbol between them. Some terms related to an equation are listed below.

- **Token**: There are nine different tokens used in an equation: “INT” (integer), “/” (fraction bar), “+” (addition), “−” (subtraction), “∗” (multiplication), “/” (division), “(” (left parentheses), “)” (right parentheses), and “=” (equal).
- **Primary**: A fraction is called a primary in the specification, which is written in the form of \( \frac{c}{b} \), where \( a, b \) and \( c \) are integers and \( b \) cannot be 0. The number \( c \) is called the numerator, and the number \( b \) is called the denominator. There are three kinds of primaries: proper and improper common fractions and mixed numbers. The fraction is called proper if the numerator is less than the denominator, and improper otherwise. A mixed number consists of an integer and a proper fraction.
- **Operator**: The Operators are “+”, “−”, “∗”, “/”, “(”, “)”, and “=”.

Definition 2 (productions). Rules of the mathematical equation expressions are defined in the productions section.

Four rules are listed below. Eqs. 1 to 3 show the productions of fraction addition. Eq. 4 shows a primary can be recognized if its format is \( \frac{INT}{INT} \).

\[
\begin{align*}
\langle expression \rangle & \rightarrow \langle expression \rangle + \langle multiplicative expression \rangle \\
\langle expression \rangle & \rightarrow \langle multiplicative expression \rangle \\
\langle multiplicative expression \rangle & \rightarrow \langle primary \rangle \\
\langle primary \rangle & \rightarrow \frac{INT}{INT}
\end{align*}
\]

(1) (2) (3) (4)

Example 1 (syntax analysis):

According to Eq. 4, the expression \( \frac{3}{5} + \frac{1}{2} \) can be recognized as two primaries with one operator combining them. Then, this expression can be recognized as an expression of fraction addition by using Eqs. 1 to 3.
Definition 3 (subroutines). The subroutines in YACC, which are rules for correctness determination of solving fraction, are used to analyze learning status of students while solving fraction questions.

Three rules are listed below for converting a mixed number to an improper fraction. Among the rules, Eq. 5 is a correct pattern. This pattern is used to recognize correct conversion of a mixed number to an improper fraction, which consists of two intersection rules: \( Y_2 = X_2 \), \( Y_3 = X_3 + X_2 \times X_1 \). Since students may have misconceptions while solving questions, Eq. 6 and Eq. 7 are illustrative misconceptions, A and B, respectively, which are mentioned in Table 1 (Idris & Narayanan, 2011; Lee & Bull, 2008; Stead, 2012; Tatsuoka, 1984; Tirosh, 2000).

\[
\frac{X_1}{X_2} = \frac{Y_3}{Y_2}
\]

**Correct pattern**

\[
Y_2 = X_2 \cap Y_3 = X_3 + (X_2 \times X_1)
\] (5)

**Misconception A**

\[
Y_2 = X_2 \cap Y_3 = X_3 + (10 \times X_1)
\] (6)

**Misconception B**

\[
Y_2 = X_2 \cap Y_3 = X_3 + X_1
\] (7)

Example 2 (correct behavior):

Following the rule of Eq. 5, an equation \( \frac{3}{5} + \frac{8}{5} \) can be recognized as a fraction question solving step of converting a mixed number to an improper fraction during the fraction solving process.

Example 3 (incorrect behavior):

By using Eq. 6, an equation \( \frac{3}{5} = \frac{13}{5} \) can be recognized as an error step of converting a mixed number to an improper fraction. Similarly, an equation \( \frac{3}{5} = \frac{4}{5} \) also can be recognized as an error step with the aid of Eq. 7.

The development

We implemented the learning diagnosis system MIT based on the specifications. Figure 2 shows the left-to-right trace process of MIT. Three different kinds of operations are adopted in this figure: raising a fraction to higher terms, adding fractions with common denominators, and converting an improper fraction to a mixed number.

![Figure 2. The trace process](http://www.w3.org/Math/)

This system characterizes not only strategies but also the misconceptions that students may acquire by mapping students’ problem-solving steps to error steps. Also, MIT gives immediate feedback right after students click on the Submit Answer button, i.e., feedback is given after the completion of every item. Then the tutor analyzes the entered answers of all the steps and determines the giving diagnosis results.

As shown in Figure 3, students key their responses into a text box. To display mathematical expressions in web browsers, Mathematical Markup Language (MathML, an XML-based markup language recommended by the W3C math working group http://www.w3.org/Math/), is used to describe mathematical notations and capture both its
structure and content. Table 2 illustrates the MathML’s representation of \((a + b)^2\). Because editing MathML directly to express mathematical expressions is difficult for students, a client-side mathematical markup language ASCIIMathML (http://www1.chapman.edu/~jipsen/mathml/asciimath.html) is used in this study to assist students in typing mathematical expressions with ASCII codes and converting the expressions to MathML syntax. In other words, students type a fraction “6 4/5” and browsers will display the typed text as \(6 \frac{4}{5}\).

As shown in Figure 3, students key their responses into a text box. To display mathematical expressions in web browsers, Mathematical Markup Language (MathML, an XML-based markup language recommended by the W3C math working group http://www.w3.org/Math/), is used to describe mathematical notations and capture both its structure and content. Table 2 illustrates the MathML’s representation of \((a + b)^2\). Because editing MathML directly to express mathematical expressions is difficult for students, a client-side mathematical markup language ASCIIMathML (http://www1.chapman.edu/~jipsen/mathml/asciimath.html) is used in this study to assist students in typing mathematical expressions with ASCII codes and converting the expressions to MathML syntax. In other words, students type a fraction “6 4/5” and browsers will display the typed text as \(6 \frac{4}{5}\).

Meanwhile, an input assistant tool is provided as another input method, as shown in the bottom of Figure 3. Students can choose to type fractions or mathematical notations to input their answers via the input assistant tool. For example, students input a fraction and then click the operator “add” to display it in the text box. If students enter an incorrect equation, a pop-up window will prompt to display the error message “Please check your equation carefully!”

<table>
<thead>
<tr>
<th>Test item no.</th>
<th>Display student’s answer</th>
<th>Input answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID: stu001</td>
<td>(\frac{3}{4} + \frac{5}{6} = \frac{8}{10} = \frac{4}{5})</td>
<td>(\frac{3}{4} + \frac{5}{6} = \frac{6}{10} = \frac{6}{4})</td>
</tr>
</tbody>
</table>

**Figure 3. A screenshot of user input interface**

Table 2. The MathML representation of \((a + b)^2\)

```xml
<msup>
  <mfenced>
    <mi>a</mi>
    <mo>+</mo>
    <mi>b</mi>
  </mfenced>
  <mn>2</mn>
</msup>
```
After students finish answering the questions, the MIT system starts to trace their operations. MIT traces not only the final result but also the operating steps. The step analyzer shows the operation result as well as the cause of error.

Figure 4 shows the diagnosis results of operations of the student in Figure 3. The student used two operations: adding fractions with uncommon denominators and reducing a fraction to lowest terms. Because of the incorrect operation of “adding fractions with uncommon denominators,” the final result is incorrect. The cause of error in this operation is “add the top numbers and the bottom numbers.”

![Figure 4. A screenshot of learning diagnosis](image)

### The experiment and evaluation

In order to evaluate the effectiveness of this approach, we conducted a three-week experiment. This research adopted a quasi-experimental design. There were 124 fifth-grade students participating in this research. The students were from four classes and taught by the same instructor under the same conditions. They had previously taken computer courses and possessed basic computer skills. The four participating classes were randomly divided into experimental group and control group, each of which consisted of sixty-two students. For the students in the experimental group, the answers of individual students were analyzed by using MIT. Each student in the experimental group can get a detailed analysis of incorrectly answered item along with the test results. The students in the control group used a conventional web-based test system and only the test results (correct or incorrect) were given. The only difference between a conventional web-based test system and a traditional paper-and-pencil test is that a conventional web-based test system can be done online.

During the three weeks, the two groups of students received instruction about fraction calculations; moreover, a paper-and-pencil pre-test was conducted to analyze the students’ knowledge. Then, the students in two groups were instructed with the tools of the learning activity. After the instruction, we conducted a 60-minute learning activity. All students participated in web-based test system. Students in the experimental group used MIT while students in the control group used conventional web-based test system. After the learning activity, all participants took a paper-and-pencil post-test.

Before the experiment began, three senior mathematics teachers who had more than five years of experience in teaching mathematics selected thirty fractional additions as item candidates. To calculate item difficulty and item discrimination, teacher participants administered these items to the students in four other different classes at the same elementary school. According to difficulty and discrimination, we chose ten items for pre-test and other similar ten items for post-test. The average difficulty of the pre-test and post-test was 0.539 and 0.532, respectively. The discrimination index of each item in both tests was over 0.500.
The scores of the students in the pre- and post-tests were analyzed to compare the learning achievements of the students in the two groups. Table 3 shows the independent t-test of the pre-test scores between the two groups. In the pre-test, the average scores of the two groups were 55.48 and 60.00, and the control group had a higher average score than the experimental group. The t-test ($t = 0.826, p > .05$) shows that there is no difference between the two groups’ average pre-test scores. In other words, students in both groups had the same prior knowledge of fractions before the implementation of the experimental treatments.

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>62</td>
<td>55.48</td>
<td>28.61</td>
<td>0.826</td>
</tr>
<tr>
<td>Control group</td>
<td>62</td>
<td>60.00</td>
<td>32.14</td>
<td></td>
</tr>
</tbody>
</table>

Before applying ANCOVA, the homogeneity of the regression coefficient was tested, which revealed that interaction $F(1, 120)$ between the covariance was 0.171 ($p > 0.05$). This confirms the hypothesis of homogeneity of the regression coefficient.

Table 4 shows the ANCOVA result on the post-test scores of the two groups, the means and standard deviations of the post-test scores, which were 69.68 and 30.00 for the experimental group and 56.61 and 34.39 for the control group. It was found that the post-test scores of two groups were significantly different, with $F = 20.27 (p < .001)$; implying that the students in the experimental group achieved a better learning performance in the knowledge of fractions after using MIT than those using a conventional web-based test system. Moreover, the adjusted mean of the experimental group’s post-test scores (71.55) was higher than that of the control group (54.74). The paired t-test result of the control group ($t = 1.126, p > .05$) indicates that the mean difference between the pre-test and post-test measures of the control group of students is not significantly different. Consequently, we conclude that the model-tracing intelligent tutor seems to be more effective than a conventional web-based test system in promoting the learning achievements of the students.

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted Mean</th>
<th>SE</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>62</td>
<td>69.68</td>
<td>30.00</td>
<td>71.55</td>
<td>2.64</td>
<td>20.27*</td>
</tr>
<tr>
<td>Control group</td>
<td>62</td>
<td>56.61</td>
<td>34.39</td>
<td>54.74</td>
<td>2.64</td>
<td></td>
</tr>
</tbody>
</table>

*Significantly different.

Discussions

The provision of feedback is believed to lead to better learning outcomes (Hattie & Timperley, 2007; Kleij, Eggen, Timmerse, & Veldkamp, 2012; Wu, Hwang, Milrad, Ke, & Huang, 2012). Timely feedback increases motivation and tends to motivate students to engage in learning (Kleij et al., 2012). Also, feedback can fill a gap between what is understood and what is aimed to be understood (Hattie & Timperley, 2007; Sadler, 1989). However, the existing web-based assessment systems mainly provide feedback based on multiple-choice question type or a restricted user interface. Through the proposed one-to-one tutoring mechanism, the instant feedback is provided to students. This system can discover students’ errors or misconceptions from their answers with model tracing approach, not only on the final results, but also on every working step.

As a result, the experimental results show that the model-tracing intelligent tutor helps students achieve significantly better effectiveness in improving students’ learning achievements than using the conventional web-based test system. That is, the proposed learning diagnosis system discovers students’ current knowledge and their weaknesses. Then students are able to make up lost ground and to improve their learning achievements. Namely, the proposed system, MIT, helps students to develop a deeper understanding of fractions. Besides, via this system, each student’s misconceptions are recorded; thus, teachers can easily understand students’ weaknesses and then help them to learn
better, faster, and more easily. In addition, it should be noted that the applications of MIT could be expanded to science, engineering, and mathematics courses that require solving equation problems.

Although this approach seems to be promising, it has some limitations in its practical application. For example, in order to trace students’ behaviors precisely and diagnose what kind of problems students have exactly, the question solving process need to be performed by typing answers into the text box step-by-step. Another limitation is that one misconception could be caused by dissimilar reasons (e.g., insufficient prerequisite knowledge). According to the researches (Chu et al., 2010; Hwang, 2003), there are prerequisite relationships among the concepts in a course. Since a lack of different prerequisite knowledge can lead to learning difficulties, future work of MIT will be directed to investigate how learning guidance could be provided for individual students by finding a set of relevant poorly learned concepts.

Conclusions

Manipulating fractions is considered one of the fundamental mathematical skills learned in elementary school. Therefore, learning about fractions has an important role in the course of children’s mathematical study (Armstrong & Larson, 1995; Barash & Klein, 1996; Behr, Harel, Post, & Lesh, 1992; Behr & Post, 1992; Behr, Wachsmuth, & Post, 1985; Hunting, 1983). However, fractions are difficult for elementary school students. Students often solve fraction problems by relying on the strategy of searching through their memories for a previously taught algorithm without understanding of the fundamental nature of fractions (Kerslake, 1986). Learning by rote memorization, however, may cause students to have misconceptions. Researchers indicated that finding students’ misconceptions or bugs is helpful for teachers and students (Brown & Burton, 1978). Schwarzenberger (1984) pointed out that mistakes help us to realize current learning situations. Also, mistakes can aid the process of mathematical discovery and assist mathematical understanding, and they can tell us more about what might be happening in a pupil’s mind than any number of correct answers (Schwarzenberger, 1984). In addition, Kerslake (1986) suggested immediate feedback could reduce students’ fear of fractions.

To assist learners in learning fractions, various cognitive tools have been devised to support, guide, and mediate the cognitive processes of learners, and meet the diverse needs of learners in comprehending procedural knowledge (Kong, 2008; Kong & Kwok, 2005). Lee and Bull (2008) introduced a learning environment with an open learner model, which models learners’ current understanding of the domain. Its aim was to help children understand their problems with fractions and help parents to help their children overcome misconceptions. To further diagnose students’ learning problems, researchers have demonstrated the benefits of applying learning diagnosis mechanisms in various courses (Hwang, 2003; Hwang, Tseng, & Hwang, 2008).

In this paper, an innovative learning diagnosis system, Model-tracing Intelligent Tutor (MIT), has been proposed to assist teachers to diagnose student learning problems. MIT is built on a context-free language parser generator YACC with input file (declarations, productions, and subroutines). By using MIT, students’ mathematical operations can be precisely traced step-by-step. When an item is incorrectly answered, MIT provides a timely diagnostic report. With the aid of this computer-assisted approach, teachers not only understand students’ learning status (diagnosis results) but also save time; therefore, teachers are able to assist weak pupils with remedial instructions.

This research adopts a quasi-experimental design to investigate the effectiveness of this system and a conventional web-based test. The research findings show that students using MIT achieve better learning than those using a conventional web-based test.

Acknowledgments

This work was partially supported by the National Science Council of the Republic of China under Grant No. NSC 98-2511-S-468-004-MY3, NSC 99-2221-E-009-130-MY2, NSC 100-2511-S-468 -002, and NSC 101-2511-S-468 -007 -MY3. The authors would like to thank Barbara Adamski for valuable comments to revise this paper.
References


Online Learning and Community Cohesion  
(Book Review)

Reviewer:  
Dermod Madden  
Associate-Superintendent, Aspen View School Division  
EDDE student, Athabasca University  
derm.madden@aspeview.org

Textbook Details:  
Online Learning and Community Cohesion  
Written by Roger Austin and Bill Hunter  
2013, 179 pages, Published by Routledge,  

“Online Learning and Community Cohesion,” is an innovative and comprehensive prospectus on the utility of Information and Communications Technology (ICT), and the extent to which it can reduce prejudice in public education, by establishing common social boundaries which bring students and teachers together for a common purpose. The book also explores and analyzes the ways in which ICT has been used in various countries around the world to promote citizenship, inclusion and community cohesion. The authors examine the theoretical frameworks of ICT initiatives in Northern Ireland, the Republic of Ireland, England, the European Union, Canada, and the United States. The rationale for the project is based on the premise that increasing globalization has resulted in the creation of greater numbers of heterogeneous groups, resulting in the potential for both rich cultural exchanges and a tendency toward tribalism. In the book the authors attempt to analyze the role of the school and the extent to which technology promotes cross-cultural interchange, and reduces prejudice by establishing common social boundaries. It is the underlying premise of the book that, “globalization has resulted in communities that are far more heterogeneous, and where intergroup contact has the potential to either foster rich cultural interchange or to provoke tension spilling over into violence” (p. 1).

The authors stress the significance of social identity and self-expression, and the importance of understanding the dynamic of a global group setting which by definition acknowledges diversity and cultural exchange. The use of ICT on a global level can promote online school projects which create deliberative environments that accommodate a global perspective. The authors point out that this has significance for pedagogy and the changing role of the teacher to accommodate twenty-first century learning practices. As such, the need for ongoing innovative teacher professional development should be deemed a priority.

The authors stipulate that meaningful online contact can be achieved using a combination of traditional teaching and learning models such as: a teaching model based on formal lessons utilizing e-books and other media; an informal contact model where students are brought together to share ideas, social activities and cultural exchange; a task model designed to work on joint projects and establish common perspectives; and a cluster model which incorporates elements of the aforementioned approaches. It is suggested that, notwithstanding the potential for meaningful global interaction in education, school divisions in cities and countries where social isolation based on race and ethnicity is the norm, need to change their practices from within before looking beyond their borders for global collaborative partners.

The expressed intent of the book was the examination of the use of ICT to foster community cohesion. Yet, according to the authors, even though all of the nations that were examined in this book have the human, economic and technological resources to support online global partnerships, very few if any programs or projects were designed to bridge societal divisions. In order to change from an exclusive to an inclusive global perspective there is a need for a greater role to be played by UNESCO to provide guidance and leadership in global educational issues. At the heart of the matter is the design of programs, “…that combine international links with connections between culturally (ethnically, racially, religiously) different local schools, might be a way to maximize the potential for this kind of learning and” (p. 144).

Instructional and program design have been identified as key to the success for online learning and community cohesion. Specifically the authors have identified four factors that will determine the success:
“the place of theoretical models in shaping the learning experience in interschool contact programs;
the role of teachers in shaping and delivering the actual experiences of the participating students;
the evolving dynamic of technology and the way that it may influence issues of program delivery; and
skillful program management to insure that quality experiences are provided for all involved” (p. 144).

The significance of this book for public education in the often termed, “developed world,” cannot be understated. The barriers that separate us identified as ethnic, political, cultural and religious have implications for educational leadership, instructional design, pedagogy and andragogy, assessment practices, teacher education programs, and most importantly, for students. If, as it is pointed out by the authors that online, ubiquitous educational access has the potential to support constructivist educational practice and inclusive universal engagement, then this book should be required reading in teacher education programs. Perhaps the relevance of this book lies in its message; that the significance of ICT and online learning for education is tantamount to a paradigm shift that will redefine public education, for the student, the teacher, the community and the world as we know it. How this will be managed is yet to be determined.

Online Learning and Community Cohesion is a thoughtful and articulate synopsis of the challenges we face in public education today and for the foreseeable future. Although the challenges are significant, the authors make a strong case that there is huge potential for improvements in public education and universal access to education through a reasoned approach to online learning, ICT and an inclusive perspective.