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 The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to Educational Technology & Society and three months thereafter.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:


Founding Editor
Kinshuk, Athabasca University, Canada.

Editors
Nian-Shing Chen, National Sun Yat-sen University, Taiwan; Demetrios G Sampson, Curtin University, Australia.

Advisory Board
Ignacio Aedo, Universidad Carlos III de Madrid, Spain; Mohamed Ally, Athabasca University, Canada; Luis Anido-Rifon, University of Vigo, Spain; Gautam Biswas, Vanderbilt University, USA; Rosa Maria Bottino, Consiglio Nazionale delle Ricerche, Italy; Mark Bullen, University of British Columbia, Canada; Tak-Wai Chan, National Central University, Taiwan; Kuo-En Chang, National Taiwan Normal University, Taiwan; Ni Chang, Indiana University 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; Athanasia 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; Rudy Lelouche, Universite Laval, Canada; Tzu-Chien Liu, National Central University, Taiwan; Rory McGrath, 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; Slavu 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; Ashok Patel, CAL Research & Software Engineering Centre, UK; Reinhard Oppermann, Fraunhofer Institut Angewandte Informationstechnik, Germany; Vladimir A Fomichev, K. E. Tsolokovsky Russian State Tech Univ, Russia; Olga S Fomichova, Studio “Culture, Ecology, and Foreign Languages,” Russia; Piet Koomers, University of Twente, The Netherlands; Chul-Hwan Lee, Inchon National University of Education, Korea; Brent Muirhead, University of Phoenix Online, USA; Erkki Sutinen, University of Joensuu, Finland; Vladimir Uskov, Bradley University, USA.

Editorial Assistant
Sie Wai (Sylvia) Chew, National Sun Yat-sen University, Taiwan; Stelios Sergis, University of Piraeus & Centre for Research and Technology Hellas, Greece; Wei-Chieh Fang, National Sun Yat-sen University, Taiwan.

Technical Manager
Panagiotis Zervas, University of Piraeus & Centre for Research and Technology Hellas, Greece

Executive Peer-Reviewers
http://www.ifets.info/
Supporting Organizations
Center for Cognition and Gesture-based Computing, National Sun Yat-sen University, Taiwan
Centre for Research and Technology Hellas, Greece

Subscription Prices and Ordering Information
For subscription information, please contact the editors at ets-editors@ifets.info.

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 ets-editors@ifets.info.

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.

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 ets-editors@ifets.info (Subject: Submission for Educational Technology & Society journal).
## Table of contents

### Guest Editorial

**Technology Enhanced Quality Education for All – Outcomes from EDUsummIT 2015**  
*Joke Voogt and Gerald Knezek*

1–4

### Special Issue Articles

**EDUsummIT: A Global Knowledge Building Community for Educational Researchers, Practitioners, and Policy Makers**  
*Kwok Wing Lai, Joke Voogt, Gerald Knezek and David Gibson*

5–15

**Advancing Mobile Learning in Formal and Informal Settings via Mobile App Technology: Where to From Here, and How?**  
*Ferial Khaddage, Wolfgang Müller and Kim Flinott*

16–26

**Infusing Creativity and Technology in 21st Century Education: A Systemic View for Change**  
*Danah Henriksen, Punya Mishra and Petra Fisser*

27–37

**Arguing for Computer Science in the School Curriculum**  
*Andrew Fluck, Mary Webb, Margaret Cox, Charoula Angeli, Joyce Malyn-Smith, Joke Voogt and Jason Zagami*

38–46

**A K-6 Computational Thinking Curriculum Framework: Implications for Teacher Knowledge**  
*Charoula Angeli, Joke Voogt, Andrew Fluck, Mary Webb, Margaret Cox, Joyce Malyn-Smith and Jason Zagami*

47–57

**Technology Enhanced Formative Assessment for 21st Century Learning**  
*J. Michael Spector, Dirk Iffenthaler, Demetrios Samspon, Lan (Joy) Yang, Evode Makama, Amali Warusavitura, Kalari Lokuge Dona, Koos Eichhorn, Andrew Fluck, Ronghui Huang, Susan Bridges, Jiingyan Lu, Youqun Ren, Xiaojing Gui, Christopher C. Deneen, Jonathan San Diego and David C. Gibson*

58–71

**A Multilevel System of Quality Technology-Enhanced Learning and Teaching Indicators**  
*Nancy Law, Dale S. Niederhauser, Rhonda Christensen and Linda Shear*

72–83

**Smart Partnerships to Increase Equity in Education**  
*Margaret Leahy, Niki Davis, Cathy Lewin, Amina Charania, Hasniza Nordin, Davor Orlić, Deirdre Butler and Olav Lopez-Fernández*

84–98

**A Smart Partnership: Integrating Educational Technology for Underserved Children in India**  
*Amina Charania and Niki Davis*

99–109

**Responding to Challenges in Teacher Professional Development for ICT-Integration in Education**  
*Jo Tondeur, Alona Forkosh-Baruch, Sarah Prestridge, Peter Albion and Shiyma Edirisinghe*

110–120

**Educational Digital Technologies in Developing Countries Challenge Third Party Providers**  
*Don Passey, Thérèse Lafertière, Manal Yazbak-Abu Ahmad, Miron Bhowmik, Diana Gross, Janet Price, Paul Resta and Miri Shonfeld*

121–133

### Full Length Articles

**The Impact of the Flipped Classroom on Mathematics Concept Learning in High School**  
*Kaushal Kumar Bhagat, Cheng-Nan Chang and Chun-Yen Chang*

134–142

**The Effect on Pupils’ Science Performance and Problem-Solving Ability through Lego: An Engineering Design-based Modeling Approach**  
*Yanyan Li, Zhinan Huang, Menglu Jiang and Ting-Wen Chang*

143–156
Improving Pupils’ Mathematical Communication Abilities through Computer-Supported Reciprocal Peer Tutoring
Euphony F. Y. Yang, Ben Chang, Hercy N. H. Cheng and Tak-Wai Chan

Surveying and Modeling Students’ Motivation and Learning Strategies for Mobile-Assisted Seamless Chinese Language Learning
Ching Sing Chai, Lung-Hsiang Wong and Ronnel B. King

Modelling Mathematics Teachers’ Intention to Use the Dynamic Geometry Environments in Macau: An SEM Approach
Mingming Zhou, Kan Kan Chan and Timothy Teo

Web-based Learning System for Developing and Assessing Clinical Diagnostic Skills for Dermatology Residency Program
Fan-Ray Kuo, Yi-Ying Chin, Chao-Hsien Lee, Yu-Hsien Chi, Chien-Hui Hong, Kuang-Lieh Lee, Wen-Hsien Ho and Chih-Hung Lee

Fostering Creativity in Tablet-Based Interactive Classrooms
Hye Jeong Kim, Ji Hyeon Park, Sungae Yoo and Hyeoncheol Kim

Effectiveness of a Learner-Directed Model for e-Learning
Stella Lee, Trevor Barker and Vivekanandan Suresh Kumar

Predictive Effects of the Quality of Online Peer-Feedback Provided and Received on Primary School Students’ Quality of Question-Generation
Fu-Yun Yu and Chun-Ping Wu

A Survey of Students’ Experiences on Collaborative Virtual Learning Activities Based on Five-Stage Model
M. Kemal Karaman and Sevil Orhan Özen

Effects of Sharing Clickers in an Active Learning Environment
Todd Daniel and Kristin Tivener

Measuring the Moral Reasoning Competencies of Service-Learning e-Tutors
Chih-Feng Chien, Ching-Jung Liao, Brent G Walters and Ching-Yieh Lee

Inquiry Based-Computational Experiment, Acquisition of Threshold Concepts and Argumentation in Science and Mathematics Education
Sarantos Psycharis

A Co-citation Network of Young Children’s Learning with Technology
Kai-Yu Tang, Ming-Chaun Li, Ching-Ting Hsien and Chin-Chung Tsai

A Decision Tool for Selecting a Sustainable Learning Technology Intervention
Maryam Raji and Imran Zualkernan

Online Metacognitive Strategies, Hypermedia Annotations, and Motivation on Hypertext Comprehension
Hui-Fang Shang

Mobile Seamless Technology Enhanced CSL Oral Communication
Yu-Ju Lan and Yen-Ting Lin

Effects of a Question Prompt-based Concept Mapping Approach on Students’ Learning Achievements, Attitudes and 5C Competences in Project-based Computer Course Activities
Hsia-Ying Wang, Iwen Huang and Gwo-Jen Hwang

ISSN 1436-4522 (online) and 1176-3647 (print). This article of the Journal of Educational Technology & Society is available under Creative Commons CC-BY-ND-NC 3.0 license (https://creativecommons.org/licenses/by-nc-nd/3.0/). For further queries, please contact Journal Editors at ets-editors@ifs.info.
Guest Editorial: Technology Enhanced Quality Education for All – Outcomes from EDUsummIT 2015

Joke Voogt\(^{1*}\) and Gerald Knezek\(^{2}\)

\(^{1}\)University of Amsterdam, Windesheim University of Applied Sciences, The Netherlands // \(^{2}\)University of North Texas, USA // j.m.voogt@uva.nl // gknezek@gmail.com

\(^{*}\)Corresponding author

ABSTRACT
There are ample studies that demonstrate how digital technologies can be used effectively to facilitate teaching and learning in the 21\(^{\text{st}}\) century. However, the insights gained from these studies often do not result in the uptake of technologies in educational practice (Voogt & Knezek, 2008). It is increasingly agreed upon that the implementation of digital technologies in education needs a systemic approach in which stakeholders at the micro, meso and macro levels of the education system actively interact with each other to align the needs of learners and the potential of technology with requirements of educational systems. EDUsummIT is a global knowledge building community of researchers, practitioners and policy makers working on this alignment with the aim to move education into the digital age. The EDUsummIT community meets every two years to discuss developments concerning technology and education from a systemic perspective with key stakeholders in the field. In this special issue the scholarly contributions of EDUsummIT 2015 are presented.

Keywords
EDUsummIT, Research-informed ICT integration strategies, Policy, Practice

Introduction
This special issue is the result of the fourth international EDUsummIT 2015, which was held from September 14-15, 2015 in Bangkok, Thailand. The EDUsummIT is an invitational summit focusing on the integration of Information and Communication Technology in education. Approximately 100 key stakeholders, policymakers, practitioners and researchers, from all over of the world were invited and discussed challenges and research-informed and practice-based strategies to effectively implement technology into teaching and learning. EDUsummIT 2015 closely collaborated with UNESCO Bangkok to determine the central theme for the summit, “Technology Enhanced Quality Learning for All.” EDUsummIT 2015 resulted in a Call to Action and policy briefs, in addition to the articles contained in this special issue of Educational Technology and Society. For further information about EDUsummIT, see http://www.curtin.edu.au/edusummit/.

EDUsummIT emerged from the International Handbook on Information Technology in Primary and Secondary Education (Voogt & Knezek, 2008). The handbook editors realized that much of the scholarly work done during forty years of research in the field of technology in education had not found its way into policy and practice. This was the initial reason for bringing researchers, policy makers and practitioners together in an international summit on technology implementation in education, which was held in 2009 in The Hague. This first summit led to EDUsummiTs in Paris (2011) (in close collaboration with UNESCO), Washington DC (2013) and Bangkok (2015). The fifth EDUsummIT is planned for September 18-19, 2017 in Borovets, Bulgaria. The EDUsummIT has been supported by international organizations such as UNESCO, the Society of Information Technology in Teacher Education, the International Society of Technology in Education, the International Federation for Information Processing (IFIP - working group named Research on Education and Applications of Information Technology), the Association of Teacher Educators, and UNESCO’s Teacher Development and Higher Education Division as well as Kennisnet (the Netherlands).

Prior to and during the summit the participants collaborated in teams of policy makers, practitioners and researchers in thematic working groups (TWGs) spanning nine themes: (1) smart partnerships; (2) advancing mobile learning in formal and informal settings; (3) professional development for policy makers, school leaders and teachers; (4) addressing gaps and promoting educational equity; (5) assessment as, for, and of learning in the 21st century; (6) creativity in a technology enhanced curriculum; (7) indicators of quality technology-enhanced teaching and learning; (8) digital citizenship and cyberwellness; (9) curriculum - advancing understanding of the roles of Computer Science /Informatics in the curriculum. Some of these themes evolved from earlier EDUsummiTs, other themes were new and suggested by the local host, UNESCO Bangkok. Within each theme actual developments and issues were explored.
discussed from system, school, classroom and student levels and their interconnections. In this special issue we present the scholarly work that resulted from EDUsummIT 2015.

Contributions to the special issue

It is increasingly agreed upon that the implementation of digital technologies in education needs a systemic approach in which stakeholders at the micro, meso and macro levels of the education system actively interact with each other to align the needs of learners and the potential of technology with requirements of educational systems. The papers in this special issue show the importance of these interactions from several perspectives. The special issue starts with a study about the EDUsummIT as a knowledge building community. This is followed by five contributions that discuss learning, curriculum and assessment in the 21st century. The subsequent five contributions elaborate upon the implementation of technology at the teacher, school and system level.

The EDUsummIT as knowledge building community

In the first paper of this the special issue Kwok-Wing Lai and colleagues discuss the EDUsummIT as a unique global knowledge building community. Data were collected from the EDUsummIT 2015 participants about the effectiveness and impact of the EDUsummIT. Based on theoretical notions about knowledge building communities (Scardamalia & Bereiter, 2003) and features of successful research practice relationships (Pieters & de Vries, 2007) the opportunities and potential threats to the EDUsummIT as a knowledge building community of researchers, practitioners and researchers were examined. This study led to recommendations strengthening the EDUsummIT and provided suggestions for the design of other internationally oriented knowledge building communities. The authors concluded that successful international knowledge building communities that aim to bring together researchers, practitioners and policy makers to solve educational problems should be designed with the following characteristics in mind: commitment to a shared goal; agency of all participants; enough opportunities for (online and face-to-face) discussions; fast disseminations of the knowledge produced; and a structure to warrant sustainability.

Learning, curriculum and assessment in the 21st century

The second paper begins from the perspective of the learners in today’s society. Khaddage, Müller and Flintoff argue that mobile learning is not so much about mobile devices but about the mobility of the learner. This implies that learners can be engaged in activities that make explicit use of the location, allowing for informal learning. A major challenge for educators in mobile learning contexts lies in bridging the gap between these informal learning experiences with the expected outcomes of formal learning as required by educational institutions. The paper starts with briefly reviewing a framework that was developed during EDUsummIT 2013, which addressed the pedagogical, technological, policy and research challenges of mobile learning (Khaddage et al., 2015) Potential solutions for these challenges are discussed. Technological developments result in mobile apps that allow for motivational design, “quiet” design and playful interfaces which make new approaches to bridging informal and formal learning settings feasible. To illustrate how mobile apps technologies can be used to connect formal and informal learning at the school level, an approach using mobile apps technologies is elaborated for the STEAM subjects.

The next three paper contributions discuss important elements of a 21st century curriculum. The contribution of Henriksen, Mishra and Fisser explores creativity as a core construct for 21st century curricula. The contributions of Fluck and colleagues and Angeli and colleagues address the role of Computer Science as a new subject in the curriculum. Based on previous research, Henriksen, Mishra and Fisser, adopt three components of creativity as a product and/or process: novel, effective and wholeness. Creativity is novel and effective when it provides useful solutions to problems (Sternberg, 2006). Wholeness refers to the esthetic component of creativity (Mishra & Koehler, 2008). It is argued that creativity is not an individual endeavor, but needs to be approached from a systems points of view: the person, the domain and the field and their interactions (Csikszentmihalyi, 1997). By making use of the affordances of technology to support teaching and learning of content domains teachers can develop a creative mindset contributing to good teaching. Recommendations for teacher education, assessment and educational policy are elaborated as means toward realizing the effective integration of creativity and technology in 21st century education.
Implementation of technology at the school and system level

Monitoring the quality of technology-enhanced learning and teaching implementations and outcomes require new indicators according to Law and colleagues. They propose a multilevel system framework of indicators at the student, school, and system level. Examples of indicators are provided. The framework emphasizes the linkages and interactions between indicators and interdependencies within and across the student, classroom, school, and system level. The authors advocate the need for longitudinal, multilevel research designs to contribute to in-depth insights of the links between intentions, implementation, and outcomes.

In the next paper by Leahy and colleagues the concept of Smart Partnerships in education is introduced. In today’s discourse “smart” in the context of education refers to the use of technology to shape teaching and learning based on the needs of the learner (Kinshuk, Chen, Cheng & Chew, 2016). Technology both creates big data and can make big data accessible to empower teachers and learners to improve education. The contribution of Spector and colleagues in this issue reflects on the possibilities of big data for formative assessment purposes. The authors argue that education can benefit from “smart” uses of technology for teaching and learning if they partner with others inside and outside the school to foster “the free sharing and exchange of knowledge and ideas to the benefit of [all] parties” (Falloon, 2015, p. 216). Smart partnerships include partners within and across education that have a shared purpose and a strategic and holistic approach, and facilitate their organizations to change. Smart partnerships aims to deploy technologies to enhance the quality of education, harness technology smartly and recognize the role of technologies in emergent processes. The assumption is that by partnering with multi-stakeholders the technological tools can be used to enhance the quality of educational outcomes. The Smart Partnership of India serves more than seventeen thousand students in India, including students from remote and deprived populations.
The development and outcomes of this Smart Partnership are analyzed using the characteristics of Smart Partnerships introduced in the contribution of Leahy and colleagues in this issue.

While the previous two papers address the implementation of technology at the system level, Tondeur and colleagues take the teacher’s perspective. Because teachers are core to the adequate implementation of digital technologies in teaching and learning, this contribution focuses on the importance of teacher professional development for technology integration. Five challenges for teacher professional development are presented: (1) contextualization: sociocultural awareness; (2) sustainability and scalability of professional development; (3) empowering pedagogy through ICT; (4) technology discernment and (5) professional development that is systemic and systematic. Exemplary teacher professional development cases from Kenya, Australia, Israel and Sri Lanka that successfully coped with these challenges are described, resulting in a model for teacher professional development on technology integration through teacher inquiry learning.

In the final contribution Passey and colleagues discuss technology planning and implementation issues of NGOs and governmental organizations in developing countries. Based on an analysis of research on technology implementation in developed countries possible explanations for the many failures (and some successes) are derived. New models for technology implementation that cater to long-term and sustainable implementation of technology in developing countries are proposed.

References


EDUsummIT: A Global Knowledge Building Community for Educational Researchers, Practitioners, and Policy Makers

Kwok-Wing Lai, Joke Voogt, Gerald Knezek, and David Gibson

1University of Otago, New Zealand // 2University of Amsterdam, The Netherlands & Windesheim University of Applied Sciences, The Netherlands // 3University of North Texas, USA // 4Curtin University, Australia // wing.lai@otago.ac.nz // j.m.voogt@uva.nl // gknezek@gmail.com // david.c.gibson@curtin.edu.au

Corresponding author

ABSTRACT

The International Summit on Information and Communication Technology (ICT) in Education (EDUsummIT) is a global knowledge building community of researchers, educational practitioners, and policy makers aiming to create and disseminate ideas and knowledge to promote the integration of ICT in education. Four EDUsummITs have been convened in The Hague, Paris, Washington, D.C., and Bangkok resulting in the development of ideas and knowledge useful and usable in educational practices worldwide. More than two-dozen presentations by EDUsummIT participants have advocated implementation of recommendations from this knowledge building community via national and international organizations such as UNESCO, IFIP, Kennisnet, ISTE, and SITE. This paper reports findings of an evaluative study conducted at the conclusion of EDUsummIT 2015 to investigate the effectiveness and impact of EDUsummIT. Evaluative findings are proposed by the authors as useful for continuous improvement of future EDUsummITs while also serving as guiding considerations for emerging communities of practice in other fields wishing to develop and disseminate recommendations for research-based best practices at the global level.

Keywords

Knowledge building community, Research practice relationships, International, EDUsummIT

Introduction

The International Summit on Information and Communication Technology (ICT) in Education (EDUsummIT) is a global knowledge building community of researchers, educational practitioners, and policy makers, aiming to create and disseminate ideas and knowledge “based on a shared understanding between researchers, policy makers and practitioners…to promote the integration of information and communication technology in education” (Voogt, Knezek, & Pareja Roblin, 2015, p. 620). EDUsummIT was founded in 2009, initially to extend and further develop the work undertaken by the authors of the International Handbook of Information Technology in Primary and Secondary Education, edited by Joke Voogt and Gerald Knezek (2008). In the first EDUsummIT the participants were closely related to the Handbook, but in the following EDUsummITs new participants who were active researchers, leading policy makers and practitioners of major educational organisations or institutions were invited by the steering committee to participate. In fact, in the most recent EDUsummIT, only about 10% of its participants were Handbook authors. Since its inception, EDUsummIT has been held four times, firstly in Hague (2009), then Paris (2011), Washington D.C. (2013), and most recently, in Bangkok (2015). As an invitational meeting, between 70 and 140 participants attended each EDUsummIT meeting.

While EDUsummIT participants meet biennially, thematic working groups (TWG) focusing on pertinent research topics in ICT and education are formed prior to the Summit to conduct research and prepare discussion papers. These papers are further developed during the Summit. The themes of the TWGs are decided by the steering committee and TWG leaders are identified. Once the TWGs are formed, specific topics of study of each TWG are negotiated between its leaders and members. There was a strong continuity of the themes of study, with only one or two themes being added or replaced in each EDUsummIT. For example, in EDUsummIT 2015, there were nine TWGs, and only two were new groups. After each EDUsummIT, TWG findings are published as research papers in international and national journals and presented at major conferences (Lai, 2015, also refer EDUsummIT website, http://www.curtin.edu.au/edusummit/index.cfm for further information). Action plans are also published in the EDUsummIT website (http://www.curtin.edu.au/edusummit/) to facilitate rapid dissemination of new ideas and knowledge created at the Summit. While EDUsummIT is an independent organisation, in the last several years it has been working closely with major international and national organisations which promote the use of information technology in education, including the Society for Information Technology and Teacher Education (SITE), the
EDUsummIT is a unique model of a knowledge building community. It connects researchers with policy makers and practitioners at the international level to co-create knowledge. EDUsummIT is not a conventional conference where the presenters disseminate knowledge and the audiences consume knowledge, but a collaborative community of practice, and so far its knowledge building process has gone through four cycles of development. This article reports findings from an evaluation conducted at the conclusion of EDUsummIT 2015 to investigate the effectiveness of EDUsummIT as a knowledge building community, as well as its impacts on local and international research, policy making, and practitioner communities. In reporting EDUsummIT as a model of how a knowledge building community could be developed and sustained, the findings of this evaluation would provide insights for designing similar knowledge building communities aiming at connecting research and practice in other fields. Theoretical underpinnings guiding the EDUsummIT endeavour will be provided in the next section.

Theoretical underpinnings

Concerns about the use of research generated knowledge for educational practice are not new and are not only voiced by researchers, but also by practitioners and policy makers (e.g., Broekkamp & van Hout Wolters, 2007; Levin, 2004; Vanderlinde & van Braak, 2010). The diversity and complexity of the contexts in which education takes place makes it difficult to make ecological generalizable statements that are only informed by the results of research. Although conferences are often thought to contribute to knowledge sharing and network building between researchers and practitioners, it is clearly shown in the literature that traditional role attitudes of the participants - researchers disseminate research results, which are consumed by intermediaries and practitioners - hinder the exchange of knowledge and experience and discourage co-creation of new knowledge and common agenda setting (De Vries & Pieters, 2007a).

A fundamental problem is that researchers and practitioners are interested in different types of knowledge. Researchers aims at contributing to generalizable statements that contribute to theory building, while knowledge generated by practitioners is often context specific and implicit. These two types of knowledge are referred to as Mode 1 and Mode 2 knowledge (Gibbons, 2000). Mode 1 knowledge is mainly generated by research done in universities or research institutions. Problems that are studied by researchers are derived from theory and accounted for within the research community. The dissemination and use of knowledge is hierarchical and linear. The production of Mode 2 knowledge, however, is not hierarchical, but based on problems defined and solved in the context of practice. This knowledge is practical in nature and context specific (Hiebert, Gallimore, & Stigler, 2002). The dissimilarities between how these two types of knowledge are generated and used lead to different perceptions on the relevance and usefulness of the knowledge that is created respectively in research and practice. As such, researchers need to acknowledge that while educational research can provide relevant knowledge useful in solving educational problems, educational practice cannot be based solely on research findings (Pieters & Jochems, 2003). A mechanistic conception of educational change where innovation is directed from “knowing based on research” to “implementing in practice” has been shown to fail (Burkhardt & Schoenfield, 2003; Vanderlinde & van Braak, 2010).

To overcome the fundamental difference between researchers and practitioners in conceptualising how knowledge is developed and used, several scholars advocate forms of collaboration based on equivalence of expertise (Pieters & De Vries, 2007). It has been suggested that new forms of collaboration and interaction are needed in knowledge creation, with a shared responsibility between practitioners, policy makers and researchers, and explicit and implicit questions and insights play a role in the process of co-creation of knowledge. For productive forms of knowledge building Scardamalia and Bereiter (2003) used the term knowledge building community as “the production and continual improvement of ideas of value to a community, through means that increase the likelihood that what the
community accomplishes will be greater than the sum of individual contributions and parts of broader cultural effort” (p. 1370). Based on De Vries and Peters (2007b), Scardamalia and Bereiter (2003), and Wenger (1998), EDUsummIT as a knowledge building community is framed by the following principles: (a) heterogeneity – the composition of the community is diverse and equivalence of expertise is respected; (b) informality – a shared goal is informally developed and agreed upon; (c) interactivity – members fully engage in dialogues and discussions; and (d) effectiveness – outcomes relevant to all members.

EDUsummIT 2015 evaluation

In the first three EDUsummITs, while the TWGs were formed prior to the Summit, pre-Summit work was primarily undertaken by TWG leaders. In 2015, however, all TWG leaders and members worked together online to prepare discussion papers and develop resources prior to the physical meeting, and some of the TWG members chose to participate in the virtual pre-Summit discussion only but not attending the meeting in person. Thus, in the 2015 EDUsummIT, there were two types of participants, the attending and non-attending. With a total of 132 participants in EDUsummIT 2015, 89 were attending and 43 were non-attending participants from over 25 countries. As can be seen from Table 1, of those who attended in person, the majority of them (51%) were from Europe and two Pacific nations (Australia and New Zealand), followed by North American (15%), and Asia (13%).

Table 1. Participants of EDUsummIT 2015 by region

<table>
<thead>
<tr>
<th>Region</th>
<th>Attending (%)</th>
<th>Non-attending (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>25 (28.4%)</td>
<td>6 (14.0%)</td>
<td>31 (23.7%)</td>
</tr>
<tr>
<td>Pacific</td>
<td>20 (22.7%)</td>
<td>9 (20.9%)</td>
<td>29 (22.1%)</td>
</tr>
<tr>
<td>North America</td>
<td>13 (14.7%)</td>
<td>6 (14.0%)</td>
<td>19 (14.5%)</td>
</tr>
<tr>
<td>Asia</td>
<td>12 (12.4%)</td>
<td>15 (34.9%)</td>
<td>27 (20.5%)</td>
</tr>
<tr>
<td>Middle East</td>
<td>3 (3.4%)</td>
<td>2 (4.7%)</td>
<td>5 (3.8%)</td>
</tr>
<tr>
<td>South America</td>
<td>3 (3.4%)</td>
<td>1 (2.3%)</td>
<td>4 (3.0%)</td>
</tr>
<tr>
<td>Africa</td>
<td>0 (0.0%)</td>
<td>3 (7.0%)</td>
<td>3 (2.3%)</td>
</tr>
<tr>
<td>International organisation</td>
<td>12 (13.6%)</td>
<td>1 (2.3%)</td>
<td>13 (9.9%)</td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td>43</td>
<td>132</td>
</tr>
</tbody>
</table>

Table 2 summarises the distribution of EDUsummIT 2015 participants by profession. It should be noted that the participants only indicated their primary profession in the questionnaire and we acknowledge that some EDUsummIT policy makers and practitioners would also undertake research, and many researchers also practiced as tertiary teachers. As can be seen from Table 2, the vast majority of the participants were researchers, and almost one-third of all the participants were returning participants (having participated in at least one previous EDUsummIT).

Table 2. EDUsummIT participants by profession

<table>
<thead>
<tr>
<th>Profession</th>
<th>Attending (%)</th>
<th>Non-attending (%)</th>
<th>Total (%)</th>
<th>Returning participants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher</td>
<td>68 (77.3%)</td>
<td>39 (90.7%)</td>
<td>107 (81.7%)</td>
<td>37 (34.6%)</td>
</tr>
<tr>
<td>Policy maker</td>
<td>14 (15.7%)</td>
<td>2 (4.7%)</td>
<td>16 (12.1%)</td>
<td>3 (18.8%)</td>
</tr>
<tr>
<td>Practitioner</td>
<td>4 (4.5%)</td>
<td>2 (4.7%)</td>
<td>6 (4.6%)</td>
<td>2 (33.3%)</td>
</tr>
<tr>
<td>Commercial &amp; Consultancy</td>
<td>3 (3.4%)</td>
<td>0 (0.0%)</td>
<td>3 (2.3%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td>43</td>
<td>132</td>
<td>42 (31.8%)</td>
</tr>
</tbody>
</table>

In EDUsummIT 2015 there were nine thematic working groups:

TWG1: Smart partnerships
TWG2: Advancing mobile learning in formal and informal settings
TWG3: Professional development for policy makers, school leaders and teachers
TWG4: Addressing gaps and promoting educational equity
TWG5: Assessment as, for, and of learning in the 21st century
TWG6: Creativity in a technology enhanced curriculum
TWG7: Indicators of quality technology-enhanced teaching and learning
TWG8: Digital citizenship and cyberwellness
TWG9: Curriculum - advancing understanding of the roles of CS/Informatics in the curriculum
In the previous EDUsummITs, participants were assigned to TWGs but in 2015 they were free to choose which group they would like to belong to. As a result, there was an uneven distribution of group membership, as can be seen from Table 3.

<table>
<thead>
<tr>
<th>TWG</th>
<th>Attending</th>
<th>Non-attending</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Organisers</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td>43</td>
<td>132</td>
</tr>
</tbody>
</table>

The online survey

After the conclusion of EDUsummIT 2015 all the attending and non-attending participants were asked to complete an online questionnaire consisting of open-ended and closed items. The questionnaire was developed by the authors and pilot tested by five researchers. The questionnaire was revised after the pilot study. Sixty-two (62) participants responded to the questionnaire, and 45 of them had attended EDUsummIT 2015 in person, accounting for 51% of all the attending participants. 17 respondents (38%) were leaders. Since several questions in the questionnaire were not relevant to the participants who did not attend EDUsummIT 2015 in person, and therefore there was a large amount of missing data in these items. Thus, the non-attending responses were not included in the analysis. Of the 45 attending respondents, 40% were male, 60% were female. In terms of profession, 70.4% of the respondents were researchers, 4.5% policy makers, 18.2% practitioners, and 6.8% had not specified their professions. The distribution of the attending respondents by profession was similar to the overall EDUsummIT 2015 population (refer Table 2). In the following sections responses to the key questions asked in the questionnaire are synthesised.

What motivated people to take part in EDUsummIT 2015?

While undoubtedly there were multiple reasons for an invitee to decide whether or not they took part in EDUsummIT 2015, nonetheless from the responses three dominant themes have been identified. One theme that stood out was the desire to be part of a knowledge building community, “to contribute to something very worthwhile,” and “to give back to society.” These respondents had a strong desire to work with other experts in the field to initiate change. They were also keen to work together to “renew working relationships,” so as to engage in something which is useful. The following comment is an example from this category of responses.

[My motivation is] the need to be a part of an expert community that exchanges ideas and aims to influence research, practice and policy-making worldwide, as a means of empowering countries as a whole and education systems in a technology-saturated world. In addition, there are joint topics that need further elaboration - this can be done only via an international summit, which combined research agenda with policy, thereby enhancing the practice… (R13, Researcher).

The second theme identified from the responses, which was closely related to the first theme, was the desire for “international networking,” to gain an “international perspective,” and “to meet new people,” in order to enhance personal growth and development, as evident from the following comments:

The most valuable part about EDUsummIT for me is - having the time to work on a specific topic - taking the time to WORK on a specific topic (not listening to presentations but actually work together) - having the opportunity to
discuss and reflect with other policy makers and researchers on the topic and get inspired by this - use this inspiration for my own work… (R30, Policy Maker)

I enjoy working with scholars from other countries to gain a more international perspective. I personally learn a great deal from other group members… (R24, Researcher)

Most participants were motivated to take part in EDUsummIT 2015 for both reasons: contributing their expertise to a worthwhile cause and using EDUsummIT as a professional learning and development opportunity. As explained by the following respondent:

[I see] the needs of UNESCO and those it represents. [EDUsummIT] also has a reputational and stimulating professional development benefit for us all… (R38, Researcher)

The third theme was identified from the responses primarily contributed by returning participants, those who have participated in EDUsummIT previously. Several respondents mentioned that they participated in EDUsummITs in the past and thus had developed a sense of belonging to this community, so they just kept coming back (“the previous one I attended was very enlightening”). The following comment captured the personal and community needs which the returning participants found EDUsummIT could help fulfill:

I started as a participant in Paris…It is a great opportunity to meet like-minded scholars and to produce new knowledge. Just personally this had led to 3 papers that have been published as well as conference presentations and hopefully more of the same from this conference as well (R23, Researcher).

From the questionnaire responses, it is clear that the majority of the respondents had a strong sense of commitment to the EDUsummIT community. When asked whether they would attend the next Summit (to be held in Borovets, Bulgaria, September 18-19, 2017), 84% gave a positive response.

**EDUsummIT as a knowledge building community**

In addition to the research papers that the TWGs were charged to develop at EDUsummIT 2015, they also worked closely with UNESCO Bangkok specialists to develop policy briefs to be distributed and used at the Asia Pacific Ministerial Forum of ICT in Education (AMFIE) in 2016. Thus, the theme of EDUsummIT 2015, *Technology Advanced Quality Learning For All*, had a special focus on the integration of digital technologies in education in Asia-Pacific countries. The TWG research briefs include:

- Defining the research, policy, and practice challenges of the theme of study
- Synthesising relevant research
- Discussing and confirming examples of innovative practices, including Asian-Pacific examples
- Discussing implications and recommendations for researchers, practitioners, and policy makers

As mentioned previously, EDUsummIT 2015 was the first Summit where both leaders and members spent a lengthy period of time prior to the Summit to develop their topic of study. It would be interesting to see how this online component would enhance the knowledge building process. One indicator to show whether a community is effective is how active its members are in contributing ideas to the community (Lai, 2015). From the questionnaire responses it was found that participation was less active in the pre-Summit phase than during the Summit phase, with only 44% of the respondents indicated that they had participated very actively or actively during this phase (but 79% during the Summit). However, almost three-quarters (72%) of the respondents considered their TWGs functioned effectively as a knowledge building community prior to the Summit, and the percentage of positive responses increased to 86% during the Summit phase.

In the open-ended responses, the key issue that has been expressed by the respondents during the pre-Summit phase was the lack of active participation in some TWGs (“Members were busy and many were slow to engage in the drafting and discussion process online” (R33, Researcher)) and thus the lack of effective co-operation. However, in other groups, there was “good participation from group members” (R19, Researcher), and “contributions…were very
busy prior to the meeting” (R21, Researcher). The importance of good organisation and preparation by the TWG leaders was a key factor affecting how the pre-Summit discussion worked. As pointed out by a respondent:

[My] TWG functions effectively because the leader gave everybody the opportunity to participate. (R9, Researcher)

The responses on the effectiveness of the TWGs as a knowledge building community during EDUsumIT 2015 were much more positive. There was a general consensus that the two full days were used productively (and a few considered it too short). “The discourse was excellent…and [it] provided new insights…and possible solutions” (R16, Researcher), “the knowledge building and discussion very inspiring and effective” (R18, Practitioner) and by and large the TWGs were run effectively, and the discussions were mostly based on the work the participants did before the Summit. While the online participation has imposed more demand on the participants, some respondents considered it worth the effort:

Working before and during the EDUsumIT is very time consuming, but very effective: it allows you to work on a specific subject for several days with others that are similarly interested in the topic. It gives you “time to think”! (R30, Policy Maker)

While the discussions during EDUsumIT 2015 were intense throughout the two-day period, most TWGs managed to develop a friendly and open atmosphere, as evident from the following comments:

Our group was characterized by a strong feeling that anything you say and any idea you bring to the table is welcome and even desired. As a result, quite open and frank discussions took place, and this was very fruitful and evident in the ideas presented in the plenary session, and before that – in the cross fertilization session. The feeling of unfamiliarity quickly changed in our group to a feeling of comfort and easiness, creating the basis for upgraded outputs based on those prepared before the summit. (R13, Researcher)

In the TWG there was an open and collaborative atmosphere where everybody can comment ideas discussed in the TWG from his/her professional background. All ideas were accepted and developed…Each voice was taken seriously. (R15, Practitioner)

Again, the importance of good organisation and leadership was highlighted:

I think our group in Bangkok was incredibly productive and effective. People came willing to share, discuss and learn. The three leaders had met previously and planned the structure well – AND (most importantly) [they] were flexible in switching plans when needed – depending on input from the other participants. (R23, Researcher)

The participation of the UNESCO specialists was helpful:

Work done prior to [the] event served as [a] very effective platform on which to build experience of two group members of [our group] in Asia Pacific proved useful backdrop to mine… (R6, Practitioner)

Taking both the pre-Summit and Summit phases into account, the following comment reflects accurately the general feeling of the community.

The fact that the participants were mostly new was a barrier to their active participation in most cases; however, those that contributed were not necessarily those that came to the summit. Hence, a real community was developed not based on physical presence, but rather on the feeling of shared interests and a sense of responsibility to the topic of the TWG. The final results of the pre-summit work was quite comprehensive and indicated that although not all group members, for reasons of lack of time or not being familiar with other group members, made their contribution – still, the result was satisfying, as well as the process itself. (R13, Researcher)

Indeed, when asked in the questionnaire whether their expectation was met, 89% of the respondents considered it fully met or almost fully met.
Has new knowledge been created?

There were 40 responses to the question of whether or not new knowledge has been created in EDUsummIT 2015. The majority of these responses (63%) were positive, affirming that new knowledge has been created in their TWGs, as evident from the following two comments:

Yes, I think new knowledge has been created by our group. The expansion of the database of creativity instruments is ongoing. The three papers we seek to write will have insights about teaching and teacher education, assessment and policy. More importantly the meeting has set the foundation for some good work that will emerge over time. (R23, Researcher)

Yes, for sure. We introduced an important concept of “technology discernment” which we think could become the main pillar to ensure effectiveness and sustainability in PD in technology. (R27, profession not specified)

In EDUsummIT 2015 a new TWG with a new theme (Curriculum – advancing understanding of the roles of CS/Informatics in the curriculum) was introduced, as a response to a new development in ICT in education. A member of this TWG confirmed that “regarding TWG 9 I would agree that new knowledge was created, because it was a new theme at EDUsummIT 2015, and just in time for many countries.” (R16, Researcher)

The following comment shows how knowledge has been developed from the online discussion during the pre-Summit phase, and then further developed during the face-to-face meeting.

The contribution of the discussions of the TWG were manifold in the sense of new knowledge creation and integration: - the contribution of the UNESCO representative was a more enhanced focus on policy, especially issues related to the Asia-Pacific area; - discussions assisted in the notion of the need to elaborate a relevant model for our topic, that was developed in the prior summit; - a clearer notion of the connection between research, practice and policy was established; - models for our topic from different countries were presented and we exchanged information and established ideas for a more generic model. (R13, Researcher)

What was the impact of EDUsummIT on practice and policy making?

While in the closed items almost three-quarters of the respondents were positive about the potential impact of EDUsummIT on educational research (73%), policy-making (73%), and educational practice (71%), in the open-ended question when the respondents were asked to elaborate on their responses, only 41% (12 out of 29) of the respondents thought that EDUsummIT 2015 would have an impact on policy making.

Most respondents thought that EDUsummIT 2015 would have an impact on educational research but they were not sure about its impact on policy making, as “this highly depends on the individual participant of EDUsummIT” (R30, Policy Maker). However, since EDUsummIT 2015 worked closely with UNESCO Bangkok, and the policy briefs developed by the TWGs would be distributed during AMFIE that “we have more chance than ever, and that there is more need as well. Collaboration with UNESCO Bangkok will be a key partnership that can make this happen more effectively” (R38, Researcher). The collaboration with UNESCO Bangkok in EDUsummIT assured that there would be impact on the policy side. To increase impact, support from international organisations is essential, as commented by the following respondent:

The EDUsummIT needs to be made more public and well-known, include more participants and work in collaboration with UNESCO and possibly more international organization. Currently, some countries are influenced dramatically by EDUsummIT’s outcomes, and some to a lesser degree – this needs to be re-examined and perhaps studied in a more cultural context (R13, Researcher).

Future of EDUsummIT as a knowledge building community

How EDUsummIT could be further developed and strengthened in order to increase its impact and sustain its development was the major question in the minds of the EDUsummIT participants. From the questionnaire responses
we have identified three ideas which would need consideration. They are the need to: (1) maintain a clear focus; (2) broaden participation; and (3) develop a permanent structure.

**Maintaining a clear focus**

The importance of maintaining a clear focus for EDUsummIT was raised by several respondents in the evaluation. Participants need to understand clearly the purpose of EDUsummIT, that is, a knowledge creating community aiming to create and recommend solutions to policy makers and educational practitioners of how ICT should be integrated into teaching and learning. To achieve this goal, and to increase EDUsummIT’s impact, several respondents expressed the need to consider carefully the composition of the TWGs, in terms of how leaders were selected and members invited. To create knowledge effectively as a community, a strong commitment to the cause of EDUsummIT is essential and participants also need to have a good understanding of the research or practice issues related to their themes of study, as explained by one respondent:

[We need] more commitment from folks who sign up to be part of EDUsummIT… I think it is great to be inclusive but I think it creates difficulty in who is committed and who has enough background and experience to contribute to the EDUsummIT (R24, Researcher).

If commitment is essential, there is a question of whether or not two tiers of participation (attending and non-attending) should be continued in the future. R24 expressed her opinion further:

There were a number of people assigned to the group and it was unclear who intended to participate in EDUsummIT and who did not. I think it should be an exception that people sign up to be in a working group and are not in attendance. It is difficult to communicate with a large number of folks who are unfamiliar and you will never meet (R24, Researcher).

The issue of whether the two tiers of participation should be continued in the future, and if it is continued, how it is managed to benefit both the attending and non-attending participants should be considered carefully. The potential EDUsummIT participants are ICT experts; it is thus natural to engage them in virtual participation and discussions. However, the time and effort required to organise and manage both the online and on-site communities may outweigh its benefits.

**Broadening participating**

As a knowledge building community, it is important that EDUsummIT participants know each other, and as such there is a need for continuity in TWG leadership and membership in order to facilitate development of ideas. However, new participants should also be recruited to avoid EDUsummIT becoming an insular group, as commented by the following respondent:

I thought [EDUsummIT] was quite insular and could be opened up to more people contributing to initial ideas through the use of technologies and these ideas refined during by the working groups. (R41, profession not specified)

In EDUsummIT 2015, only a small proportion of the participants were policy makers and practitioners. The questionnaire respondents clearly expressed the need for broadening participation, as explained by one respondent:

EDUsummIT needs its own infusion of broad-based support partnerships that bring in the commercial/governmental business elements… UNESCO provided a much needed balance between academics and policy makers but there was still an obvious hole. More educators should be offered the chance to interface with those outside academia if they are going to increasingly be expected to incorporate their needs into research and planning programs. (R8, Practitioner)

Also, as a global community, EDUsummIT should have a more significant representation from less economically advanced countries:
It is important that (1) members of this community origin from all over the world. (2) [in] the community there are experts from different branches related to education. (3) the EDUsummIT community is opened [sic]. These principles would be important to apply also for the future. (R15, Practitioner)

This is challenging, as EDUsummIT has not been able to secure long-term sponsorships to support travel for increasing participation from developing nations.

*Developing a permanent structure*

Several respondents raised the need to develop a permanent organisational structure for EDUsummIT. As pointed out by a leader respondent, this would increase the impact of EDUsummIT on policy making.

The community is too small with too few high level decision makers and policy leaders involved. It is a diverse and interesting community with many good ideas who enjoy talking and listening to each other. There is no permanent structure for the community, and things are decided on an ad-hoc basis. To have an impact, there needs to be a permanent organizational structure with a transparent decision making body with some sponsorship from prestigious bodies and agencies. The goals and ambitions are great – realizing them even to a small degree remains a serious challenge (R2, Researcher).

With a more permanent structure, it may be easier to get funding support from international organisations, since “expecting so many to come at their own expense is expecting quite a lot…that practice is simply not sustainable” (R2, Researcher), and funding support will also allow participants from economically less advanced countries to participate.

However, some EDUsummIT participants may not welcome such a move, since a loosely structured organisation may be more conducive to creating knowledge, as explained by the following respondent.

It seems clear from the plenary discussions about creating a management structure that currently EDUsummIT is more of a network than a professional organisation. So it operates on shared understandings rather than explicit guidelines. While guidelines might give direction to those responsible for planning TWGs they might also be resisted by those who prefer just to meet and see what emerges when interested professionals work together. (R4, profession not specified)

The following comment seems to provide a good summary of the key issues that EDUsummIT should consider in the future:

A careful nomination process for invitations for the EDUsummIT. Putting in an effort to get more representatives from developing countries. A good balance between veterans and newcomers. Having good TWG leaders that preferably attended earlier EDUsummiTs. Making sure that we address new topics; leave room for innovations in the topics we address. Be action/output oriented (R16, Researcher).

*Discussion and conclusion*

Integrating knowledge generated from researchers and practitioners is a demanding task and EDUsummIT was established to tackle the task of developing understanding of how research and practice knowledge could be integrated in the field of ICT in education. In this section we use the four principles for developing effective knowledge building communities as discussed in the “theoretical underpinnings” section to discuss the implications of our findings (Wenger, 1998; De Vries & Pieters, 2007b). First, this study showed that EDUsummIT was highly interactive. EDUsummIT was a community where knowledge was created through an open and ongoing dialogue among researchers, policy makers and practitioners. Knowledge creation happened both during the online discussions and in the face-to-face meeting where participants discussed problems and offered solutions of what they experienced in the implementation of information and communication technology in practice. This discourse reflected Mode 2 knowledge more than Mode 1 knowledge (Gibbons, 2000). Second, the findings of the study showed that participants had three major reasons for participating in EDUsummIT. They wanted to be part of a collaborative effort to help change education given the
technological developments that impacted on our society. They wanted to be involved in an international network, and for returning participants, they appreciated being part of the community as such. These reasons pointed to the informally
shared and agreed upon goals of EDUsummIT, which can even be strengthened in the future by keeping a clear focus on
the purpose of EDUsummIT. The researchers, policy makers, and practitioners wanted to be equal partners in this
unique knowledge building community. Third, although EDUsummIT intends to be heterogeneous, our findings showed
that this was a potential threat to the effectiveness of the EDUsummIT, as researchers participating in EDUsummIT 2015
outnumbered practitioners and policy makers. To realize a more balanced distribution of membership should therefore be
a major concern of the EDUsummIT organisers. From the evaluation reported in this paper, it is clear that to sustain
and further develop EDUsummIT, there is a need to broaden its participation to involve more educational policy
makers and practitioners, and develop a more permanent structure in order to increase its impact. As well, whether
virtual participation should be restricted to eliminate distraction should be considered carefully. Fourth, participants of
the EDUsummIT community had ownership of the outcomes of the EDUsummIT. By inviting key stakeholders in the
field of ICT in education, i.e., policy makers, researchers and practitioners, it is assumed that EDUsummIT outcomes
would have impact in the communities in which these stakeholders are active. Thus, in addition to the scholarly papers
and policy briefs EDUsummIT produced, the impact needs also to be visible in the key stakeholders’ own local
communities. Although participants were rather positive about the impact of the EDUsummIT on research, policy and
practice, the findings of this study did not provide us with clear evidence of the extent of the impact. Thus further
research on how the EDUsummIT community impacts the networks and communities of its participants is needed to fully
understand the impact of EDUsummIT outcomes.

This evaluation study showed that EDUsummIT could be used as a model for developing and sustaining knowledge
building communities aiming to create and disseminate knowledge through close collaboration of researchers, policy
makers and practitioners. To design and develop a similar knowledge building community in other fields, it is
recommended that attention should be put to the following elements:

- Shared goal - participants need to commit to a shared goal developed by the community.
- Agency - there needs to be a strong voice of the practitioners, and thus it is important to maintain a more balanced
distribution of researchers, policy makers, and practitioners.
- Discourse – provide ample opportunities for online and face-to-face dialogues and discussions.
- Dissemination – ideas and knowledge produced have to be disseminated rapidly.
- Structure – a permanent structure is needed to sustain the development of the community.

In conclusion, this study clearly showed that EDUsummIT was successful in networking global policy makers,
researchers, and practitioners to support knowledge creation, and knowledge created in the last seven years had been
used in a number of local, regional and international institutions and organisations. In fact, more than two-dozen
presentations by EDUsummIT participants since 2009 have advocated implementation of recommendations from this
knowledge building community via national and international organizations such as UNESCO, IFIP, Kennisnet,
ISTE, and SITE. Evaluative findings are proposed in this paper as useful for continuous improvement of future
EDUsummITs while also serving as guiding considerations for emerging communities of practice wishing to develop
and disseminate recommendations for research-based best practices at the global level.

Acknowledgements

EDUsummIT 2015 was co-hosted by UNESCO Bangkok (the Asia and Pacific regional bureau of UNESCO) and
Curtin University, and sponsored by Blackboard. The authors are also grateful to the researchers who participated in
the pilot study, and the respondents who completed the questionnaire. The authors would also like to thank the two
anonymous reviewers who have provided useful comments to an earlier draft of this paper.

References


Burkhardt, H., & Schoenfield, A. H. (2003). Improving educational research: Toward a more useful, a more influential, and better-
funded enterprise. Educational Researcher, 32(9), 3-14.


Advancing Mobile Learning in Formal And Informal Settings via Mobile App Technology: Where to From Here, and How?

Ferial Khaddage1*, Wolfgang Müller2 and Kim Flintoff3

1Deakin University, Australia // 2University of Education Weingarten, Germany // 3Curtin University, Australia // f.khaddage@deakin.edu.au // mueller@md-phw.de // K.Flintoff@curtin.edu.au

*Corresponding author

ABSTRACT
In this paper a brief review of the framework that addressed mobile learning implementation challenges (pedagogical, technological, policy and research) that was developed by Khaddage et al. (2015) is briefly discussed, followed by possible solutions that could be deployed to tackle those challenges. A unique approach is then applied to bridge the gap between formal and informal learning via MAT (Mobile Applications Technology). This approach is based on STEAM (Science, Technology, Engineering, Art and Mathematics) as subjects to be taught and the specific skills needed to achieve the RLOC (Required Learning Outcome) that can support student learning informally. This specific approach shows HOW to advance mobile learning in formal and informal settings.

Keywords
Mobile learning, Mobile apps, Informal learning, Challenges, STEAM, Mobile solutions

Introduction
Mobile learning (or m-learning) as a concept and theory has evolved rapidly, and it is no longer considered just a technocentric trend, attractive for those interested in devices and technologies. This becomes obvious due to the increased reception of mobile learning in reviews on current trends in education (e.g., Johnsson, Levine, Smith, & Stone, 2010). The most recent discussions tend to assert there has been a shift from defining mobile learning as based on the devices used (Soloway et al., 2001) towards the inclusion of context (Sharplies, Taylor, & Vavoula, 2007, p. 4). Mobile learning is accepted to represent a technological advance, enabling rich, distributed and contextualized approaches to learning (Crompton, 2014). Moreover, it is accepted that m-learning is about the learner’s mobility, and how we as educators can engage students and in learning activities without them being restricted to a physical location. Nevertheless, it seems that the understanding of mobile learning and m-learning is still evolving, and that there are several considerations that should be included in trying to define the term. Specifically, an educationally relevant definition of mobile learning seems to be required (Laouris & Eteokleous, 2005). Along this line, Crompton, Muijlenburg, and Berge’s define m-learning as “learning across multiple contexts, through social and content interactions, using personal electronic devices” (Crompton, 2013). The authors would like to extend that definition by including notions of agency and timeliness, in the following way: “Mobile learning accommodates and supports personal agency of the learner in a way that the learner can decide when, where and how he or she will learn; as such, mobile learning is instrumental in just in time and on demand learning.” With this definition, we summarize notions of several research sources (Baker III, 2016; Boese, 2016; McLean, Attardi, Faden, & Goldszmidt, 2016).

A major potential of mobile technologies for learning lies in the ability to provide timely access to learning in authentic working contexts (Herrington et al., 2012; Herrington, Ostashewski, Reid, & Flintoff, 2014). Chan et al. (2006) coined the term seamless learning for this, which they define as the “ubiquitous access to mobile, connected, personal, handhelds creating the potential for a new phase in the evolution of technology-enhanced learning, marked by a continuity of the learning experience across different environments.” However, this relates to the challenge of finding appropriate and effective methods to blend formal and informal learning, as seamless learning can occur anytime – in-classroom, or outside the classroom, in formal settings, or incidental within a peer group. The corresponding challenges can be classified into four categories: pedagogical challenges, technological challenges, policy challenges and research challenges (Khaddage et al., 2015). Evolutionary change usually takes place in response to ecological interactions that operate on the overall ecosystem, Zhao and Frank (2003) suggested that the process of technology integration is evolutionary, and they stated that pedagogy, and technological skills slowly build upon each other and evolve as technology is introduced into the learning environment. Therefore these four challenges, enabling the understanding of the structure and function of each one of them. Understanding the relationships between these challenges is essential for a proper mobile learning integration and a successful mobile learning ecology (Zhao & Frank, 2003; Khaddage et al., 2015).
The adoption, support and integration of mobile learning was also a topic at the Fourth International Summit on ICT in Education (EduSummit 2015), which was held in Bangkok, Thailand, members of the Thematic Working Group 2 (TWG2) discussed methods, strategies, and guidelines for issues and challenges in the design, implementation, evaluation, and policy development of mobile learning. Some major key challenges were highlighted and discussed along with issues that policy makers, teachers, researchers, and students are facing in mobile learning. A theoretical framework for mobile learning emerged during EduSummit 2013 (Khaddage et al., 2015), which identified barriers and limitations along with dynamic criteria for mobile learning implementation. This paper presents results from discussions at this summit, reviewing the framework’s major challenges, and identifying possible solutions that could be applied to solve these challenges.

Background

New technological innovations have often been attributed with the potential to have a large impact on the field of education. For instance, when film was first used in instruction in 1913, Thomas Edison was optimistic of the potential that this could bring to education and he claimed then that “the motion picture is destined to revolutionize our educational system and that in a few years it will supplant largely, if not entirely, the use of textbooks” (Cuban, 1986, p. 9). Although instructional films did contribute a great deal in some military training (Noble, 1991), films have never replaced the traditional book. ICT and computers came up with similar optimistic promises for revolutionizing the classroom, which so far apparently have not been fulfilled. The question remains as to why mobile technologies and the concept of mobile learning could provide a similar successful concept, revolutionizing education as the introduction of books did some 2000 years ago, despite the also apparent critiques at that time.

So far, it is apparent that despite the results presented from many research studies such as Ooms et al. (2008) followed by Attewell, Savill-Smith, Douch, and Parker (2010) and many others who identified the positive effects on engagement from the introduction of mobile and handheld technologies, and the positive impact of mobile technologies integration in general – the fusion of mobile technologies into educational settings has not yet been widely adopted. Many teachers in schools and colleges are still reluctant to allow widespread access to mobile devices in a formal classroom setting (Khaddage, Lanham, & Zhou, 2009), often for reasons of lack of control of student activities and general safety concerns. This has resulted in many students being bored in classrooms and has added to the already high dropout rate, since today’s students would favor an engaging, creative and collaborative learning environment (Bonk, 2009), and mobile technologies could help in setting up such environment (Khaddage et al., 2009). Many students feel that the materials provided are somehow irrelevant for them, not engaging and do not satisfy their needs, as these materials are outdated and do not fit into today’s society (Knezek, Lai, Khaddage, & Baker, 2011; Khaddage, Knezek & Baker, 2012).

What is emerging now is a class of mobile applications that support synchronous collaboration: so called Social 3.0 apps. Examples of such apps include those with Web 2.0 support and asynchronous collaboration (e.g., SMS, Twitter) while Social 3.0 supports synchronous collaboration (e.g., Google Docs Editor). Upon reflection it is not surprising that schools and particularly teachers do not really show much interest in any technology (e.g., synchronous collaboration technology); their objective is pedagogical impact, and technologies as such are just understood as a means to achieve this, e.g., to facilitate social learning. As such, technologies such as mobile apps will still have difficulty finding their way into the primary/secondary classrooms and becoming a valuable component of the curriculum, until they provide an added value from the perspective of the responsible teachers and school administration, and until the introduction of such technology into the classroom is no longer tied with obstacles to overcome, such as technological problems, additional technical skills, and further challenges with respect to privacy and security issues. The question remains as to whether mobile learning is finally poised to make the level of impact on teaching and learning that mobility is having on most other areas of human endeavor? The rise of the Internet of Things (IoT), the scope of m-learning and apps are poised to be redefined in a very significant way. Wearable devices and IoT interactions introduce a whole arena of new considerations, and all of these combined pose challenges on informal learning. These challenges are summarized and illustrated in Figure 1, and are discussed in more detail in Khaddage et al. (2015).
Key challenges for informal learning

Informal learning lacks a clear definition, and it is often being explained by contrasting it with formal learning (Marsick & Watkins, 2001). Thus informal learning is often used synonymously with non-formal learning. Also, the boundaries between informal and incidental learning seem to be blurry. There exists agreement that informal learning represents learning that takes place outside of formal learning and educational establishments, that informal learning does not follow a specified curriculum and that it is not necessarily pedagogically planned or organized (McGivney, 2006). More concisely, the distinction between formal and informal learning can be seen in the presence or absence of a formal curriculum and the frequency with which we visit the places where learning occurs. (Looi et al., 2016).

A major part of learning takes place outside of formal learning. This is well known in adult learning (McGivney, 2006), but also applies to children and young people. Therefore it is possible to consider that school is simply one context for learning and that formal and informal learning can occur within and outside of school and institutional settings. In some ways informal learning could be considered autodidactic, or indeed incidental (Kerka, 2000), non-formal, or random (Connal & Sauvageot, 2005) learning, but some of these terms have inherited a negative connotation and are discouraged by some scholars. For the purposes of this paper the most significant differentiation between formal and informal relates to who determines the what, where and when of learning, i.e., is an accrediting body defining the required learning (formal) or not (informal).

Information technology and computers specifically allow children and young people a wide variety of activities and experiences of informal learning. It has been suggested that these hold the promise of transforming the nature of education, possibly leading to a “wider ecology of learning,” by changing the scope and the nature of learning, providing opportunities for learning new kinds of skills, and offering different and new ways of learning traditional knowledge (Sefton-Green, 2004). In this context, mobile learning represents a modality that very naturally integrates, bridging naturally the domains of formal and formal learning with its applicability in both domains, representing the basis for a perfect personalized learning environment (PLE; van Harmelen, 2006), and apparently providing the perfect learning tool for both formal and informal modes.

Still, acknowledging such informal learning represents a challenge faced by societies and by schools in particular. Griffiths and Peñalvo (2016) discussed the difficulties in recognising informal learning such that the requirements of formal settings do not formalise the learning; and they discussed ways of bridging both modes in an effort to enhance engagement in both modes through an awareness of each (Griffiths & Peñalvo, 2016).

Valuing informal learning should be a crucial element to consider when developing and adapting educational policies. So far only a few countries (such as South Africa and Ireland) award credential based on knowledge gained via informal learning, while the rest still have no formal policy framework for this type of learning (Werquin, 2010). However, there also exists an inherent risk that approaches to recognising informal learning lead to restrictions that...
might again formalise the learning, such as the requirement of formal settings. At first sight, the issue of
deployment of formal recognition of informal learning seems not connected to the field of mobile technologies and learning.
However, any assessment of learning gains is always connected to a corresponding documentation of competency
development. Mobile devices as a central computing and communication device provide the opportunities to log
learner performance on a regular basis, thus providing a perfect basis for the assessment and honoring of user skills.

From the point of view of formal education, it is also tempting to try to extend learning outside the classroom and to
connect to informal and incidental learning. Once again, this should not introduce new restrictions and obstacles,
hindering informal learning. Instead, approaches should be sought and required that ease linking of formal and
informal learning phases, and which foster the transfer of acquired knowledge and skills.

In the following section we will address more precisely the recommendations and possible solutions to the identified
challenges, as a new approach that can map informal and formal learning strategies to real work. The authors propose
that mobile technologies can bridge the gap between informal and formal learning via application of the MAT
(Mobile App Technology) model (Khaddage, Knezek, & Baker, 2012).

**Recommendations and possible solutions**

Making informal learning a valued and visible component of the education system is very important. Hence, schools
should re-evaluate the current educational framework and determine how best to merge it seamlessly with informal
learning. Informal learning should be embedded in educational contexts by training teachers via professional
development on how to help learners share knowledge gained through informal learning activities and tasks. Schools
should let teachers see the potential of this sharing and collaboration activities amongst learners. This may help to
broaden the acceptance of this type of learning. Figure 2 illustrates possible solutions to the identified challenges.

![Figure 2. Solutions to the identified challenges regarding informal learning](image)

Solutions for informal learning can be categorized according to the types of the identified challenges in Figure 1. In
the field of pedagogy, the major challenge is in preparing and opening schools for innovative learning opportunities
incorporating informal learning. We summarize this challenge to “reinventing our schools.” This includes a
reevaluation and development of school curricula to allow for more collaborative learning activities, student-focused
learning, and self-directed learning. All of these we understand as preconditions for a better linking to informal
learning. At the same time, these strategies also support novel learning activities that exploit the potentials of ICT
and mobile learning technologies specifically. In addition, teachers and schools need to get prepared to handle the
challenges and to exploit the opportunities associated with both informal learning and mobile learning. A specific
aspect in this context is the enhancement of awareness of the values of both worlds of learning, formal and informal – on the side of teachers, and, in the end, also the students’ sides – in order to enhance an engagement in both
domains and to foster a bridging of both modes of learning (Griffiths & Peñalvo, 2016).
Technology represents the second area of challenges. While we would like to argue that mobile technologies represent a major innovation enabling new forms of learning, further challenges exist. On a very practical level, current school infrastructures typically do not fulfill the requirements to allow for a seamless integration of mobile learning. In addition to relating at the level of school policies (e.g., allowing the use of personal mobile devices in the classroom) this also relates to the availability of appropriate technological infrastructures (e.g., connectivity and bandwidth to allow for a parallel connection of up to 40 devices in parallel in a classroom); this is also a question of appropriate funding and technological support. Further challenges relate to availability of technological infrastructures on a more general level. Cloud infrastructures represent a further requirement to allow for a seamless exchange of learning material and information between teachers and learners and within learner communities. In addition, accepting private devices at schools (e.g., BYOD: Bring Your Own Device) requires a certain level of standardization and interoperability, e.g., with standards from W3C, to allow for a seamless integration on class activities. These standardizations on a technological level seem to be a necessary requirement for an efficient and effective integration for both, classroom usage and usage for informal learning.

On the policy level, we see different challenges, which may be further distinguished on a school level, regional level, national level, and even international level. As mentioned before, novel approaches for honoring the increase of skills and performances from informal learning are required. Corresponding frameworks are required also.

Finally, we also need further efforts in research which is required, namely in the areas of effect case studies in impact on informal learning, and on the acquisition of competencies, both in formal and informal learning, based on comparison studies.

Figure 2 also depicts the massive change that needs to take place in order to find solutions to the existing challenges, and provides readers an overview of the big picture regarding what aspects need to be addressed on the different levels of pedagogy at schools, technology, policy, and research. It represents a bottom up approach that requires solid implementation from the foundation level to cover all levels of the school hierarchy. This may involve all stakeholders on different levels, such as students, administrators, teachers, policy makers, researchers, curriculum developers, IT designers and developers, parents, principals, coordinators, and others.

What we have been witnessing so far is that the current zeal around mobile application solutions sometimes distracts schools from the meaningful purpose of mobile integration; that is being portable, interactive, engaging, on demand, collaborative etc. (Benham et al., 2014). This is because some schools do not follow the bottom up approach by starting at the foundation level and covering all aspects, as proposed in Figure 2. Instead schools tend to bring in and integrate the technologies and deliver the same static content without any changes to existing traditional methods. Examples of such approaches are numerous – such as the many iOS/iPad projects worldwide, where some schools brought in specific mobile devices to deliver the same static content (Khaddage, 2013; Lai, Khaddage, & Knezek 2012).

**Bridging the gap via mobile app technologies**

The way schools provide and deliver teaching and learning should change to cater for the demand of our iterating and mobilized society. Mobile learning gets positioned as the enlightened and “correct” solution for the modern school settings, and can if applied seamlessly bridge the gap between inside class/ formal learning and outside class or informal learning (Kukulska-Hulme, 2007; Sharples et al., 2005). Currently one of the most flexible and portable ways to access information is via mobile technology and devices, and in particular via apps and social networking. Mobile social networking apps are being used as informal learning spaces in collaborative learning environments (Crompton 2013; Khaddage et al., 2015). Educators and schools have become interested in the application of mobile learning via apps that bridge the gap between formal and informal learning, in an effort to extend their reach into spaces and times that have previously been regarded as outside the scope of formal education. Such approaches offer opportunities to flexibly deliver learning content that complements formal class settings.

While figure 1 is a representation of the big picture and illustrates a general overview of the change in education, we try to sketch a possible approach on a more concrete level that could be implemented by schools to bridge the gap between formal and informal learning, based on evidence from the current research. This approach is proposed as being capable of offering informal learning effectively and efficiently.
Researchers in ICT (Information and Communication Technology) in education have been reporting for years that formal learning is getting boring and this is resulting in dropouts in schools. Formal learning is slowly disappearing, as it follows a static linear and routine approach that the 21 century digital native learners find boring and no longer useful (Voogt, Knezek, Cox, Knezek, & ten Brummelhuis, 2013). Some would argue that both formal and informal learning are valuable for gaining skills and obtaining knowledge (Billet, 2001). Other researchers have shown that strong foundation skills allow learners to make the most of their informal learning (Nordman & Hayward, 2006). They suggest that informal learning may not be useful to those without such a foundation. The authors believe it is crucial to find a unique approach capable of bridging the gap between formal and informal learning via mobile technologies and applications. As we see from the literature, there is no doubt that new technological innovations are changing the whole ecosystem and the economic direction; hence changing our education system.

In Table 1 the authors propose a unique approach based on STEAM (Science, Technology, Engineering, Art and Mathematics) since the main themes of STEAM are fostering innovation, the need for twenty-first century skills, and divergent and convergent thinking amongst learners (Kim & Park, 2012). Schools could identify Subjects to be taught and the Skills along with the Required Learning Outcome (RLOC) that can support student learning. As seen in Table 1 subjects are categorized, and skills with the RLOC are identified. This may vary from school to school, depending on the subjects being taught. But for an illustration of how this method would work, the focus is on STEAM subjects, to show the categorised list and use it as a benchmark.

Instead of generalizing and following the same approach for all subjects, we present a flexible approach that could be altered depending on the subjects and skills, and could bridge the gap between formal and informal learning. The main focus is on the skills that students need to gain in order to address the RLOC rather than the subject itself, and then on applying the appropriate methods to support using innovative mobile technology, in an attempt to advance mobile learning in formal and informal settings. This method is to focus on the important skills we want students to gain from studying those STEAM subjects: Science, Technology, Engineering, Art, and Mathematics. Do they all demand a singular approach to developing skill and capability necessary so that RLOC are met for each particular subject? Is formal learning or informal learning the most appropriate method for ALL of the skill types? Of course the answer would be NO.

Table 1. Subjects are categorised and skills with the RLOC are identified

<table>
<thead>
<tr>
<th>Skills and RLOC</th>
<th>Subjects</th>
<th>Science</th>
<th>Technology</th>
<th>Engineering</th>
<th>Art</th>
<th>Mathematics</th>
</tr>
</thead>
<tbody>
<tr>
<td>ainting and interpreting data</td>
<td>Observation</td>
<td>Coding</td>
<td>Assessment and estimation</td>
<td>Perseverance</td>
<td>Problem solving</td>
<td></td>
</tr>
<tr>
<td>Making models</td>
<td>Logical thinking</td>
<td>Assembly and awareness</td>
<td>Focus</td>
<td>Analysing, estimating,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measuring and predicting</td>
<td>Creativity</td>
<td>Teamwork and leadership</td>
<td>Collaboration</td>
<td>Measurements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computation</td>
<td>Analytical and adaptability</td>
<td>Dedication</td>
<td>Interpret and processing information</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Now we can classify the skills from STEAM that could be gained via a formal setting versus those skills that could be gained in informal settings. MAT (Mobile App Technology) is applied to see how it can be used in order to advance mobile learning in formal and informal settings. This is briefly illustrated in Table 2.

Table 2. Formal settings skills versus informal settings and MAT application

<table>
<thead>
<tr>
<th>Type of learning and MAT</th>
<th>Informal learning</th>
<th>Formal learning</th>
<th>Some (MAT) that can be used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation, observation collecting data</td>
<td>Coding interpreting data</td>
<td>Mobile cloud-based apps</td>
<td></td>
</tr>
<tr>
<td>Teamwork collaboration</td>
<td>Logical thinking</td>
<td>Social spaces on mobile apps</td>
<td></td>
</tr>
<tr>
<td>Making models and creativity</td>
<td>Leadership and adaptability</td>
<td>Hybrid mobile apps</td>
<td></td>
</tr>
<tr>
<td>Measuring and predicting</td>
<td>Computation</td>
<td>Native mobile apps</td>
<td></td>
</tr>
</tbody>
</table>
As shown in Table 2, some of MAT that can be used could be applied to bridge the gap between formal and informal learning to help students gain the required specific skills. Students can use mobile app technologies to demonstrate a variety of tasks. Apps come in different types, hybrid, native, web/or cloud apps. The list below identifies each type:

- **Hybrid apps**: are cross-platform and combine both characteristics of native and web/cloud apps, they run on the device itself, but they run inside a native container within the operating system of the device itself (Mudge, 2012; Khaddage et al., 2015)
- **Native apps**: are device specific and operating system specific, they are most suitable for gaming, as they provide high quality graphics and speed since they run within the device engine, thus making them the most engaging robust apps (Korf & Oksman, 2012).
- **Web/Cloud apps**: they are server side apps (cloud-based or web-based), are device independent and run on all platforms. They only use the browser of the mobile device and they support BYOD (Bring Your Own Device) (Boulos et al., 2011; Khaddage, 2013).

Khaddage and Knezek (2012) discussed how the use of already-existing educational apps is important and can be used to empower informal learning. Many authors have described and discussed methods of mobile technology integration (Boden, 2001; Candy & Edmonds, 2002), and how it improves creativity, collaboration and engagement amongst students. Khaddage and Knezek (2012) stated that students can use mobile apps to demonstrate a variety of tasks in formal and informal settings. They described the benefits that can be offered, as illustrated in Figure 3.

![Figure 3. Skills learned via mobile apps (Khaddage & Knezek, 2012)](image)

Students can use some of the already existing educational apps to complete learning activities. By using mobile apps, these tasks could be performed anywhere and anytime and the RLOC for any specific subjects could be met. But the issue is how to get this acknowledged and recognized by teachers? This proposed unique approach can be considered as one solution, especially when schools ensure that those mobile apps’ initiatives for teaching and learning are aligned well with subject objectives and the RLOC as listed in Table 1. Then those mobile apps could be used across variety of subjects in and outside of classrooms. These mobile apps are capable of running on the learner’s mobile device to deliver learning that can offer skills in a hybrid/blended-learning environment, where the main focus is on the required skills regardless of how they are obtained, formally or informally. For example, with the advent of the mobile phones and apps, informal learning can now take place in mobile cloud space, in hybrid mobile apps, and in social networking spaces, as illustrated in Table 2. Virtual meeting points can be created where learners can collaborate with each other via apps on their phones (Fotouhi-Ghazvini et al., 2009). Learners can learn across spaces via apps and virtual interaction. They can communicate, engage and create, then gain ideas and learning resources that they obtained in one location and apply or develop them in another – and that could be in class in a formal setting (Sharples et al., 2005). By doing so, this gap between formal and informal learning can be bridged with MAT, thus forming a new and innovative learning setting that can be hybrid, dynamic and synergetic. It is also important for schools to design and integrate tools and devices that can be used across subjects in classrooms, thus allowing flexibility and portability (Khaddage & Knezek, 2012).
While there are indications of the educational transformation possible with informal learning, particularly with MAT integration, more needs to be understood about this phenomenon. The change should start at the government level by re-allocating more funding for informal learning initiatives, including academic studies to pinpoint the benefits of this type of learning.

A review of published literature on mobile learning highlights the fact that informal learning is indeed possible with mobile apps (Jones et al., 2006; Scanlon, Jones, & Waycott, 2005; Clough Jones, McAndrew, & Scanlon, 2007; Khaddage, Knezek, & Baker, 2012). The only limitation to mobile app integration for informal learning is that it can be difficult to capture informal learning when it occurs, and there are no common key performance indicators against which to measure the progress of learners. But the opportunities to evaluate the skills and RLOC of informal mobile learning have increased with the wide availability of tools and technologies, combined with easy access to Internet via Wi-Fi and 3G. There is also an increase in the number of educational apps for different platforms, although iOS is still leading in terms of the total number of native educational apps available on any platform (Mobile matters). This is largely the result of some schools teaming up with Apple for mobile technology integration. Higher Colleges of Technology iPad initiative is one example of such integration in an effort to offer a flexible informal learning environment (Khaddage, 2013).

Innovative practices and future considerations

When it comes to design challenges of mobile learning, leading mobile apps are delivering exceptional user experiences (UXs) achieved with a variety of techniques including motivational design, “quiet” design, “playful” interfaces and new methodological approaches (Gartner, 2015). Designers are also creating apps that can accommodate mobile challenges, such as partial user attention and interruption, or exploit technologies with novel features in an attempt to hook the learner into using the technology to complete the learning task. All these newly developed apps do support STEAM subjects and offer a challenging informal learning environment. A good example of this is augmented reality and virtual spaces via mobile apps. According to Gartner (2015), by the year 2020, an affluent household will contain several hundred smart objects, including domestic appliances, sports equipment, medical devices and controllable power sockets. These domestic smart mobile objects will be a part of the Internet of Things (IoT), with the majority of them being able to communicate in some way with an app on a smartphone or tablet. Smartphones and tablets will perform many functions, including acting as remote controls, displaying and analysing information, interfacing to social networks to monitor “things” that can tweet or post for learning activities and tasks informally. This combination of smart objects and mobile apps and technologies will enable an even wider range of informal learning opportunities (Gartner, 2015). So far only a small number of smart objects and appliances are available in 2015, such as sensors. However, the range of domestic smart objects will continue to grow. How this will affect the learning environment in informal settings is still not clear, as new technologies bring along new issues. Wearable technologies such as watches displaying email and messages via an app on those devices will pose new security and management challenges. Devices that can record video will raise many privacy concerns, as has been demonstrated by Google Glass. Schools are still fretting about mobile learning policies as well as pedagogies, research and technologies, and they are struggling to find ways of proper integration. It is crucial for schools to advance in their offering and how they deliver the learning content; otherwise they are going to be left behind. Hopefully the proposed approach/solution in this paper, if deployed properly, may help in solving the identified key challenges and help in preparing schools to find unique approaches to blend informal learning seamlessly into their existing settings.

Conclusion and future work

When bridging formal and informal learning, schools should not be aiming to unintentionally formalize informal learning, but rather they should be looking to find unique methods and approaches to incorporate mobile learning and blend it seamlessly into their settings to create an engaging informal learning environment. The approach proposed in this paper could be used as a good starting point. While there are potentials with informal learning and particularly in low resource contexts, more research is needed to further understand this shift in technology and in educational settings (formally and informally). More funding for informal learning initiatives should be made available in order for educators, researchers, policy makers and practitioners to highlight the value and benefits of
this type of learning. The continuous and consistent work of EduSummIT’s Thematic Working Group 2 via the application of the mobile learning framework (see Khaddage et al., 2015) to challenges identified is considered useful techniques that can be used to test the ecological theory in the mobile learning framework. The unique approach proposed here that combines skills RLOC and the inclusion of STEAM to test MAT, could be considered a first step towards bridging this gap between the two types of learning (formal and informal), in an effort to keep advancing mobile learning in all different settings, shapes and forms, and hence assisting researchers, policy makers and educators in the practical implementation within a mobile learning environment.

Acknowledgments

This work is grounded in the discussions of Thematic Working Group 2 (TWG2) during the Fourth International Summit on ICT in Education (EDUsummIT 2015) which was held in Bangkok, Thailand. The authors would like to acknowledge the contribution and active participation of all members of TWG2 during the summit, and they are listed in alphabetical order by first name: Aunken Tungtarova, Barry Quinn, Dolores Zambrono, Immo Kortelainen, Linda Fang, Lucila Perez, Rowland Baker, and Yidda Marcial.

References


Kukulska-Hulme, A. (2007). Mobile usability in educational contexts: What have we learnt? The International Review of Research in Open and Distributed Learning, 8(2), 1-16.


Infusing Creativity and Technology in 21st Century Education: A Systemic View for Change

Danah Henriksen1*, Punya Mishra2 and Petra Fisser3

1Arizona State University, Tempe, Arizona, USA // 2Michigan State University, East Lansing, Michigan, USA // 3National Institute for Curriculum Development, The Netherlands // danah.henriksen@asu.edu // punya@msu.edu // p.fisser@slo.nl

*Corresponding author

ABSTRACT

In this article, we explore creativity alongside educational technology, as fundamental constructs of 21st century education. Creativity has become increasingly important, as one of the most important and noted skills for success in the 21st century. We offer a definition of creativity; and draw upon a systems model of creativity, to suggest that creativity emerges and exists within a system, rather than only at the level of individual processes. We suggest that effective infusion of creativity and technology in education must be considered in a three-fold systemic manner: at the levels of teacher education, assessment and educational policy. We provide research and practical implications with broad recommendations across these three areas, to build discourse around infusion of creative thinking and technology in 21st century educational systems.

Keywords

Creativity, Technology, Teacher Education, Teacher Professional Development, TPACK, Policy, Assessment

Introduction

In this article, we explore creativity alongside educational technology, for 21st century education. Creativity has seen heightened discussion in fields such as psychology and education (Sternberg, 2000; Sweller, 2009), and in popular interest in broader culture as well. Lewis (2008) noted that creativity is a coveted quality of thinking often an important aspect of innovation and change. There has also been increasing educational research to support the importance of creativity in fields of thinking and learning (Henriksen & Mishra, 2015; Robinson, 2011; Williams, 2002).

Much of the research on creativity has focused on individual creativity, or psychological, psychometric or personality approaches. There has been comparatively little research on creativity in classrooms (DeSouza Fleith, 2000). The field of education must consider the applications and rationale of creative educational practice and policy, especially for 21st century, technology-rich contexts. New technologies have altered teaching and learning rapidly, with innovations and affordances for creating and sharing ideas and content. We must consider the development and impact of learning technology not in isolation, but rather alongside opportunities for creative education.

We begin by considering the global context for an emphasis on creativity, then describe the foundations for creativity in society and in education, alongside educational technology. This emphasis on creativity and its curricular integration requires forethought and planning. In drawing on a systems model for creativity in broader culture, we suggest that there are three threads of importance for creative education with technology: teacher education, assessment, and educational policy. In this three-pronged approach, we describe how each has a role in building appropriate educational contexts to meet the needs of 21st century learners and teachers.

Context for change

The rapid pace of new technology development has presented a challenge for classroom technology integration (Zhao, 2012). Creativity is deeply connected to issues of technology integration, so these issues of creativity and technology can be considered in tandem.

While new technologies and discoveries have been a constant through human history, digital technologies rapidly scale up the technological growth. We have seen an incredible flowering of creativity and innovation fuelled by the capabilities of such technologies. From Google to Facebook, from cloud computing to YouTube channels, digitality has altered how we live, work and connect with each other (Mishra & Henriksen, 2013). Technological change is
driven by human creativity, and in turn provides new contexts and tools for creative output. Given this reciprocal relationship between creativity and technology we suggest that teaching and learning must emphasize their connection (Henriksen, Hoelting, & The Deep-Play Research Group, 2016). It is important to explore the relationship between these constructs across varied, global educational contexts.

This is a challenge, because even as standalone issues, both have confounded attempts at common, effective educational approaches. Yet a better understanding is vital. Creative thinking is essential for 21st century success, as societal problems become more interdependent, global and complex. Daniel Pink (2005) has stated that the skills that were important in the past (the popularly termed “left-brain” skills) are still important but not enough. He suggests that “the ‘right brain’ qualities of inventiveness, empathy, joyfulfulness, and meaning—increasingly will determine who flourishes and who flounders (Pink, 2015, p. 3).”

While there has been increased interest around creativity in education, this has not always translated into practice. Traditional “drill and kill” approaches or standards-based teaching have often squeezed creativity out of the curriculum or areas of policy and assessment (Giroux & Schmidt, 2004). For all its importance, creativity is a concept that has not been well understood, framed, or defined. Education needs a frame to help students and teachers develop creative thinking skills that span disciplines, and use technology tools for creative solutions and outcomes.

In the next section, we consider some key literature on creativity, and situate our thinking in a definition of “creativity.”

Examining creativity in its foundations

Research has shown intellectual, educational, and talent-building advantages associated with creativity throughout life (Blicbau & Steiner, 1998). Educational psychologists have noted connections between creativity and other significant areas including: life success, leadership in the workplace, psychological functioning, and intellectual/emotional growth (Williams, 2002).

Sternberg (1999) noted that creativity is available to everyone, but is most prevalent in young children. As he states, creativity “may be harder to find in older children and adults because their creative potential has been suppressed by a society that encourages intellectual conformity” (p. 93). Robinson (2011) also suggested that conventional approaches tend to crush students’ natural inclinations toward creative and divergent thinking. Sternberg (2006) has stated that, “When students are taught in a way that fits how they think, they do better in school. Children with creative or practical abilities, who are almost never taught or assessed in a way that matches their pattern of abilities, may be at a disadvantage in course after course, year after year” (p. 94).

The role of the teachers and classroom settings is an important influence upon student beliefs and development of their own creativity. Amabile (1996) asserts that when all variables that influence creative development are considered, most factors are classroom-related. Creative teachers show a willingness to try new things, give real-world assignments, and use cross-disciplinary approaches (Henriksen & Mishra, 2015). As a result, their students tend to be enthusiastic and engaged (Kiely, 1998), building the skills and habits of mind for success in problem solving and applying knowledge (Zhao, 2012).

Despite its importance, theorists have struggled to find common ground in a concrete definition of the term (Sternberg, 1999). Yet having a definition is essential to the instantiation of an idea—in order to apply a working shared understanding of its themes.

A definition of creativity

Creativity can be viewed as a process and/or a product, and is generally thought of as the production of useful solutions to problems, or novel and effective ideas (Amabile, 1996). An idea that has novelty, but lacks in value or effectiveness to other people, cannot be considered “creative” (Cropley, 2003).
Two factors in most discussions of creativity are “novelty” (or newness, originality, freshness, uniqueness, etc.) and “effectiveness” (or value, usefulness, quality, etc.) (Sternberg, 2006). But while these two recur in many definitions for creativity, some scholars have called for the inclusion of a subtler, third component.

Sternberg and O’Hara (1999) argued that “task appropriateness” should be added to the definition, speaking to the contextuality in creative work. Based on this, creativity lies in the ability to create ideas or works that are “novel, high in quality, and task appropriate” (p. 255). This suggests that creative work is dependent on context, because it is assigned value in relation to the domain it is created within. Mishra, and Henriksen (2013) note that an innovative mathematical proof or a unique beautiful painting are incredible different things, yet they are both “creative.” They both have an aesthetic context that goes beyond novelty and utility. Mishra and Koehler (2008) describe this aesthetic sensibility in context as “wholeness,” which is a third, crucial component of creativity. Thus they offer a “NEW” (novel, effective, whole) definition of creativity (Mishra & Koehler, 2008; Mishra, & Henriksen, 2013).

Here, we suggest this NEW definition for educational contexts, and as our definition in this article. We define creativity as both the oft-noted “novel,” and “effective,” in addition to the subtler component of “wholeness” (or context, important to education). Recent scholarship has focused on this definition of creativity, along with attempts to develop rubrics to measure creative student output (Henriksen, Mishra, & Mehta, 2015; Mishra, Henriksen & Mehta, 2015).

**Going beyond “what is creativity” to “where is creativity”**

While our definition of NEW (novel, effective, and whole) is supported by research (Henriksen, Mishra, & Mehta, 2015), we acknowledge that any shared definition of a subjective concept is challenging. So we propose that we must also go beyond defining “what is creativity?” To ask, “where is creativity?” Asking this helps us locate creativity in specific realms that impact practice, a key point for the field of education.

In this we draw upon Csikszentmihalyi’s (1997) discourse about: “where is creativity?” Csikszentmihalyi moves beyond typical definitions to emphasize a systems model of creativity, in which creative production is an interaction of systemic elements.

To understand creativity as a complex phenomenon “we need to abandon the Ptolemaic view of creativity, in which the person is at the center of everything, for a more Copernican model in which the person is part of a system of mutual influences and information” (Csikszentmihalyi & Csikszentmihalyi, 1988, p. 336). We must consider how creativity arises from a dynamic interaction of “a system composed of three elements: a culture that contains symbolic rules, a person who brings novelty into the domain, and a field of experts who recognize and validate the innovation” (Csikszentmihalyi, 1997, p. 6).

Csikszentmihalyi (1997) asserts that creativity lies in the interaction of the individual, the domain, and the field. This system is where creative work emerges (as represented in the diagram below) and functions among these three arenas. Creativity is developed and disseminated based on the judgments and interactions of members of those levels. As the diagram indicates, the Individual, the Field and the Domain work together reciprocally to decide if something is creative (i.e., novel, effective & whole).

At the level of the individual, individual people (or groups/teams) produce creative work, ideas, art, or new discovery. But creativity does not happen there alone. Creativity is affected at the level of domains, which are areas of specialized knowledge (e.g., mathematics, biology, physics, art, law, and more). A domain is the symbol system that an individual (or group) working in an area uses. It includes the tools, rules, conventions, knowledge, norms, and systems of techniques, codes, or symbols that help a person create or discover new things in the domain. At the level of the field, creative work is disseminated to an audience to make impact. The field is a collection of “experts,” communities of practice, or people with the knowledge capital and clout to make judgments influence the domain (about what is valuable work, at the cultural or social level) (Csikszentmihalyi, 1997).

Each of these three components – person, field and domain – exerts and receives influence from the others. Each component is a necessary factor in creativity (and even expertise) but not sufficient in itself to produce impact or valuable novelty. Creativity exists as a dynamic process emerging through a system of interactions.
In any systemic interaction we can see how creativity has multiple entry points within the system. So, this informs our considerations of how we think about the systemic aspects of creativity in education. Given this, we must consider how and where it might be infused through entry points. In identifying categories of stakeholders that play a role in this process, we decide where to place our emphasis. In later sections, we elaborate on how this type of systems-thinking model might be defined in education. We suggest creativity must become systemic, at the levels of teacher education, assessment, and educational policy.

We must note that technology plays an important role in every aspect of this transaction among the domain, the individual and the field. It impacts construction of knowledge, its sharing with a wider community and its acceptance by the field. Therefore, we explore this relationship between creativity and technology next.

**Creativity and technology in education**

Creative teaching alone is a complex and open-ended arena. Incorporating effective uses of technology for teaching is also complex on its own terms. So things become more complex when these two intersect, as they must in 21st century classrooms. Mishra, Koehler, and Henriksen (2011) have argued that the best uses of educational technology must be grounded in a creative mindset that embraces openness for the new and intellectual risk-taking. This is a tremendous challenge for any teacher, but especially new teachers. It has not been addressed in great detail by teacher education, professional development, or educational policy.

Contemporary technologies often bring new possibilities for people to be creative. In classroom terms teachers must understand the range of ways in which technology can present content creatively, and see how this intersects with different pedagogies. Since technologies emerge and shift continuously, a tool-based focus is akin to a moving target. Creative real-world approaches to teaching might allow us to also consider how technology helps us view and learn content in original or compelling ways. It allows us to create content, rather than summarize and repeat it.

TPACK (Mishra & Koehler, 2006) has become a central framework for using technology well in the classroom. It focuses on knowledge types for effective teaching with technologies (from chalkboards to smart boards, pencils to Pixlr). The TPACK framework suggest that that teachers have a specialized brand of knowledge for using classroom
technology, involving an integrated combination of Technological, Pedagogical, and Content knowledge. While TPACK framework has received a significant level of scholarly and theoretical attention (Herring, Koehler & Mishra, 2016), it has been argued that the framework overall is neutral with regard to the goals of teaching. As Mishra, Koehler and Henriksen (2011) write:

It is also important to note that the TPACK framework offers no specific directives about what content to teach (science or music), which pedagogical approaches are useful (didactic or constructivist), and what kinds of technologies to use in teaching (digital or analog). Thus given the changing world we live in, it becomes critical for us to ask ourselves what it is that today’s students need to know in order to succeed… Once we identify these larger purposes and goals, the TPACK framework helps us consider how to achieve them... (p. 24).

Focusing on the affordances of tools, and how tools can serve the content in novel and effective ways, helps us use creativity as a driver for good teaching with technology. One of the key affordances of digital technologies is that content or knowledge can be created, shared, and discovered much more quickly and easily (Henriksen et al., 2016).

With digital media contributing to globalization and diversification of ideas and content, there has been a rethinking of how we communicate and share ideas, art, culture, and other forms of content. New technology has much to offer to the world of creative sharing—as seen in internet crowd-sourcing of data or ideas, new applications for content development, creating unique or remixed work, sharing video/audio/images/text across global contexts, and websites that allow diverse creators to share content (for example, YouTube, Sound Cloud, Vimeo, to name a few) (Henriksen, et al., 2016). This explosion in technologies for content sharing and development is transforming how culture, art, and knowledge emerge within disciplines.

Contemporary digital and networking technologies can play a significant role in the systems view of creativity we presented earlier. The field, the domain, the individual and their interconnections are transformed by content creation and sharing technologies. For instance, an individual can sidestep the gatekeepers of the field and connect with an audience directly. Thus the gate-keeping function now shifts from the members of the field to the audience. Moreover, the content audience can speak back to the individual creator. A good example of this is the YouTube channel, Veritasium, created by Derek Muller. Muller’s channel focuses on physics, and the choice of topics is often driven by questions sent in by his audience—allowing learners all over the world interact directly with science ideas and each other via his site. Individuals can also follow their interests within a domain more promptly and easily – given the wide access to resources and information, and tools to create new knowledge/content. Thus we see the rise of YouTube superstars, individuals who have sidestepped standard approaches to creative success. This suggests a possible reconfiguration of the standard systems view of identifying creativity.

Much of our discussion here focuses on creativity “in the wild,” i.e., creativity in the broader world. This is different from creativity in educational contexts. We suggest there are two key aspects to the role of technology and creativity in the classroom. The first is that educators must be creative in devising new ways of thinking about technology, particularly for teaching specific content. Most digital tools (be it Facebook or a smart-phone or any other tool) have usually not been designed for educational purposes. It becomes an opportunity for the teacher to creatively repurpose existing tools for educational purposes (Koehler et al., 2011). Second, it is also clear that technologies afford new ways of constructing, representing, communicating, and sharing knowledge, providing opportunities for creative output by and between students that were not possible before.

These two approaches complement and support each other. An example of the first approach would be when teachers use tools (such as Facebook or Twitter) not designed for education in creative ways in the classroom. While an example of the second would be the opportunities that tools such as VoiceThread allow, for students to engage in creative multi-modal writing. Within this new context, there is a reciprocal relationship between creativity and digital technologies. Here we mean that technologies allow for new and creative pedagogical practices, but also that educators must develop a creative mindset to teaching and learning. Only then can they fully exploit the potential of these new technologies. Of course the idea of TPACK is key—in that teachers need to always ask themselves how these new creative use of tools fits with the content to be covered and their pedagogical approaches.

It is important to note the significance of teacher beliefs about creativity and technology. There is research indicating that teacher beliefs about subject matter, learning, teaching and technology influence the way they approach practice (Ertmer, Ottenbreit-Leftwich, Sadik, Sendurur & Sendurur, 2012; Kim, Kim, Lee, Spector, & DeMeester, 2013;
Richardson, 1996). As Borko and Putnam (1995, p. 37) write, “to understand teaching, we must study teachers’ knowledge systems; their thoughts, judgments, and decisions; the relationships between teachers’ knowledge systems and their cognitions; and how these cognitions are translated into action.”

The reciprocal relationship between creativity and new technologies has implications for (a) teacher education and professional development; (b) for how we evaluate student learning and output; and finally, (c) for policies we enact to support teachers and students in this arena. There are many ways to categorize the systemic elements of creativity in education. But for practicality and clarity, we suggest these are the core areas that must be considered and we look at each in greater detail below.

**Three stands of influence for approaching creativity in 21st century education**

Our goal in this paper is to lay out a broad plan for action. We do not provide this in micro-detail because to do so would be challenging (if not impossible), given the range of settings and variables in education. But we do wish to introduce the idea that each of these three arenas of teacher education, assessment and policy are crucial to moving these ideas forward. Specifically, teacher education focused on creativity is necessary for creativity to be infused into classrooms. Teacher training must support repurposing of technologies in the classroom and teaching approaches that creatively engage students with content. However, creative student work must also be assessed—requiring an emphasis on the assessment of creative work. Finally, none of this is possible if we do not focus on the broader policy goals of integrating technology and creativity across the policy framework of education. Thus we argue that a focus on these three areas is the first step towards locating creativity within educational systems.

**Teacher education**

A teacher’s pedagogy is often a primary driver of how students develop and learn. Teachers who model creativity tend to fluidly enhance, support and develop the tendency in their own students (Amabile, Conti, Coon, Lazenby, & Herron, 1996). We must build teaching dispositions that take advantage of the affordances of new tools for learning and thinking creatively, in ways not possible without new technologies (Ertmer et al., 2012). But effective teaching is difficult in itself, even without the added elements of creative and technology-savvy practices. How do we support the development of creative pedagogy, along with effective use of classroom technology, to support the 21st century teacher and student?

Teacher education programs are often the core driver of how new teachers see the profession, how they interact with students and develop their classroom practices and repertoire. Therefore, it becomes important that we see teacher education as a key site in developing a creative mindset and practices that encourage the use of new technologies creatively in the classroom. Yet the role of creativity and technology in teacher education is rarely clear, varying at the school/program level. It is essential to build a platform for teacher education programs that addresses creative, technology-rich approaches and pedagogies. In brief, the research and scholarship in this area suggests the following key recommendations.

**Teacher education / teacher professional development recommendations**

- *Develop Teacher Education curriculum that integrates technology and creativity across the program*
  
  Current teacher education curricula may give some emphasis to teaching creatively with technology – though even there it appears spotty at best. The other aspect that of teaching to enhance creativity in students, and to explore the affordances of technology to do so, has received even less attention. Integration of ideas related to creativity and technology need to be across the program and curriculum. Research has shown that highly creative teachers tend to engage in a variety of creative pursuits that they draw into their teaching practice (Henriksen & Mishra, 2015). Teacher education students could be encouraged to actively spend time in creative interests, and incorporate these into lessons and activities through technology. This might include coursework that specifically asks new teachers to “play” with approaches to using technology in the curriculum in creative lessons on content. Opportunities to engage in lesson planning focused on real-world, cross-curricular and novel approaches to content and technology (TPACK) would help build creative teaching skills, as a part of a teacher
education curriculum. Examples of such practices can be found in the special issue devoted to teacher education, creativity and technology (Henriksen & Mishra, 2015), and in Koehler et al. (2011).

- **Specific course / programs focusing on creativity and technology**
  Even as we seek to suffuse a “creativity mindset” across programs, we see the need for specific courses that target creativity and technology and their use in the classroom teaching/learning context. This includes more theoretical knowledge of creativity into teacher education curricula for pre-service teachers, particularly in emphasizing the relationship between creativity and student achievement or teacher effectiveness and impact (DeSouza Fleith, 2000; Henriksen & Mishra, 2015). Other researchers have highlighted the ways that TPACK can be used as the basis of creative approaches to technology integration (Koehler et al., 2011). A theoretical understanding of creativity is something that should connect to practical applications. Teacher education students must have the opportunity to consider how creativity works in their own lives and practices, particularly with regard to technology and tools for teaching. See Henriksen and Mishra (2015), and Koehler et al. (2011) for examples.

- **Identify / use a framework that connects creativity and technology to curriculum guidelines**
  Curriculum guidelines are overarching structures that determine how specific curricula are designed. It is important that the dual-goals of teaching creatively with technology, and teaching for enhancing creativity with technology, be incorporated in these guidelines. The use of theoretical frameworks (such as the systems model of creativity or TPACK) give cohesion to any research study or paradigm. Without a good framework guiding the work, it is hard to move beyond ad-hoc ideas and examples. While frameworks currently exist for creative education, or for technology infusion in education, it is difficult to find a framework that integrates the constructs. Developing such a framework would guide teacher education programs on a path that blends these ideas into their curricula.

**Assessment**

Creativity, due to its open-ended nature, is difficult to evaluate and assess. However, if creativity is to become a part of the educational process, developing a range of assessments is essential. The arena of assessment of creativity is rife with multiple challenges, which tend to present as dichotomous tensions. We see these tensions as inherent and not ones that can be wished away. As educators we have to contend with these dichotomies, and find a resolution or compromise that works in our specific context. We list a few of these tensions below.

- **Psychometric vs. Behavioral:** Most creativity research has focused on identifying psychometric characteristics of creative individuals (such as affinity for risk-taking, cognitive flexibility, etc.). At the other extreme are behavioral measures of creativity (such as the alternative uses test, where participants are asked to come up with as many alternative uses for a random object).

- **Process vs. Product:** This distinction is important in the classroom where the teacher may focus learning processes for creative solutions (sometime irrespective of whether the final result was judged creative or not) OR may focus on the output of the creative activity. Most of the current discourse on assessment has emphasized the process aspect, though there are a few approaches to considering the final product developed by students. A product is concrete and more amenable to evaluation, but process may be more important to teachers since its respects the whole learner (process may be idiosyncratic and playful, which also brings up another challenge). For example, an ICT approach tends to focus on the final output or product, compared to the art teacher concerned with process. In considering both creativity and technology, assessment has to consider how to navigate between and consider both product and process, for effective, creative uses of classroom technology.

- **Individual vs. Group:** This is particularly important in contexts where teachers give students open-ended, group projects. These projects are more authentic (in being similar to actual work-place situations) but prevent the easy assessment of the individual contribution, which has typically been the mainstay of assessment in schools. In other words, how do we get students to engage in the kinds of collaborative and open-ended products that support creativity, while also assessing their individual performance?
• Domain general vs. domain specific: This is an ongoing problem and dispute among most creativity researchers. Evaluation becomes more challenging unless we start from a place of solid agreement on whether creativity is located specifically within domains, or whether it is a more general and extendable thing.

The overwhelming theme of creativity and assessment revolves around the challenge of navigating tensions in evaluating and assessing creativity. It is important that we not focus on just one approach towards assessment but explore a range of alternative assessment formats that consider how creativity and technology intersect. These allow for the dynamic, flexible, application of idea across learning contexts.

Assessment recommendations

• Recognize that assessment of creativity (with or without technology) exists within a range of tensions/dilemmas
   Issues related to the assessment of creativity have a range of dimensions (individual – group, process – product, domain general – domain specific etc.). These are not problems to be solved, but essential tensions to be resolved in a context sensitive manner. This is important when we consider technology-driven activities and assignments, where often the mere inclusion of technology is seen as being creative. Clearly this is a far more complex problem, where the role of technology needs to be better understood and articulated so that creative teaching, and teaching for creativity, (using ICTs) can be better understood.

• Use alternative forms of assessment – triangulation through technologically sophisticated, dynamic and flexible approaches
   It is essential that we explore a range of different and alternative forms of assessments (i.e., open ended versus more constrained tasks) to allow for the dynamic, flexible, triangulation of the construct as it plays out in different learning contexts. Technology can play an important role in terms of allowing teachers and learners to both easily construct creative artifacts. However evaluating these artifacts (and the process that led to their construction) is complex. It requires evaluating the artifacts (or the process) through multiple assessment techniques (formal and informal; process and product; formative and normative).

• Evidence based research on creativity and technology from the classroom
   Finally, there is clearly need for research connected to these different learning contexts where creativity and technology co-exist. These contexts differ in a range of dimensions (e.g., formal–informal; disciplinary–transdisciplinary—multi-disciplinary). This requires a new form of research, that both honors the complexities of practice as well as the broader goals and themes of learning and creativity, and the role of technology in the process. Only through such research can we offer sound, data-driven guidelines for future educators, scholars and researchers.

Educational policy

Creativity can be learned, but since it is a thinking skill it can only be “learned by doing” or as “learning in action.” Creativity involves approaches to thinking rather than a set body of knowledge that can be taught. However, we can reinforce and support sustained creativity as a “habit of the mind.” However, this also means that the education system and educators must recognize and support a sustained facilitation of creativity as a habit of the mind, and agree upon what that is and how to engage it. This can vary greatly across contexts and cultures. So essential challenges involve convincing policy makers, who often prefer clear answers and objectivity that it is important to infuse curricula with creativity, an area that does not have one “right” answer. Along these lines, policy must also begin to consider the intersection of technology with creativity, and offer guidelines for how these ideas can intersect in the classroom.

We must realize that policy texts at all levels in education (macro, meso, and micro) are predominantly indicative of practice, rather than definitive, because policies are mediated by schools, teachers and other actors in education (Ball, 1997; Wyse & Ferrari, 2014). But as we acknowledge this, policy texts are also representations of discussions on certain topics. They are important enough to be emphasized in a document, and often the basis for further curriculum development. Along with curriculum development, policy documents are often used to compare countries, regions and schools with each other. For instance, Heilmann and Korte (2010) carried out a content
analysis of national curriculum texts to study the role of creativity and innovation in compulsory education in 27 countries of the European Union. The outcomes of such studies can promote new policy texts and approaches. Wyse and Ferrari (2014) state: “The inclusion of explicit reference to creativity [in all national curricula of the EU27] is an indication that creativity is valued by policy-makers and curriculum developers” and “It is likely that creativity will have a more significant impact on pupils’ learning if the choices made to include creativity in national curricula are coherent throughout different types and sections of texts” (p 13).

An additional challenge involves how to implement something as context-driven as creativity, and as ever-changing as technology, in ways broad enough to speak to curriculum across varied settings. The variation in language and conceptualizations of creativity, the integration of creativity across disciplines, the relation between technology and creativity, and the professional development of teachers are just a few examples of complexities to consider.

**Policy/Curriculum recommendations**

- **Creativity and technology need to be featured in policy at all levels (macro / meso / micro)**
  It is clear that creativity is complex and works across all aspects of the teaching learning process, particularly when coupled with the potentials of technology. Thus it is important that educational policy emphasize creativity across all levels: macro, meso and micro, (i.e., at the level of national policy, state or school district-wide, or individual schools and classrooms). The policy texts should in turn be incorporated into other aspects of curriculum and documentation that teachers and other stakeholders use. Policy should extend beyond the document it is built into, so that it can be operationalized and instantiated throughout the education system – particularly in documents read by teachers.

- **Creativity and technology should be embedded across the curriculum**
  Creativity is not a domain by itself but a way of thinking and approach to problem solving that cuts across disciplines. Thus creativity is as important in the sciences and mathematics as it is in the arts. Technology in turn has dramatically changed the work and creative process of almost every domain of human activity. This is often forgotten and needs to be part of every policy-makers thinking. Creativity is also not a skill that is limited to few individuals. Similarly, technology is not something that is limited to a few individuals or in a few select domains. In policy and curricular documents these issues related to creativity and technology should be part for all learners, not just for the “special” or “talented” ones.

- **A greater push for research to identify models and practices**
  Though creativity research has received greater attention recently, there is much we still do not know about it in formal and informal learning contexts. The addition of technology also complicates the picture. Clearly there are models and practices that work, but more systematic research is the pressing need, both in theory and practice. The use of new technologies and their reciprocal relationship with creativity needs to be studied. We need to learn more about creativity and technology and how both can be integrated in education at all levels.

**Conclusion**

We began by reiterating the reasons for why creativity has been receiving increased attention in education. We offer a definition of creativity as being a process of developing something that is novel, effective and whole (NEW), and suggest that it is a complex skill prevalent across domains and practices. Moreover, we argue that a productive way of thinking about creativity not only considers what it is but also where it is located. In this a systems view of creativity captures the complexities of identifying creativity. The advent of new technologies can initiate, stimulate, broaden and expand how we think about creativity systemically. New digital and networking technologies with their dual affordances of ease of creation and ease of sharing complicates the standard systems model. In a reciprocal way, technologies support creativity even as creative approaches create new ways to use (repurpose) technologies for pedagogical purposes.

Despite the increasing importance of creativity and ICT in education, neither area has had broad-based, significant impact on teaching and learning. Part of the reason for this is in the complexity of the process of integrating both into the curriculum. The inclusion of creativity is hampered by the fact that educators have to focus on both teaching
creatively and teaching for creativity. Both of which need teacher training, new approaches to creative assessment, and broader policy frameworks that support the integration of creativity in the curriculum. We have provided broad recommendations for all three of these aspects.

These recommendations are aimed across education: teachers, scholars, curriculum designers, policymakers, and researchers. It is only when all of these different stakeholders work together, for the broader goal of integrating creativity and technology in education in a system-wide manner, that we can have hope for making a change. In this way, research, practice, and policy come together. Such an alignment is necessary, if we truly believe that creativity is important for the future of education.

Acknowledgements

The authors thank the organizers of the EduSummit 2015 conference in Bangkok. It was there that many of the ideas in this article were first discussed. We also thank all the members of TWG6, the Thematic Working Group on Creativity. This group’s willingness to discuss and share ideas around creativity and technology provided the foundations for this piece. In alphabetical order, TWG6 members are: Miroslava Cernochova, Janet Cochrane, Sue Cranmer, Sacha DeVelle, Michael Henderson, Susan Hopper, Leah Irving, Eugenia Kovatcheva, Don Krug, Tim Patston, Nicholas Reynolds and Paolo Tosato.

References


Henriksen, D., & Mishra, P. (2015). We teach who we are: Creativity in the lives and practices of accomplished teachers. Teachers College Record, 117(7), 1-46.


Sternberg, R. J. (2000). An Evolutionary interpretation of intelligence, creativity, and wisdom: A Link between the evolution of organisms and the evolution of ideas. Behavioural and Brain Sciences, 23(1), 160-161


Arguing for Computer Science in the School Curriculum

Andrew Fluck1*, Mary Webb2, Margaret Cox2, Charoula Angeli3, Joyce Malyn-Smith4, Joke Voogt5 and Jason Zagami6

1University of Tasmania, Australia // 2King’s College London, UK // 3University of Cyprus, Nicosia, Cyprus // 4Education Development Center, USA // 5University of Amsterdam, Windesheim University of Applied Sciences, The Netherlands // 6Griffith University, Australia //Andrew.Fluck@utas.edu.au // mary.webb@kcl.ac.uk // mj.cox@kcl.ac.uk // cangeli@ucy.ac.cy // jmsmith@edc.org // J.M.Voogt@uva.nl // j.zagami@griffith.edu.au

*Corresponding author

ABSTRACT

Computer science has been a discipline for some years, and its position in the school curriculum has been contested differently in several countries. This paper looks at its role in three countries to illustrate these differences. A reconsideration of computer science as a separate subject both in primary and secondary education is suggested. At EDUsummIT 2015 it was argued that the major rationales for including computer science as a subject in the K-12 curriculum are economic, social and cultural. The paper explores these three rationales and also a beneficence matrix to assist curriculum designers. It also argues computer science is rapidly becoming critical for generating new knowledge, and should be taught as a distinct subject or content area, especially in secondary schools. The paper concludes by looking at some of the key questions to be considered when implementing computer science in the school curriculum, and at ways its role might change in the future.

Keywords

Computer science, Curriculum, K-12, Rationale, Primary school, Secondary school

Introduction

This paper looks briefly at the positioning of computer science in the curricula of several countries, and discusses the reasons for varying levels of priority from place to place and time to time. These reasons are then examined holistically to come to some broader sense of how and why the subject can be understood in context. From these reasons an argument is constructed to provide a rationale for including computer science in school curricula that can be tuned to fit particular contexts. Two ideas are critical to this paper: the contexts for school curricula considered in this article, and the definition of computer science. The school curricula considered in this article are primary and secondary schools, typically providing education for young people from the age of 5 to 18 years.

Computer science emerged from Babbage’s Analytical Engine with Ada Lovelace’s programming (to 1871) and from mathematics with Gödel’s incompleteness theorem (1931). Its importance as a discipline became more widely accepted during the Second World War. The first university course in computer science was established at Cambridge (UK) in 1953, and the first university department of computer science at Purdue University (USA) in 1962. Subsequently, the growth of information technology has percolated through business, education and almost every aspect of society. This paper argues for computer science to become more widely taught in schools.

Computational thinking (Wing, 2006) is often seen as a key component of computer science in school curricula, and its importance was elaborated at EduSummit 2013 (Voogt et al., 2015). Definitions and arguments were made for its inclusion in school education. However, computational thinking is generally accepted as something which can be developed in computer science and which should also be integrated across the curriculum. In so doing, computational thinking is likely to transform what is taught in other subject areas (Fluck, 2003). We argue here that computer science should be formally taught in schools and that computational thinking is just one element of computer science, albeit an important one.

What is computer science?

Computer science is the term used in this article for a particular identified subject in the school curriculum. While “subject” is a common term, this varies from one country to another, so we use the word here to mean a particular
element of a school timetable, or a section of a regular report to stakeholders on a student’s learning achievements. A subject can also be referred to as a domain or field of study, a discipline or a content area in the school curriculum.

Different names are given to around the world. Some alternatives are: informatics, digital technologies, or computing. Not all of these cover the same ground, but there is a great deal of commonality between them. The key focus for this article is on computer science (which we use for consistency throughout) as a separate subject distinct from the use of computers to support learning in other areas of the curriculum. As a separate subject, computer science has its own curriculum documents defining what students will know, the skills they will master and the attitudes they will acquire. Initially in this article we are concerned with this content rather than the way it is delivered. Later we will discuss approaches to implementation and delivery. The content will probably be assessed and achievement reported to stakeholders as an identified separate subject and may be taught in identified lessons in the school timetable but teaching arrangements may vary depending on age and other factors.

By comparison, there are general uses of computers in schools which are excluded from consideration here. Administrative uses (for recording attendance, managing resources etc.) are clearly excluded. Teacher construction of digital learning materials such as visual presentations (powerpoint slides) and other uses of office suite software are generally excluded, as is student word processing and similar activities. In particular, the general use of computers to support learning across the curriculum, often known as ICT (information and communication technology) is distinct from “computer science.”

In some countries both ICT and computer science are taught side by side. In the United Kingdom, the “ICT curriculum” was phased out before “computing” with a strong foundation of computer science, was introduced as a curriculum subject.

Here are some popular definitions of computer science:

- It seeks to answer the following questions: What is information? What is computation? How does computation expand what we know? How does computation limit what we can know? (Denning, 2007).
- The study of computers and algorithmic processes, including their principles, their hardware and software designs, their applications, and their impact on society (Seehorn, Carey, Fuschetto, Lee, Moix, O’Grady-Cunniff, Owens, Stephenson & Verno, 2011).
- The scientific and practical approach to computation and its applications. It is the systematic study of the feasibility, structure, expression, and mechanization of the methodical procedures (or algorithms) that underlie the acquisition, representation, processing, storage, communication of, and access to information (Wikipedia, 2015).

In moving forward we favour the definition from Wikipedia, reflecting a more recent consensus. Disciplines are recognised by six features; terminology, non-routine intellectual work, axioms, openness, a body of knowledge, and identified sections of universities (Fein, 1959, p. 11). Chief amongst the axiomatic foundations of computer science as a discipline are the undecidability theorems (Gödel, 1931); the Turing–Church thesis (Turing, 1936); information theory (Shannon, 1948); and more recently, the empirical Moore’s Law (1965). These theorems and their intellectual consequences have had widespread application in many other areas of intellectual endeavour. As we shall demonstrate later in this article, computer science is rapidly becoming crucial for progress in many other areas.

**How are curricula determined?**

Curriculum design is interwoven with pedagogy, and considered briefly here in order to position this current discussion within a broader context of theory and research. Key curriculum theorists such as Dewey (epistemology and personal growth), Popham (behavioural objectives), Foucault (power-knowledge), Vygotsky (social constructivism), and Bruner (psycho-cultural learning) have described the overt and hidden curriculum, and the aims, subject matter, methods and evaluation of learning. However, as in many other areas, theory and practical implementation are often in tension (Smith, 1973). Curriculum design is often a contested process, blending together what we know about educational maturation and pedagogy, social trends and political machinations tied to voting aspirations. Some view schools as resistant to change (Cuban & Tyack, 1995, p.7) and others have charted teacher change fatigue (Dilkes, Cunningham & Gray, 2014). Students spend up to 14 years in school, but curricula can be re-designed almost every five years, often with little regard to a teach-out or any phasing-in process. Any curriculum
has a political dimension, as well as technical and professional dimensions (Stenhouse, 1975). Depending upon national governance structures, education systems can be centralised or decentralised and the curricula they use can be designed top-down or bottom-up (Braslavsky, 2003).

Several drivers of curriculum development can be identified in the literature. Standards and accountability have generated a great deal of discussion (Taubman, 2009; Slattery, 2013). Entitlement is another driver, attempting to identify what a student needs to know and master in order to function in society (Young, 2013). Devising a rationale for computer science in school education requires an approach which can take account of how this range of issues and drivers may play out in different curriculum contexts. In addition to these more general considerations a major tension identified in relation to Computer Science is between the rapidity of technological change for computers and the general tendency of schools or curricula to respond slowly to such changes.

**Computer science in different countries**

We now look at three countries with differing trajectories of computer science curriculum development in schools. Cyprus has had such a subject in schools since the 1980s; the United Kingdom made a significant shift from ICT to computer science in 2014; and Australia recently endorsed its curriculum document for the subject in late 2015.

Cyprus is one of twelve European countries with computer programming and coding (as parts of computer science) as part of the curriculum as revealed by a survey for the European Schoolnet (Balanskat & Engelhardt, 2014). Computer science (in Cyprus) has been taught as a separate school subject in the secondary school grades 10-12 since 1987 and in grades 7-9 from 2001-2003. There is no distinct computer science subject in the primary school curriculum, but computers support other learning. The content of the course focuses on algorithmic thinking and programming which is compulsory for students aged 13-16. Older students can take it as an elective. Scratch and Alice are used to introduce the development cycle using drag and drop programming. Students compare the efficacy of different algorithms. Knowledge and achievement are assessed “through day-to-day assessment exercises, written assessment exercises and examinations” (Balanskat & Engelhardt, 2014, p. 19). All teachers of computer science in secondary education are required to be computer science graduates or graduates of a computing-related field.

The United Kingdom reviewed education in computing in schools from 2009-2012, culminating in a report by Furber (2012). This concluded that the positioning of ICT, as a subject, in the National Curriculum was counter-productive, and it was discontinued in 2012. In 2014 a new subject called “Computing” was introduced for students aged 5-16 (Department for Education, 2013). This encompasses computer science, information technology and digital literacy (Berry, 2013, p. 5). Students are expected to understand and apply the fundamental principles and concepts of computer science, including abstraction, logic, algorithms and data representation; to analyse problems in computational terms, and have repeated practical experience of writing computer programs in order to solve such problems; evaluate and apply information; and be responsible, competent, confident and creative users of information and communication technology, including new or unfamiliar technologies, analytically to solve problems (Berry, 2013). The curriculum specification is quite brief at three pages but assessment approaches and teaching methods are provided as advice to teachers by the subject associations (see for example, Berry, 2013). New entrants to computer science teaching are expected to demonstrate their knowledge of the subject, and at secondary level they would normally be graduates of computer science or a computing-related field. An extensive professional development program has been developed for existing “ICT” teachers to equip them for the new curriculum.

Australia is a federated country, where responsibility for education was devolved to the eight states and territories in 1901. A statutory body was established in 2009 to frame a national curriculum which began implementation in 2013 with four subjects and ICT as a general capability to be taught across and in support of the other subject areas. In November 2015 “Digital Technologies” was endorsed as an additional subject in the Technologies area, and will be gradually introduced into schools region by region, with reports on student achievement being sent to stakeholders in about 2018. Digital technologies looks at digital system components and data representation; and students learn to collect, manage and analyse data, and to create digital solutions. Computational thinking is key to the syllabus, encompassing abstraction, data collection, representation and interpretation, specification, algorithms and implementation. Methods of teaching or assessment are unspecified, but content is described and achievement standards have been set at five age levels for students aged from about 5 to 16 years old. The subject will be compulsory from about age 6 to age 14, and elective thereafter. Teacher professional development will be conducted.
by individual states and territories, but central resources such as the learning object repository SCOOTLE will probably be useful.

Given the different pathways computer science has taken in these three countries, it seems timely to examine the arguments for its inclusion as a separate subject in school curricula.

The arguments for computer science in the curriculum

At EDUsummit 2015 it was argued that the major rationales for including Computer Science as a subject in the K-12 curriculum are economic, social and cultural, partly echoing Hawkridge’s (1990) justification for using computers in schools.

The economic rationale is a strong driver for inclusion of computer science in the curriculum. The economic rationale rests not only on the need for a country to produce computer scientists to sustain a competitive edge in a world driven by technology, but also on the requirement for computer science-enabled professionals in all industries to support innovation and development. World trade in information technology products and services continues to grow. Worldwide exports of computer and information services grew between nine and 32 percent from 2005 to 2015 and imports grew from 6 to 28% in the same period (World Trade Organization, 2014a, p. 134). Computer and information services were identified as the seventh largest trade sector by value, particularly between developed countries, from 2008 to 2013 (UNCTAD, 2014, p. 7). Export growth in the sector was up to 20% in the period 2011-13 (UNCTAD, 2014, p. 11), and over the period 2005-13 was the highest of any global trade sector (WTO, 2014b, p. 25). Both Australia and Europe forecast shortages of computer scientists of 100 to 500 thousand (ACS, 2015, p. 34; Husing & Korte, 2010). Businesses in the sector compete for staff and intellectual property development against other sectors. Alongside these growth trends must be placed the increasing stress in trade negotiations on intellectual property rights and the trend for software to be protected by both copyright and patent legislation. Recent trade agreements elongate intellectual property ownership terms for up to 70 years from the death of the author (AU FTA 2013). This underlines the economic importance of computing professionals and the intellectual property they produce. The economic rationale is therefore compelling for governments devising national curricula, and who see future trade prospects dependent upon computers. Computers have also made previously un-tradable services such as banking become more globally mobile. This makes sourcing local talent more crucial in controlling and building the local national economy.

This leads to the social rationale for computer science. The social rationale emphasises the value in society of active creators and producers rather than passive consumers of technology. Such capability provides people with power to lead, create and innovate within society and therefore is also an issue of entitlement to “powerful knowledge” (Young, 2013) giving individuals opportunities to choose their role in society. Since so much of our lives depends upon computers, surely students should become confident in their use? The social rationale goes further, and discusses the impact of technologies on societies. Computing makes possible new forms of governance (distributed daily or enhanced democracy) and redeﬁnes our sense of privacy (Tsohou, Lee, & Irani 2014). Robots comfort senior citizens in senior care institutions (Sharkey & Sharkey, 2012), and taxis or trucks cease to need drivers (Ticoll, 2015, p. 48). Technology can change our sense of ethical behaviour. Rather than being oppressed by innovation shock, a society equipped with its own creative proponents of new ideas is more likely to sift them and control their impact. This is a strong argument for teaching computer science in schools, making it possible to generate new content and applications locally, and enhancing the agility of each nation to deal with inevitable changes in the world. Competition for resources with other nations is just one such change pressure, so falling behind is not a desirable option. In some ways the social rationale overlaps the economic rationale, but the emphasis is on attitudinal aspects rather than knowledge or skill.

The cultural rationale for computer science in schools looks at the integrity of local values and customs. The cultural rationale enables people to be drivers of cultural change, rather than having change imposed or mediated through technological developments from outside their society (Webb et al., 2015, p.61). For example, Hollywood produces many movies every year which are successfully exported all over the world. Embedded in these films are language, customs, attitudes, ethical values and mores which reflect the USA context. Surprisingly, these attributes are not universally shared, and hence other centres of production have emerged, for example, Bollywood, which made twice as many films in 2011 (Matusitz & Payano, 2011). Nollywood, films made in or for Nigeria and Africa, is also
developing (Oyeniya, 2013). Given computer games now outsell the film industry, there is additional reason for local creative talent to provide competition for imported embedded values and language. In 2014 the global video game market was estimated to be worth $US102 billion, compared to global box office receipts for films of $US36 billion (van der Meulen, 2013; MPAA, 2014, p. 2). The local talent pool becomes more imperative when we take into account the greater cognitive engagement in computer games, and therefore their greater capacity to transmit culture.

In addition, social media are connecting groups of people across the globe in personal ways almost instantaneously. The social and semantic webs facilitate data creation, contribution and sharing on an unprecedented scale. Online shopping is changing the structure of production and distribution of goods. Friendships are made online (Ofcom, 2008). Therefore governments are motivated to ensure national cultural values are preserved through locally educated computer science professionals participating in the arts and related fields.

Alongside these three rationales, we can frame the inclusion of computer science in school curricula in a broader context in order to give curriculum designers greater scope to determine its feasibility and desirability in schools curricula. We suggest two dimensions for evaluating the contribution of computer science. The first dimension is the beneficial context: the individual learner; the society in which they live; humanity and the ecology upon which we all depend; and the wider universe. The other dimension concerns the period for the benefits of the learning to be experienced: immediately; the lifetime of the individual learner; years within a social system transformation; the expected duration of humanity; or the lifetime of the universe.

These dimensions are deliberately drawn large to ensure we do not omit any potential benefit from the proposed learning.

We also posit computer science as necessary for education because of its increasing importance for knowledge generation in many areas of human endeavour. Depriving school students of the terminology, the axioms and other components of the discipline would impair their capacity to understand or contribute to this new knowledge. Examples include the Nobel prizes for chemistry in 1998 and 2013 (Nobelprize.org, 2015) where computers were fundamental; astro-physics looks at information flows in black holes (Braunstein et al., 2013); data mining helps solve problems in medicine (Yavlinsky, 2015) and parallels have been drawn between the entropies of energy and of information in the universe (Avery, 2003; Shannon, 1948, p. 11). Not only is computer science heralding new developments in chemistry, physics and biology, but data science is providing new methods for knowledge discovery. These new methods of Data Science compare with the Scientific Method as shown in Table 1.

<table>
<thead>
<tr>
<th>Scientific method</th>
<th>Data science method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Observe phenomena</td>
<td>1. Become aware of a problem</td>
</tr>
<tr>
<td>2. Propose a hypothesis</td>
<td>2. Identify data sources</td>
</tr>
<tr>
<td>3. Design an experiment as a fair test of the hypothesis</td>
<td>3. Propose an automated analysis process</td>
</tr>
<tr>
<td>4. Collect data, analyse results</td>
<td>4. Test and validate analysis process</td>
</tr>
<tr>
<td>5. Confirm or modify the hypothesis, start over</td>
<td>5. Share results, start over</td>
</tr>
</tbody>
</table>

Data science links to machine learning to make new discoveries (Phoboo, 2014; Adam-Bourdarios et al., 2015, p.27) and lies at the heart of many recent global businesses such as eBay, Amazon, Facebook, Twitter and Spotify (Gershkoff, 2015). Therefore the link between computer science and innovation makes a strong case for inclusion in the school curriculum.

There are other immediate benefits to students learning computer science. Some examples illustrate these benefits which strengthen the argument:

Coding is about thinking. Putting a process into a particular code (writing a program for a computer) requires precision. This is analogous to the precision required in literacy skills of writing, but while people can understand intent or infer spelling, computers are far more literal. Therefore a child skilled at coding, may by transference, be more precise in their thought and have greater capacity to communicate. The beneficence is immediate and personal to the learner.

Generally, with modern personal computers, students can get fairly immediate feedback on the accuracy of their initial coding attempts. Naturally, in the learning process, some early efforts result in less than complete success – or
failure. Handled well within the classroom, this can be an opportunity to build resilience. The nature of the failure (perhaps printing out all but the last number in a sequence) can indicate the corrective editing needed to achieve success. This is an important educative step of understanding for many students. In this case the beneficence may not be so immediate, and affects both the learner and the society in which they live.

The strength of the arguments for including computer science into school curricula will vary from nation to nation, and time to time. In the second decade of the twenty-first century, there are strong economic and social reasons for its adoption, culminating in an entitlement perspective benefiting both the individual and the society in which they live.

Implementation of computer science as a subject

Literacy and numeracy are often woven into many subject areas of the school curriculum, creating themes which are often the focus for much primary (elementary) schooling. Therefore there is a suggestion that computer science could be integrated into other thematic programs of work, particularly in this primary sector. After all, most curriculum specification documents identify what should be learned, not how it should be organised and taught. Arguments for the integration of computer science need to be looked at in context. If teachers are knowledgeable, confident, fully trained in computer science and have adequate resources, it would be very difficult to mount an argument against integration. However, for those countries introducing or re-introducing the subject into school curricula, an integration policy could mask the extent to which the new content is being mastered by students. A large scale study of 37 countries (Plomp et al., 2009) concluded that integration of ICT use into other subjects was spasmodic and ineffective. Voopt and ten Brummelhuis (2014) illustrated how information literacy in the Netherlands disappeared as a separate subject, because of poorly trained teachers and a vague place in the curriculum. These examples indicate computer science could have a similar fate if only taught through integration.

There are several challenges for schools where computer science is (re-)introduced. New pedagogical content knowledge may be required, as well as teaching resources such as programming development environments and textbooks or their equivalents. While computers themselves are getting less expensive, access and particularly individual access, cannot be assumed. A great deal of computer science can be taught “unplugged” (Bell & Newton, 2013) but greater alignment with the three main rationales is likely to be greater if computers or BYOT (bring your own technology) like smartphones are available.

Another tension is the rapid rate at which innovations are changing our understanding of computer science. To become entrenched in schools, the subject needs clear specifications for learning outcomes, assessment and standards which go beyond the tendency to have the curriculum build around technology innovations instead of a deep understanding of computer science (Webb et al., 2015, p.64). Whilst we may need to live with a continually renegotiated curriculum (Webb, 2014), this is not easy to communicate and interpret between the policy level of education systems and implementation in the classroom. France is planning an integrated approach (Balanskat & Engelhardt, 2014, p.17) so will be worth monitoring.

Several other challenges and solutions were considered at EduSummit 2015 (Webb et al., 2015), which made recommendations for policy makers, educators, industry partners and researchers.

Into the future

Many things in education oscillate in the style of a pendulum. Since the year 2000 computer science has been less popular, but in 2015 it is becoming resurgent. Several countries are re-introducing computer science (under this or a similar subject title) from 2015 onwards. Australia, USA, UK, have committed or initiated computer science as a subject, whilst the Czech Republic, Denmark, Lithuania, Poland and the Netherlands are in the process of doing so, mostly from the first classes of primary school.

This brings us to some of the more important questions educators need to answer with respect to computer science as a separate subject in the school curriculum. Some of these questions were raised at EduSummit 2015 and many of
them interact with one another. Winch’s (2013) concept of “epistemic ascent” is key to understanding how these questions can be used to frame computer science curriculum design:

- What is the range of skills and understanding that should be developed in computer science?
- Are such skills and understanding necessary for everyone? Should it be and remain compulsory?
- At what age should computer science education commence?
- How many computing languages or frameworks should a student be exposed to in the span of schooling from K-12?
- How varied should these languages be? Should a variety of paradigms be explored?
- How closely should the curriculum match computers available to schools and students?
- What consideration in curriculum design should be given to emerging technologies such as quantum networks and computing?
- What pedagogical approaches are likely to be appropriate, and how do they vary with age and other factors?

Finally, to bring together some of the issues briefly mentioned in this paper, it may be worth making a case study from quantum computing. Let us assume all the above questions have been answered in a particular national context, and a clear specification for computer science in the curriculum has been defined and implemented. Teachers have been trained and assessments of achievement show targeted students are largely attaining the intended learning outcomes. Many current curricula focus on procedural or algorithmic thinking when devising digital solutions. However, this may not be appropriate with emerging quantum computers where the fundamental components are qubits which can be in a probabilistic superposition of several states simultaneously. How can our agile curriculum design accommodate such a radical departure from established content knowledge? Touted applications for quantum computers include cryptanalysis, typically by factoring very large numbers using the Shor algorithm (1997). The essential approach is to try all possible solutions simultaneously. This has parallels in the declarative programming paradigm where code specifies what computation should be performed, and not how to compute it. Prolog was once a typical declarative language, and functional languages such as Haskell may be similarly suitable. This makes an argument for computer science curricula to have an eye to the future, with content chosen to some extent with awareness of emerging technologies.

References


A K-6 Computational Thinking Curriculum Framework: Implications for Teacher Knowledge

Charoula Angeli*, Joke Voogt2, Andrew Fluck3, Mary Webb4, Margaret Cox4, Joyce Malyn-Smith5 and Jason Zagami6

1University of Cyprus, Cyprus // 2University of Amsterdam, The Netherlands // 3University of Tasmania, Australia // 4King’s College London, UK // 5Education Development Center, USA // 6Griffith University, Australia // 
cangeli@ucy.ac.cy // J.M.Voogt@uva.nl // Andrew.Fluck@utas.edu.au // mary.webb@kcl.ac.uk // mj.cox@kcl.ac.uk // j.zagami@griffith.edu.au

*Corresponding author

ABSTRACT

Adding computer science as a separate school subject to the core K-6 curriculum is a complex issue with educational challenges. The authors herein address two of these challenges: (1) the design of the curriculum based on a generic computational thinking framework, and (2) the knowledge teachers need to teach the curriculum. The first issue is discussed within a perspective of designing an authentic computational thinking curriculum with a focus on real-world problems. The second issue is addressed within the framework of technological pedagogical content knowledge explicating in detail the body of knowledge that teachers need to have to be able to teach computational thinking in a K-6 environment. An example of how these ideas can be applied in practice is also given. While it is recognized there is a lack of adequate empirical evidence in terms of the effectiveness of the frameworks proposed herein, it is expected that our knowledge and research base will dramatically increase over the next several years, as more countries around the world add computer science as a separate school subject to their K-6 curriculum.

Keywords

Computational thinking curriculum, Pedagogical content knowledge, Technological pedagogical content knowledge, Teacher preparation, K-6

Introduction

In a world in which digital technology plays an important role in carrying out essential daily-life tasks, it is imperative individuals have the education, knowledge, and skills to critically understand the technological systems they use, as well as to be able to troubleshoot and problem solve when things go wrong (Wing, 2006; Czerkawski, 2015; National Research Council, 2010). Czerkawski (2015) argues the knowledge that individuals need to have in order to competently respond to the challenges of the 21st century goes beyond the acquisition of mere skills with immediate application, to knowledge with long-term value that will enable them to understand the basics of computer structures and practices. In essence, the society needs citizens who understand the true affordances of computers in terms of what they can and cannot do, so they themselves become effective authors/creators of computational tools. Wing (2006) broadened the idea of computation, and proposed that computational thinking should be considered as a basic skill taught across the curriculum. She defined computational thinking as the thought process of formulating and solving problems with the use of computers. According to Wing (2006), the teaching of computational thinking, as a basic skill across the school curriculum, will enable K-12 students to learn abstract, algorithmic and logical thinking, and be prepared to solve complex and open-ended problems.

How do we then prepare our students to develop the knowledge they need to survive and effectively cope with the technological challenges of the 21st century? As many educators strongly argued, this educational goal can be achieved by integrating computer science as a distinct discipline and a school subject in the K-12 curriculum (Barr & Stephenson, 2011; Fluck, Webb, Cox, Angeli, Malyn-Smith, Voogt, & Zagami, 2016; Goode, Chapman, & Margolis, 2012; Hazzan, Lapidot, & Ragonis, 2011; Tucker, Deek, Jones, McCowan, Stephenson, & Verno, 2003). Fluck et al. (2016) stated that there is a strong case for integrating computer science in the K-12 curriculum with arguments from both the educational and economic sectors. Succinctly, the educational case asserts that computer science: (a) develops and promotes a unique way of thinking about problems, namely computational thinking, that uses the power of logic, algorithm, abstraction, and precision; (b) empowers individuals to create new artifacts and to move from being consumers of technology to producers of technology; and (c) redefines the way learners think about other disciplines, and, this can have a major impact on teaching practices, such as, for example, interdisciplinary teaching in school. The economic case stresses the critical shortage of applicants in IT-related jobs, especially in Europe,
while at the same time the European Commission predicts that major European countries, such as UK, will need an additional 500,000 IT professionals by 2015 (Husing & Korte, 2010).

Adding computer science as a separate school subject to the core K-12 curriculum is, however, a complex issue that involves many legislative, administrative, political, and educational challenges. The latter are the focal point of this paper. In particular, there are two major educational challenges related to: (a) what computer science content to teach across different educational levels, and (b) what body of knowledge do teachers need to have to be able to teach the computer science curriculum. Over the years, a variety of computer science curricula, representing different views about what is important to teach in computer science and when, have been proposed in the literature and or enacted in different countries, such as UK, USA, Austria, Germany, Mongolia, Israel, Greece, Cyprus, and recently Australia. Well-known efforts in the United States are, amongst others, the Computer Science Principles, Exploring Computer Science, Beauty and Joy of Computing, Project Lead the Way (PLTW), and Code.org. Computer Science Principles is part of a larger national effort in the United States, namely the CS 10K Project that aims to develop effective high school computing curricula enacted in 10,000 high schools taught by 10,000 well-prepared teachers by 2016. Computer Science Principles constitutes a framework of standards from which high school computer science courses can be built (Astrachan & Briggs, 2012). The framework is specified through a set of six Computational Thinking Practices (i.e., connecting computing, developing computational artifacts, analyzing problems and artifacts, abstracting, communicating, and collaborating), and a set of seven Big Ideas of computer science (i.e., creativity, abstraction, data and information, algorithms, programming, Internet, and global impact), and has been adopted by several high schools in the United States for developing computer science courses, such as the Beauty and Joy of Computing, Code.org, and PLTW (Astrachan & Briggs, 2012; desJardins, 2015). The Beauty and Joy of Computing course focuses on the Big Ideas of computing, and its main objective is to expose students to the beauty and joy of programming by engaging them in meaningful projects using the Snap! programming language. Similarly, Code.org is a high school course with lessons and programming projects around the seven Big Ideas of computing as well, whereas, the PLTW uses Python as its primary programming environment to expose students to different computational thinking projects.

Analogously, in various other countries similar initiatives have also been undertaken for introducing computer science to high school students (van Diepen, Perrenet, & Zwaneveld, 2011; Micheuz, 2008; Furber, 2012). Undoubtedly, during the last two decades, a lot of work has been done by the computer science education community in promoting computer science as a school subject in secondary education. Unfortunately, the same conclusion cannot be reached about the status of computer science in the elementary school curriculum (grades K-6, approximately from 6 to 12 years old).

A number of computer science education researchers have written about their concerns in regards to teaching computer science in K-6 (e.g., Armoni, 2012). These concerns are primarily linked to the incompatibility between abstraction, an essential process in computer science, and children’s weakness to understand abstraction because of their very young age. Armoni (2012) explained that abstraction is an inherent component of computer science that is always encapsulated during the process of thinking about and automating a solution to a problem. From a Piagetian perspective, children before the age of seven cannot really understand concrete logic, whereas children between seven and eleven years old can solve problems that apply to concrete objects, but not problems that apply to abstract concepts or phenomena. Conversely, Gibson (2012) argued that high school is too late for exposing students to computer science for the first time, and stated that early exposure during kindergarten is necessary. In his research, Gibson (2012) found that young children can think abstractly when concrete reference systems are used to situate their thinking.

Recently, there has been much impetus in bringing computer science experiences to elementary school children (Kumar, 2014). Kumar (2014) wrote about the proliferation of app development startup companies that have targeted “early childhood computing education as the next emerging frontier” (p. 52), and about formal deliberative initiatives for developing computer science curricula for K-6 students. Succinctly, we acknowledge the effort by Prottsman (2014) who reported on the development of the Thinkersmith curriculum in 2011, which introduced a stand-alone set of unplugged activities for K-8 specifically designed to provide students with strong computer science foundations without using computers. Lessons in this curriculum, such as Binary Baubles, used materials found in games and crafts to teach authentic computer science concepts. In 2013, Code.org expanded on what ThinkerSmith created, and offered a 20-hour unplugged curriculum for grades K-8. After the wide adoption of this curriculum, in 2015 Code.org developed further the existing 20-hour unplugged curriculum, which now includes
more than 55 lessons. CS Unplugged, another unplugged computer science (CS) approach proposed by Bell, Witten, Fellows, Adams, and McKenzie (2015), is a collection of activities that teach computational thinking through engaging games and puzzles that use cards, string, crayons and lots of physical movement. Students learn about binary numbers and algorithms without using computer programming.

Clearly, early computing education is now at the forefront, and, studies toward this line of research are urgently needed in order to develop an informed body of knowledge about learning and teaching computer science in K-6. Accordingly, the authors propose a curriculum framework with a focus on promoting computational thinking skills for ages 6 to 12. While computational thinking is just one element of computer science, albeit an important one (Fluck et al., 2016), the authors suggest a curriculum for K-6 with an explicit focus on computational thinking, before covering more theoretical and applied concepts of computer science in secondary education. Particularly, this study sought to address the following questions: (a) what computational thinking skills should a curriculum promote in K-6? and (b) what knowledge do teachers need to have to be able to teach a computational thinking curriculum in K-6?

**A definition of computational thinking**

While the concept of computational thinking in education can be traced back to the work of Seymour Papert (Papert, 1980), Wing’s (2006) article has rekindled the interest for promoting computational thinking in K-12. Other efforts aiming at developing a definition for computational thinking include, among others, the National Academy of Sciences workshop (National Research Council, 2010), the initiative undertaken by Furber (2012), and workshops organized by the Computer Science Teachers Association (CSTA) and the International Society for Technology in Education (ISTE).

Succinctly, the 2010 National Research Council’s report differentiated computational thinking from computer literacy, computer programming, and computer applications (i.e., games), and broadened the term to include core concepts from the discipline of computer science, such as abstraction, decomposition, pattern generalization, visualization, problem-solving, and algorithmic thinking.

Similarly, Furber (2012) offered a concise definition of computational thinking as “the process of recognizing aspects of computation in the world that surrounds us, and applying tools and techniques from computer science to understand and reason about both natural and artificial systems and processes” (p. 29).

CSTA and ISTE, in collaboration with leaders from higher education, industry, and K-12 education, developed an operational definition of computational thinking as a problem-solving process that includes, but is not limited to, the following elements: (a) Formulating problems in a way that enables us to use a computer and other tools to help solve them; (b) Logically organizing and analyzing data; (c) Representing data through abstractions, such as, models and simulations; (d) Automating solutions through algorithmic thinking (i.e., a series of ordered steps); (e) Identifying, analyzing, and implementing possible solutions with the goal of achieving the most efficient and effective combination of steps and resources; and (f) Generalizing and transferring this problem-solving process to a wide variety of problems.

Despite the fact that currently there is not one unanimous definition of computational thinking, it seems fair to conclude that, based on the literature reviewed in this study, researchers have come to accept that computational thinking is a thought process that utilizes the elements of abstraction, generalization, decomposition, algorithmic thinking, and debugging (detection and correction of errors). Abstraction is the skill of removing characteristics or attributes from an object or an entity in order to reduce it to a set of fundamental characteristics (Wing, 2011). While abstraction reduces complexity by hiding irrelevant detail, generalization reduces complexity by replacing multiple entities that perform similar functions with a single construct (Thalheim, 2000). Abstraction and generalization are often used together as abstracts are generalized through parameterization to provide greater utility. Decomposition is the skill of breaking complex problems into simpler ones (National Research Council, 2010). Algorithmic thinking is a problem-solving skill related to devising a step-by-step solution to a problem and differs from coding (i.e., the technical skills required to use a programming language) (Selby, 2014). Additionally, algorithmic notions of sequencing (i.e., planning an algorithm, which involves putting actions in the correct sequence), and algorithmic notions of flow of control (i.e., the order in which individual instructions or steps in an algorithm are evaluated) are
also considered important elements of computational thinking (Selby, 2014). Debugging is the skill to recognize when actions do not correspond to instructions, and the skill to fix errors (Selby, 2014).

Table 1 shows the elements of computational thinking as these have been discussed and defined in this section. Accordingly, this conceptual framework is the one that was adopted for developing the computational thinking curriculum framework for K-6 presented in the next section.

<table>
<thead>
<tr>
<th>Element</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Abstraction</td>
<td>The skill to decide what information about an entity/object to keep and what to ignore (Wing, 2011).</td>
</tr>
<tr>
<td>2. Generalization</td>
<td>The skill to formulate a solution in generic terms so that it can be applied to different problems (Selby, 2014).</td>
</tr>
<tr>
<td>3. Decomposition</td>
<td>The skill to break a complex problem into smaller parts that are easier to understand and solve (National Research Council, 2010; Wing, 2011).</td>
</tr>
<tr>
<td>a. Sequencing</td>
<td>The skill to put actions in the correct sequence (Selby, 2014).</td>
</tr>
<tr>
<td>b. Flow of control</td>
<td>The order in which instructions/actions are executed (Selby, 2014).</td>
</tr>
<tr>
<td>5. Debugging</td>
<td>The skill to identify, remove, and fix errors (Selby, 2014).</td>
</tr>
</tbody>
</table>

A computational thinking curriculum framework for K-6

Based on the five computational thinking skills shown in Table 1, a computational thinking curriculum framework is developed and presented in Table 2. Table 2 shows indicators of competence for all five computational thinking skills, namely, abstraction, generalization, decomposition, algorithmic thinking, and debugging, in a progression from simple to complex across the educational levels of K-2, 3-4, and 5-6. Succinctly, the framework aims at engaging children in thinking and problem solving by developing a solution to a problem, automating the solution through algorithmic thinking, and generalizing this solution to new problems when common patterns are identified or recognized. In essence, the framework aims at introducing students of a very young age to the thinking processes of computational thinking so they become competent to learn more advanced theoretical and practical topics of computer science in secondary education. In addition, the framework targets the development of all five computational thinking skills across all K-6 levels, albeit at different levels of competence, through the use of examples and tasks that are within the reach of children either with or without external support (external reference systems). It is noted that the boundaries specified for each level may possibly vary from school to school and from classroom to classroom. By the same token, it is also expected that refinements to the curriculum framework will be ongoing once data become available from pilot offerings of different curricula, aligned with the proposed framework, in diverse contexts.

Table 2. A computational thinking curriculum framework for K-6

<table>
<thead>
<tr>
<th>Skill</th>
<th>Grade level (age level)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K-2 (ages 6 to 8)</td>
</tr>
<tr>
<td></td>
<td>3-4 (ages 9 to 10)</td>
</tr>
<tr>
<td></td>
<td>5-6 (ages 11 to 12)</td>
</tr>
<tr>
<td>Abstraction</td>
<td>• With the use of external reference systems, create a model/representation* to solve a problem (i.e., using specific directional language - forward, left turn, right turn, back - and turns of a given degree (90, 180, 270, 360), children create a path and write instructions to enable others to follow the path, or children design a mat based on a story, and have their Bee-Bot</td>
</tr>
<tr>
<td></td>
<td>• Create a model/representation to solve a problem (i.e., create an object and assign properties to it during an activity of digital game design and creation).</td>
</tr>
<tr>
<td></td>
<td>• Create a new model/representation to solve a problem (i.e., create a simulation using Scratch).</td>
</tr>
</tbody>
</table>
follow the path from the narrative).

**Generalization**
- Identify common patterns between older and newer problem-solving tasks, and use sequences of instructions previously employed, to solve a new problem (i.e., use a sequence of instructions from an older path, to program the Bee-Bot to follow a new path that includes the older path).
- Remix and reuse (by extending if needed) resources that were previously created.
- Remix and reuse (by extending if needed) resources that were previously created.

**Decomposition**
- Break a complex task into a series of simpler subtasks (i.e., break a longer path into a series of smaller paths that the Bee-Bot can follow).
- Break a complex task into simpler subtasks.
- Break a complex task into simpler subtasks.
- Develop a solution by assembling together collections of smaller parts.
- Develop a solution by assembling together collections of smaller parts.

**Algorithmic thinking**
- Define a series of steps for a solution.
- Put instructions in the correct sequence.
- Define a series of steps for a solution.
- Put instructions in the correct sequence.
- Repeat the sequence several times (iteration).
- Repeat the sequence several times (iteration).
- Make decisions based on conditions.
- Store, retrieve, and update variables.
- Formulate mathematical and logical expressions.

**Debugging**
- Recognize when instructions do not correspond to actions.
- Remove and fix errors.
- Recognize when instructions do not correspond to actions.
- Remove and fix errors.
- Recognize when instructions do not correspond to actions.
- Remove and fix errors.

*Note.* *model/representation* = can be conceptual, mathematical, mechanical, textual, graphical, etc.

**Curriculum design issues: A focus on a holistic design approach**

The framework presented in Table 2 constitutes a general framework that can be used to develop various computational thinking programs, courses, or modules in K-6. The curriculum framework is conceptualized in a generic form to allow teachers the freedom and agency to adapt and customize the framework as they see fit for their own classrooms and students. According to van den Akker (2010), this enactment perspective, where teachers create their own curriculum realities, is increasingly replacing the fidelity perspective on implementation where teachers faithfully follow curricular prescriptions from external sources. Accordingly, this trend “puts even more emphasis on teachers as key people in curriculum change” (van den Akker, 2010, p. 185), underlining the utmost importance of relevant teacher preparation. In view of that, the authors herein propose the holistic design approach as one method that teachers can use to enact the computational thinking framework proposed in this paper.

A holistic design approach attempts to “deal with complexity without losing sight of the separate elements and the interconnections between those elements” (van Merriënboer & Kirschner, 2007, p. 6). It is the opposite of an atomistic design where complex contents and tasks are reduced to simpler elements, promoting this way content compartmentalization and fragmentation. Compartmentalization and fragmentation support the separation of a whole
into small, distinct, and often isolated parts. For example, teachers teach children to think computationally by teaching them abstraction, then decomposition, followed by generalization, algorithmic thinking, and debugging. It is doubtful if in the end children will have the opportunity to practice the whole complex skill (computational thinking, in this case) in its entirety, and doubtful if they will ever learn to think computationally. On the other hand, a holistic design approach aims at eliminating compartmentalization and fragmentation by focusing on whole complex and authentic learning tasks, without losing sight of the individual elements that make up the complex whole. Thus, with this approach, if implemented correctly by the teacher, children learn to think computationally to solve a problem, and also learn all other constituent and interconnected pieces of knowledge (theoretical and or practical) that are directly related with the computational thinking task. We support the holistic design approach for teaching computational thinking and emphasize here two design steps in the process, namely, (a) the design of problem-solving tasks with a focus on real-life issues, and (b) the sequencing of problem-solving tasks from simple to complex. We do acknowledge that more design steps exist in the literature.

With regard to the first design step, it is argued that the sources of the computational thinking curriculum ought to be problems, issues, and concerns directly related to life itself. A curriculum of this kind will result in usable knowledge - that is, knowledge that can be applied directly in the context of real life, problems and concerns at hand - and not inert knowledge (Voogt, Fisser, Good, Mishra, & Yadav, 2015; Webb, Fluck, Cox, Angeli-Valanides, Malyn-Smith, Voogt, & Zagami, 2015). Educational researchers have found that a curriculum that is focused on problem solving around real-world problems can result in greater intellectual curiosity, motivation, improved attitude toward schooling, and higher achievement in college (Wolf & Brandt, 1998). Consequently, a curriculum designed around real-life problems can be a way to make computational thinking relevant to students’ lives, and, thus, a way to keep them interested in the subject matter. Ultimately, this may end up in increasing substantially the number of students who will eventually pursue computer science as their major field of study later in college.

From an implementation point of view, a curriculum designed around real-life problems demands a wider range of content, simply because authentic real-world problems are usually multidisciplinary in nature. As a consequence, a curriculum from this perspective poses new demands on teaching often requiring close collaboration among teachers with different content expertise. It should be noted that real-life problem-solving tasks constitute challenging design endeavors, and, a curriculum designer may approach the design process through the means of rapid prototyping before designing an entire educational program, course, or module.

With regard to the sequencing of the problem-solving tasks, a sequence from simple whole tasks to more complex whole tasks is recommended. It is made clear that each problem-solving task, irrespective of complexity, engages the learner in whole-task problem-solving experiences. In the context of computational thinking, this means that each learning task, simple or complex, confronts the learner with all or almost all of the constituent computational thinking skills for a real-life computational thinking experience. All tasks are meaningful, authentic, and relevant to children’s life. A sequence of tasks constitutes the backbone of the computational thinking curriculum. It is also evident that children may need guidance and support as they start working on more challenging tasks. Support may be provided in the form of external reference systems to help students gradually develop abstractions. Students may also need guidance with the problem-solving process itself.

**The knowledge that teachers need to teach the curriculum**

As Gal-Ezer and Stephenson (2010) stated, having a curriculum is important, but preparing teachers to teach the curriculum is also critical. Amongst computer science teacher educators, the framework of pedagogical content knowledge (PCK) has been highly regarded as an appropriate framework for defining the knowledge teachers need to have to be able to teach computer science (e.g., Hubwieser, Magenheim, Mühling, & Ruf, 2013; Saeli, 2012). Succinctly, PCK refers to a body of knowledge, which is highly context sensitive, cannot be conceptualized in isolation from children’s classroom and teaching experiences, and is beyond and above a simple synthesis of knowledge of subject matter and pedagogy (Shulman, 1986; Shulman, 1987). PCK is an amalgam of knowledge that “embodies the aspects of content most germane to its teachability” (Shulman, 1986, p. 9), and refers to the transformation of content into forms that are understandable to learners. According to van Driel and Berry (2012), having a good PCK means that teachers have several representations of the most commonly taught topics within a certain subject. The more representations teachers have at their disposal and the better they recognize learning difficulties, the more effectively they can deploy their PCK (van Driel & Berry, 2012).
Within the domain of computer science, a number of computer science education researchers attempted to define PCK for computer science, either in general ways (Hubwieser et al., 2013; Saeli, Perrenet, Jochems, & Zwaneveld, 2011; Stephenson, Gal-Ezer, Haberman, & Verno, 2005) or more specific ways (Saeli, 2012). Saeli et al. (2011) concentrated on the teaching of programming in secondary education, and provided a general conceptualization of PCK for the domain of programming in terms of its constituent elements (i.e., what to teach about computer programming, how to teach programming, and what are learners’ difficulties in programming). In a following study, Saeli (2012) was able to provide a more specific conceptualization of PCK for the domain of programming in the context of secondary education, which included details about each constituent knowledge base. In terms of the content to be taught, she mentioned loops, data structures, arrays, problem-solving skills, decomposition, parameters, and algorithms amongst others. Regarding teachers’ pedagogical knowledge she mentioned offering a simple programming language to better facilitate students’ effort to learn the syntax of the language, and choosing several worthy problems to solve. Lastly, she identified learners’ difficulties about different programming concepts, such as loops, arrays, variables, and general problem-solving skills.

In the early 2000s, though, a number of educational researchers undertook systematic efforts for extending and enriching the concept of PCK by adding Technology Knowledge as another essential category of teachers’ knowledge base (Angeli & Valanides, 2005; Koehler & Mishra, 2008; Niess, 2005). From this perspective, the introduction of Technology Knowledge in the existing framework of PCK successfully expanded PCK to TPCK - that is, Technological Pedagogical Content Knowledge (Angeli & Valanides, 2005; Angeli & Valanides, 2009; Koehler & Mishra, 2008; Niess, 2005). A conceptualization of the framework of TPCK is proposed by Angeli and Valanides (2005; 2009) as shown in Figure 1. According to Figure 1, TPCK is conceptualized as a unique body of knowledge that is formed by the contribution of five distinct knowledge bases, namely, content knowledge, pedagogical knowledge, knowledge of learners, knowledge of the educational context, and technology knowledge (Angeli & Valanides, 2005; Angeli & Valanides, 2009). This body of knowledge grows when teachers are engaged systematically in useful educational practices, either in their own classrooms or teacher professional development programs.

**Figure 1. Technological Pedagogical Content Knowledge (adopted from Angeli & Valanides, 2005)**

TPCK is an important body of knowledge for the field of computer science, because technology is at the center of the computer science domain, either, as a means in itself (i.e., to learn to use the technology as a goal), or as a means for achieving or teaching something else (i.e., to use technology in order to solve a problem or to teach a computer science concept). For the purposes of this study, the authors provide a conceptualization of TPCK for the construct of computational thinking, as it is defined in Table 1, in order to better explain what teachers need to know to be able to teach a computational thinking course aligned with the framework proposed in Table 2.
Analytically, content knowledge (CK) is defined as knowledge about computational thinking (CK\textsubscript{CT}). This includes knowledge and understanding about the skills of abstraction (including modeling), denoted as CK\textsubscript{CT(A)}; generalization, denoted as CK\textsubscript{CT(G)}; decomposition, designated as CK\textsubscript{CT(D)}; algorithmic thinking, designated as CK\textsubscript{CT(Algo)}; and debugging, denoted as CK\textsubscript{CT(Debug)}. CK\textsubscript{CT(Algo)} includes knowledge of several computational thinking concepts, such as, data, processing, information, sequencing, loops, parallel processing, events, conditions, operators, variables, and dataflow of control.

Learner knowledge for computational thinking (LK\textsubscript{CT}) includes knowledge about learners’ difficulties in (a) developing abstractions that are beyond of any particular programming language or tool, denoted as LK\textsubscript{CT(A)}, (b) generalizing from one solution to another by identifying common patterns, denoted as LK\textsubscript{CT(G)}, (c) decomposing complex problems to simpler ones, designated as LK\textsubscript{CT(D)}, (d) thinking algorithmically to solve a problem (including difficulties in understanding relevant concepts, such as sequencing, loops, flow of control, conditions, etc.), denoted as LK\textsubscript{CT(Algo)}, and (e) debugging, denoted as LK\textsubscript{CT(Debug)}.

Pedagogical knowledge for computational thinking (PK\textsubscript{CT}) includes the general pedagogical knowledge applicable to all other content domains (i.e., the use of questions to promote understanding, use of examples, explanation, demonstration), in addition to knowledge about subject-specific pedagogical practices pertinent to computational thinking. PK\textsubscript{CT} is defined in terms of the following teaching tactics: (a) model how to problem solve or think about a problem in iterative and incremental ways, (b) present or explain a solution to a problem in terms of a series of steps, (c) model decision making based on conditions, (d) do something based on (and expanding) what others or you have done (reuse and remix), (e) show how a complex problem can be decomposed into simpler problems and develop a solution in increments, (f) show how to design a model before writing a computer program for solving the problem, and (g) try things out as you go and make revisions based on what happens.

Technology knowledge for computational thinking (TK\textsubscript{CT}) includes knowledge and skills about how to (a) operate/use a variety of technologies, (b) invent new technologies/tools, (c) solve a task using technical processes, methods, and tools, and (d) learn and adapt to new technologies.

Context knowledge for computational thinking (CX\textsubscript{CT}) is defined from the point of view explicated by Porras-Hernández and Salinas-Amescua (2013) who proposed to regard context knowledge along two important dimensions, namely (a) scope (macro, mezzo, and micro level context) and (b) actor (students’ and teachers’ inner and external context). Macro context is defined by social, political, technological, and economic conditions at a global level that influence the value and worth of adding computer science and computational thinking to the school curriculum. Mezzo context is defined by the social, cultural, political, organizational, and economic conditions settled in the local community and the educational institution about the value of computational thinking in children’s lives. Finally, micro context is the level that deals with in-class conditions for learning (e.g., available resources for computational thinking, available technologies, norms and policies, beliefs, expectations, teachers’ and students’ goals about computational thinking). In addition, Porras-Hernández and Salinas-Amescua (2013) argued that in order to comprehend teachers’ uses of technology, it is important to consider teachers’ and students’ (actors’) unique characteristics, as they are brought in the context as separate objects of knowledge with internal (e.g., students’ needs, preferences, misconceptions, learning difficulties, prior knowledge, teachers’ self-efficacy, pedagogical beliefs) and external contexts (e.g., ethnicity, culture, community, and socioeconomic background).

Lastly, TPCK for computational thinking (TPCK\textsubscript{CT}) is defined as knowing how to: (1) Identify a range of creative and authentic computational thinking projects; (2) Identify a range of technologies with an appropriate set of affordances in terms of providing the necessary technological means for practicing/teaching the whole range of computational thinking skills with each project; and (3) Use the affordances of technology to transform CK\textsubscript{CT} and PK\textsubscript{CT} using representations that make the overall computational thinking experience comprehensible for all learners.
The question that naturally arises at this point is: “What form should teacher preparation take, so that teachers develop their TPCK_{CT} competencies adequately?” In the next section, we provide preliminary research evidence from a teacher education course on preparing teachers how to teach computational thinking.

**Teacher preparation in developing TPCK competencies for computational thinking**

In the fall of 2015, fifteen elementary school teachers pursuing a master’s degree in instructional technology were enrolled in a course on learning how to teach computational thinking in their K-6 classrooms. All teachers were unfamiliar with computational thinking and had no prior experiences with computer programming. The teachers participated in 13 three-hour weekly meetings. The participants were engaged in hands-on design activities with the Scratch computer programming environment. The learning-by-design approach, which has been shown to be effective in contemporary teacher development studies (McKenney, Kali, Markauskaite, & Voogt, 2015), was used in the course to engage teachers in designing models of different problem situations before constructing computer programs for solving the problems.

The course instructor initially engaged the teachers in authentic problem solving by asking them to think about the city/town they were living and identify ways of how people's lives in those places could be improved. The teachers explained their thinking about possible improvements and then the instructor asked them to think about how computers could be used for solving some of the problems they identified. A brainstorming activity resulted in ten different ideas that constituted the real-life tasks that the course instructor used to teach the teachers about computational thinking. The ten tasks were sequenced from simple to complex based on the involvedness of the solution.

For each problem, teachers were taught how to create a model first before writing a computer program for solving the problem. Creating a model proved to be extremely difficult for the teachers and often times they asked their instructor for help. Early attempts in creating models resulted in models containing lots of unnecessary information, but, gradually teachers, with the help of the course instructor learned that models are abstractions of something free from inessential detail. The teachers were taught how to create models through a process that was explicitly taught to them and involved identification of the important entities of the model, their characteristics (parameters in the model), and relationships, either quantitative or qualitative, between the parameters of the entities. The teachers showed commitment in developing the best models they could possibly create, and, often times they exhibited lots of creative ideas of how to make them better.

In regards to teaching teachers computer programming, the course instructor used systematically the following pedagogical strategies: (a) decide what sprites are needed for your project, (b) decide what scripts are needed for your project, (c) organize the scripts in meaningful ways for you and others, (d) develop some code, try it out, then develop some more, (e) test and debug, and (f) build or extend on existing projects or ideas. During the programming tasks, computational concepts such as, data, processing, information, sequencing, loops, parallel processing, events, conditions, operators, variables, and dataflow of control, were explicitly explained and illustrated with lots of programming examples. Teachers had no difficulties with understanding programming concepts, even though they found the concepts of variables and conditional logic more challenging than the others.

**Concluding remarks**

In conclusion, the authors in this paper presented a computational thinking curriculum framework for designing a curriculum for K-6, an area of research that is still in its infancy, described design guidelines for enacting the curriculum framework, and defined TPCK_{CT} as the body of knowledge that teachers need to have to be able to teach the curriculum in K-6. In addition, the authors provided an example of a teacher preparation course that was specifically designed to promote teachers’ TPCK_{CT}. It is recognized that more empirical evidence in the form of rich educational cases is needed in terms of further investigating the effectiveness of the frameworks proposed herein in a variety of contexts. It is expected that with the gradual adoption of computer science as a distinct school subject in the K-6 curriculum of countries around the world, our knowledge and research base regarding the issues discussed in this paper will dramatically expand over the next several years.
References


Selby, C. C. (2014). *How can the teaching of programming be used to enhance computational thinking skills?* (Unpublished doctoral dissertation). University of Southampton, Southampton, UK.


Technology Enhanced Formative Assessment for 21st Century Learning

J. Michael Spector1*, Dirk Ifenthaler2, Demetrios Sampson3, Lan (Joy) Yang4, Evode Mukama5, Amali Warasvitarana6, Kulari Lokuge Dona7, Koos Eichhorn8, Andrew Fluck9, Ronghui Huang10, Susan Bridges11, Jiingyan Lu11, Youqun Ren12, Xiaoqing Gui12, Christopher C. Deneen13, Jonathan San Diego14 and David C. Gibson3

1University of North Texas, USA // 2University of Mannheim, Germany and Deakin University, Australia // 3Curtin University, Australia // 4Hong Kong Institute of Education, Hong Kong // 5University of Rwanda, Butare, Rwanda // 6Institute of Advanced Technological Education, Sri Lanka // 7Monash College, Australia // 8AB-ZHW, The Hague, Netherlands // 9University of Tasmania, Australia // 10Beijing Normal University, China // 11University of Hong Kong, Hong Kong // 12East China Normal University, China // 13National Institute of Education, Singapore // 14King’s College London, UK // mike.spector@unt.edu // dirk@ifenthaler.info // sampson@iti.gr // yanglan@ied.edu.hk // emukama01@gmail.com // amali57@live.com // Kulari.Lokuge@monashcollege.edu.au // keichhorn@abzhw.nl // Andrew.Fluck@utas.edu.au // huangrh@bnu.edu.cn // sbridges@hku.hk // jingyan@hku.hk // yqren@admin.ecnu.edu.cn // xqgu@ses.ecnu.edu.cn // Christopher.d@nie.edu.sg // j.p.san_diego@kcl.ac.uk // david.c.gibson@curtin.edu.au

*Corresponding author

ABSTRACT

This paper is based on the deliberations of the Assessment Working Group at EDUsummIT 2015 in Bangkok, Thailand. All of the members of Thematic Working Group 5 (TWG5) have contributed to this synthesis of potentials, concerns and issues with regard to the role of technology in assessment as, for and of learning in the 21st century. The group decided to focus primarily on formative assessment rather than summative assessment and high stakes testing. Formative assessments and feedback provide an important opportunity to support and enhance student learning. Recognizing shifts in education towards blended and online learning with problem-based and inquiry-based approaches led to considerations of technologies that could effectively support formative assessment and informative feedback to 21st century learners. The paper concludes with a summary of conclusions and recommendations of the working group to be taken up in subsequent efforts.

Keywords

Formative assessment, Inquiry-based learning, Problem-based learning, Summative assessment, Technologies for assessment

Introduction

The importance of formative assessment in student learning is discussed at various levels in the context of teaching and learning environments. With the known purpose of assessment in education being to support learning, some educational reforms clearly recognized the significance of formative assessments and feedback (Black & William, 2006; CDC & HKEAA, 2007). Brown, Hui, Yu, and Kennedy (2011) also have highlighted the importance of teacher cooperation and staff development to initiate the changes to formative assessment and feedback. In Hattie’s (2009; 2012) research on visible learning, formative assessment was rated as one of the most effective methods to encourage student achievement. Wiliam, Lee, Harrison and Black (2004) also found that integration of formative assessment into teaching resulted in enhanced student performance. Bulunuz et al. (2016) indicated that formative assessment creates opportunities to promote the development of student's skills of explaining, interpreting and reasoning. Narciss (2008) provided a comprehensive overview of the benefits of timely and informative feedback (i.e., formative assessment). The shift in pedagogy towards dynamic problem-based and inquiry-based learning is gaining increasing attention as a means to support the development of 21st century skills (Lu, Bridges, & Hmelo-Silver, 2014; P21, n.d.), and this creates a need for changes in formative assessments. The opportunities to use technology to support formative assessment for learning have also grown with the development of new learning technologies (Johnson et al., 2016; Woolf, 2010). For example, recent review identified how various technological affordances are being integrated into the next generation of problem-based designs, to enhance both online and face-to-face interactions and collaborative knowledge building (Jin & Bridges, 2014; Narciss, 2008; Roskos & Neuman, 2012). This paper discusses some of the current issues with formative assessment and considers how the tools and technologies that are available can develop new formative assessment practices to support 21st century learning (P21, n.d.). The paper concludes with trends, directions and general conclusions emphasizing the significance of formative assessment.
assessment along with recommendations for future research and development in this area consistent with the emphasis on 21st century skills and new technologies.

Formative assessments and current issues

Davis (2015) highlights the importance of formative assessment, as it is critical to a teacher’s ability to adapt lessons and check for student understanding. Under the umbrella term of assessment, formative assessment refers to the appraisals of student performance intended to help learners attain intended goals as distinguished from formative evaluation, which refers to judgments about how to improve program effectiveness as an effort evolves over time (OECD, 2005). Formative assessment is also defined as assessment for learning (Bennett, 2011) because the emphasis is on forming judgments about learners’ progress that then affects the subsequent flow of instruction, whereas summative assessments have been viewed as focusing on making judgments about how well individuals did at the end of an instructional sequence, which might be considered assessments of learning (Ecclestone, 2010).

Because formative assessments, in the form of timely and informative feedback, are aimed at helping learners improve, formative assessments can be considered a form of learning or assessments as learning (Spector, 2015). According to Sadler (1989) formative assessments involve making judgments about the quality of students’ responses and using these judgments immediately to guide and improve students’ understandings and skills. In the current context of teaching and learning, Ecclestone (2010) argued that formative assessment or assessment for learning is now considered an integral component to good teaching, student motivation, engagement and higher levels of achievement, and that view has been supported by many others (see, for example, Johnson et al., 2016; Narciss, 2008; Spector, 2015; Woolf, 2010). Also, timely and informative feedback (also known as formative assessment) is known to enhance and expedite learning (Bransford, Brown, & Cocking, 2000; Clariana, 1990; Epstein et al., 2002; Hannafin, 1982; Kluger & DeNisi, 1996; Narciss, 2008). Formative assessments are usually determined by teachers or automated learning environments, allowing them to be directly controlled by a teacher or a learning management system. However, when learning tasks involve critical thinking and complex problem solving, determining relevant feedback for learners becomes more challenging and time consuming. Sometimes, overemphasis on summative assessments such as grades, standardized test scores, comparative rankings and annual performance ratings have resulted in too little emphasis on and support for formative assessment providing individualized and constructive feedback during learning (Baker, 2007; Ecclestone, 2010; Harlen & Deakin Crick, 2002; Sadler & Good, 2006). In the presidential address to the American Educational Research Association, Eva Baker (2007) argued that tests should generally be used to help teachers and learners improve individual learning (formative assessment) and to help administrators and policy makers identify areas in need of improvement (formative evaluation). One issue, then, concerns the need for additional focus and emphasis on formative assessment.

A second issue involves sources of formative assessments. The one-way route of feedback from teacher to student is not the only way to provide feedback for students to improve their learning. Both self-assessment and peer-assessment can be meaningful forms of formative feedback. The feedback from a student about his or her own performance or feedback provided to other students about their performance is often sufficient for students to improve their work (Ecclestone, 2010; Ross, 2006). Peer- and self-assessments then comprise a major form of assessment as learning. Webb, Gibson, and Forkosh-Baruch (2013) argue that combining evidence-centered design (Mislevy, Steinberg, & Almond, 1999), formative assessment frameworks (Black, Harrison, Lee, Marshall, & Wiliam, 2004) and the use of computerized assessments can harness learning experiences for both teachers and students. The integration of evidence-centered design, formative assessment, and automation assist the development of their evaluation and validation processes through technology-enabled assessments to address simultaneously, assessment for learning and assessment of learning. Sadler (1989) argued that for students to be able to improve they must develop the capacity to monitor the quality of their own work by learning to appreciate high quality work along with the knowledge and skills to be able to compare their work against high quality work (i.e., effective self-assessment). Self-assessment can foster and strengthen meta-cognition and self-regulation skills that are important learning skills (Bransford, Brown, & Cocking, 2000).

A third issue involves the so-called 21st century skills (P21, n.d.). While 21st century skills are discussed by many educators in a variety of contexts, the Partnership for 21st century Learning (P21) provides a framework that captures most of what is being said about these skills in terms of four clusters: (a) key subjects and themes (e.g., global awareness and multiple literacies, such as financial, civic, health and environmental; (b) learning and innovation
skills (e.g., creativity and innovation, critical thinking and problem solving, communication and collaboration); (c) information, media and technology skills (e.g., information, media and ICT literacy); and (d) life-long learning and career skills (e.g., adaptability, initiative, social skills, accountability and leadership). P21 stresses the need for ongoing formative assessments throughout a learner’s education to develop those skills. Many of these skills and the need for assessing them are emphasized in other sources (e.g., Bransford, Brown, & Cocking, 2000; Johnson et al., 2016; Woolf, 2010). In this paper, we focus on a few that we believe are especially challenging and critical – namely, critical thinking and problem solving (Bulunuz et al., 2016; Spector, 2016) as they are connected with nearly all of the other P21 skills as well as with new technologies which are discussed in the next section.

There are additional issues that present challenges for formative assessment. One situation that present a challenge for timely and meaningful assessments include large and multi-grade classrooms backgrounds, especially in developing countries (Wagner, 2011). Another challenging situation involves filtering and managing the wealth of resources, data, services and technologies available at little or no costs on the Internet (Spector, 2016). These additional issues are touched upon in a subsequent section but remain ongoing issues to be addressed by researchers, developers and educators to wish to realize the full potential of formative assessment and new technologies.

**Formative assessments in technology-enhanced learning**

Ellis (2013) and others note that in the past one of the key limitations of improving formative assessment was that the vast majority of courses were delivered face-to-face and that prevented capturing the learning interactions and outcomes into a system for identification and analysis of formative feedback and assessment. However, in the 21st century and with new technologies, there are multiple opportunities to capture both performance and assessment data and analyse them to understand how students are progressing with various forms of activities and then determine what adjustments might be made to help different learners (Bichsel, 2012). Some automated systems are capable of capturing such data; however, it is the effective use of those data that will have a positive impact on learners and teachers (Spector, 2014a; Spector, 2016). Learning analytics projects that are spread across various educational institutions can provide closer identification of learner habits and key areas that need additional support for progressing with learning (Siemens & Baker, 2012). The analysis of large data sets containing information about student learning and performance has significant potential to impact formative feedback as well as the design of instruction and educational practices and policies (Spector, 2016).

Other educational technologies mentioned by the New Media Consortium and other groups include MOOCs (Massive Open Online Courses), Serious Games and Gamification (Spector, 2014a; Hooshyar et al., 2016; Johnson et al., 2016; Woolf, 2010). These new technologies have the ability to generate and make use of large sets of data. As indicated by Ellis (2013), Bull et al. (2014) and Colvin et al. (2015), these data sets can be used to represent learning patterns, areas in which students need improvement, and likely interventions to be considered by a teacher or teaching system. Making use of big data requires sophisticated learning analytics applied to performance and assessment data collected from many different learners in a wide variety of learning situations (Spector, 2014a; Spector, 2016). Moreover, formative assessments can motivate individual learners, help teachers adjust individual learning paths, and inform parents and others of progress (Spector, 2016). From a learning perspective, there are clearly links between engagement, efficiency and effectiveness (Spector & Merrill, 2008). From a design and systems perspective, there are clearly links between interaction, integration, and intelligence. When learning becomes inefficient, slow or tedious, learners tend to lose interest. Integrating resources, technologies and activities can contribute to efficiency and keeping learners engaged. Using intelligent and adaptive technologies tend to sustain interest and engagement.

In addition to providing the means to support personalized learning and a smart education system responsive to the needs of learners and teachers and their learning environments, new technologies can support key 21st century skills discussed in the previous section – notably critical thinking and problem solving. Along with the development of 21st century skills, emphasized has been placed on specific domains – notably the so-called STEM disciplines (science, technology, engineering and mathematics). That emphasis is evident in many of the sources already cited (e.g., Johnson et al., 2016; Woolf, 2010) and is reflected in new standards such as the Next Generation Science Standards that emphasis the integration of engineering into science learning (see http://www.nextgenscience.org/). Such emphasis is consistent with advances in pedagogical approaches such as anchored instruction (Bransford et al., 1990), inquiry learning (Linn & Eylon, 2011), and problem-based learning (Hmelo-Silver & Barrows, 2000), along
with the ability of new technologies (e.g., augmented and virtual reality, context aware sensors, and automated feedback mechanisms) to support those approaches in the form of inquiry- and problem-based learning. The ability of new technologies to provide support for formative assessment has risen considerably in recent years with the advent of intelligent agents (Ehimwenma, Beer, & Crowther, 2014; Greer, & Mark, 2015; Harley et al., 2015), smart devices (Shorfuzzaman et al., 2015; Spector, 2014a; Spector, 2016) and cloud-based resources (Armstrong, & Llorin, 2015). These and other developments will be taken up in a subsequent section. Next we discuss new technology-enhanced learning environments and environments that create new demands for timely and informative feedback.

Formative assessments in the form of timely and informative feedback can support effective problem-based and inquiry-based learning that is especially sensitive to and dependent on formative feedback to learners (Salomon & Globerson, 1987; Shute, 2007). A particular approach to such support in a dynamic and online learning context will be discussed in the next section. That approach involves eliciting how learning is thinking about a challenging problem or inquiry process, and then generating both directive and suggestive feedback automatically to help the learning succeed in finding a suitable resolution to the problem situation. Before discussion that and other new approaches, the issue of open resources and the general purpose of formative assessment in a technology intensive environment require further comment.

The use of Open Education Resources (OERs) that are free online learning materials accessible to anyone across the world for teaching and learning purposes can be used to enhance formative assessments. The multitude of open education resources available to support learning can only be fully realized when coupled with meaningful formative assessments, especially in areas involving critical thinking skills. As the use of OERs spreads, emphasis on digital literacy becomes increasingly important. Not everything that is openly accessible is reliable or easily found. Moreover, there are associated reasoning skills pertaining to the selection, validation and use of open resources that might be considered basic skills that also require the support of new assessment strategies and techniques. Supporting learning in a connected world in which students have easy access to rich sets of data and information is a new and challenging area for formative assessment. New tools and technologies for formative assessments can be used to leverage OERs and eventually also create open assessment repositories that could been be mined using learning analytics to design learning scenarios as well as personalize learning experiences for individual learners.

To conclude this section on assessments in technology-enhanced learning situations, we wish to emphasize several points. First, nearly all learning environments in the 21st century involve and depend on digital technologies, such as computers, hand-held devices, the Internet, internet whiteboards and so on. As a result, the purpose of formative assessment has not changed on account of new technologies. Rather, the significance of formative assessment has grown on account of new technologies and 21st century learning demands. As more and more learning activities involve the internet, the timeliness of feedback becomes increasingly important. A week’s or even a day’s delay in proving feedback can jeopardize the effectiveness of the feedback in terms of improving learning and performance. Providing such timely and meaningful feedback without making use of advanced technologies is difficult to imagine. Those possibilities are discussed in the next two sections.

As a closing reminder to this section, it is worthwhile to recall the key purposes of formative assessments identified by Ecclestone (2010): (a) determine starting points and current knowledge levels of learners, (b) identify strengths and weaknesses of learners in relation to desired learning outcomes, ((c) review the learning progress of individual learners, (d) set targets and goals for future efforts, and e) guide and advise individual learners to address their weaknesses and build on their strengths and interests. While a formative assessment supports those five areas, the primary purpose of formative assessment is to help learners attain intended goals, which involves indicating weaknesses and deficiencies while encouraging specific activities and approaches to promote engagement and address weaknesses. The outcomes of formative assessments should enable teachers to make instructional decisions and enable students to set goals, allow students to evaluate their performance and change their approaches to learning, while actively engaging the students in their learning; improving their skills and knowledge; motivating and allowing students to become independent learners (Crookall, 2010; Ecclestone, 2010; Slavich, & Zimbardo, 2012). As indicated earlier, as formative assessments are emphasized and developed for challenging situations (large and multi-grade classrooms, inquiry- and problem-based learning), there is an opportunity for technology to play a key role in formative assessment. Recent studies in technology-enhanced problem based learning (PBL) have indicated the continuing centrality of the teacher or tutor in facilitating feedback whether in blended environments (Bridges et al., 2015), online meeting technologies (Ng et al., 2013) or virtual learning environments (Savin-Baden et al., 2015).
New trends and directions in formative assessments in technology enhanced learning

In this section, a few representative formative assessment efforts in a number of technology-enhanced learning contexts are discussed in this section which is far from exhaustive. General conclusions are addressed in the next section followed by a number of recommendations.

Formative feedback mechanisms

Technology can be used to support formative feedback in many different ways. In general, one can conceptualize formative feedback as one form of scaffolding for learners. The simplest scaffolds provide feedback based on the performance of a learner on a specific learning task, as was done in various intelligent tutoring systems. That form of feedback requires a database containing the subject to be learned along with common problems encountered with links to remediation. Automatic direction to a question and answer database based on a learner response is another simple formative feedback mechanism. Such mechanisms already exist and can be expanded. While those mechanisms still have appropriate uses, their application is generally limited to very well defined domains and straightforward learning tasks.

More advanced mechanisms depend on additional information about learners that goes beyond their performance on a specific learning tasks. An effort at the Hong Kong Institute of Education aligned to formative assessments with individual differences and formative feedback orientation (Yang, Sin, Li, Guo, & Lui, 2014). To gain a deeper understanding of the impact of formative feedback orientation on students’ learning, they took account of students’ important goal orientations: (a) performance goal orientation, and (b) learning goal orientation. Yang et al.’s (2014) findings indicated that students with a learning goal orientation (e.g., master that learning tasks) were more likely to find teacher feedback useful and feel personally responsible to respond to teacher feedback (i.e., high accountability of implementing feedback) than those with a performance goal orientation (e.g., get a desired grade).

When a robust and dynamic learner profile is available that includes interests, biases, preferences, as well as performance data, then feedback can be generated either by a teacher or by a smart learning engine that leverages those characteristics in suggesting next steps and resources to explore. Such an intelligent feedback mechanism has yet to be deployed on a large and sustainable scale, so it is one of the recommendations mentioned in the final section.

Formative feedback for problem-based and inquiry learning

Ongoing research in technology-enhanced problem-based learning (PBL) in health sciences education are now aimed at integrating multiple approaches to learning activities and formative feedback (Kazi, Haddawy, & Suebnukarn, 2009). These designs employ multimodal problems, scenarios, or/and cases (videos, 3-D imaging) in re-designed physical and virtual learning spaces. In this new iteration of the traditional PBL format, students collaborate in face-to-face groups via mobile devices, interactive whiteboards and online resources (including open educational resources) using collaborative tools and concept mapping software. Findings are indicating new opportunities and challenges for PBL designers and facilitators in scaffolding deeper approaches in blended PBL environments (Bridges, 2015; Chan et al., 2012; Chan et al., 2015) including harnessing the potential of new digital texts in supporting visual semiosis across the PBL cycle (Bridges et al., 2015). Of course PBL has extended far beyond medical education and training and these approaches are generally applicable in any PBL context.

Formative feedback for ePortfolios

ePortfolio tools are increasingly used to support PBL, inquiry learning and other subjects and are now embedded within learning management systems (Fisher, & Hill, 2014); they provide a holistic approach to recording achievements (Bell, & White, 2013; Ambrose, & Chen, 2015) and providing formative assessment in line with professional standards. While many institutions are still working through the known challenges in relation to ePortfolios (Gibson, 2006), learning management systems and other learning technologies have provided elegant solutions (Smart, Sim, & Finger, 2015; Brown, 2015). While ePortfolios provide opportunities for authentic
Formative and summative assessments, they require significant human time with regard to feedback and assessment (Smart, Sim, & Finger, 2015). While ePortfolios are promoted for their potential benefits to teaching, learning, assessment, and curricula, they are seen as especially useful for extending and deepening assessment value (Ambrose & Chen, 2015; Fisher & Hill 2014). However, empirical research into ePortfolio initiatives suggests the complexities and challenges are significant. One obvious challenges involves the variety of information types and representations that comprise portfolios (text, graphs, links, audio, video, images, etc.). There are also issues involving privacy and intellectual property rights. In addition, it is not clear that large collections of ePortfolios will be available for use within the context of an intelligent assessment system, so this is another area recommended for further exploration.

Formative feedback to improve motivation and engagement

When it comes to developing knowledge and skills in complex and challenging domains (discussed next), issues pertaining to motivation and engagement emerge as important feedback areas. The reality is that motivation and engagement are important areas to support with formative feedback in nearly every learning situation because there is a clear link between motivation, engagement, time-on-task and learning outcomes (Keller, 2009). While several learning technologies and tools support engaged learning and formative assessments, promoting student engagement and motivation has become challenging with regard to complex learning tasks and in developing countries for large student cohorts with limited resources. A group of academics at the Sri Lanka Institute of Advanced Technological Education (SLIATE) have been experimenting with new teaching and assessing methodologies; combining team-based learning and guided inquiry learning (Lokuge Dona & Gregory, 2015a; Lokuge Dona, Gregory & Pechenkina, 2016). The assessments are designed in accordance with established pedagogical principles (e.g., provide scaffolding according to individual learner needs) so as to initiate discussion among the academics about the key concepts related to the unit, to identify a strategy to embed the idea and to identify the suitable technology for a successful implementation (Gregory & Lokuge Dona, 2015; Gregory, Lokuge Dona & Bailey, 2015). This approach has enabled SLIATE staff to provide an engaging learning experience and develop a collaborative work environment, representing real world situations while integrating digital assessments (Lokuge Dona & Gregory, 2015a). The ability to provide immediate meaningful feedback has facilitated significant improvements in learning. However, substantial training and mentoring of teachers with a change management process is required for successful implementation on a larger scale (Lokuge Dona, 2015; Lokuge Dona et al., 2015b), which is discussed more generally in the section on teacher training below.

Tools for complex learning

As mentioned in the discussion of 21st century skills, critical thinking is key to success in a number of areas in which inquiry- and problem-based approaches are appropriate. Within those approaches, complex learning tasks are frequently involved - that is, tasks that are somewhat ill-defined and that lend themselves to multiple solution approaches and acceptable solutions. It is precisely in such areas that critical thinking is important to develop, and those tasks tend to recur throughout a person’s life in both job and personal contexts. The most promising recent advances in providing meaningful just-in-time, just-when-needed formative assessment for complex learning tasks involve a series of research efforts in Germany and the USA and tools that were consolidated in Highly Integrated Model Assessment Tools and Technology (HIMATT; Ifenthaler, 2010; Pirnay-Dummer, Ifenthaler, & Spector, 2010; Spector & Koszalka, 2004). HIMATT provides a learner with a problem situation and then prompts the learner to indicate (in the form of text or an annotated graph) the key factors and their relationships involved in addressing the problem. This problem conceptualization can be compared to an expert conceptualization or reference model and analysed to indicate things for the learner to consider (Ifenthaler, 2010; Pirnay-Dummer, Ifenthaler, & Spector 2010; Spector & Koszalka, 2004). A further development of HIMATT is AKOVIA (Automated Knowledge Visualization and Assessment; Ifenthaler, 2014). These formative assessment tools require refinements and user-friendly interfaces to be used in face-to-face and online settings. The tools were originally developed for research purposes and have now been shown to be effective and useful in more than 15 research studies. However, to be converted into formative assessment tools for use by teachers, a number of enhancements need to be made (Ifenthaler, 2011; Shute, Jeong, Spector, Seel, & Johnson, 2009). First, a way for teachers to easily enter representative problems and to develop reference models based on their own responses and those of recognized domain experts. Second, a decision algorithm needs to be included that determines which parts of a reference model might be suggested to a learner for consideration or whether and when to show the entire reference model and prompt a learner for a comparative
analysis. This research needs further development and enhancement to create a new version based on the currently validated methodology for educators to be able to use it.

**Adaptive formative assessment**

Computer-assisted learning environments made use of branching based on learner interactions that were the same for all learners in that same situation (Wauters, Desmet, & Van den Noorgate, 2010). Intelligent tutoring systems were a step forward but once a learner had been profiled with a specific deficiency in the learning environment, the remediation was again the same for all learners with that particular deficiency (Spector, 2014a; Spector, 2016). Those systems were static with regard to the knowledge domain and feedback mechanisms involved. However, with the development of new learning technologies, large sets of data, learning analytics and advanced artificial intelligence tools more can be done to create dynamic formative feedback systems (Murray & Pérez, 2015; Spector, 2014a; Spector, 2016). As discussed by Tyton Partners (2013), an adaptive learning system is capable of creating a personalized learning experience that employs a sophisticated data-driven, often non-linear approach to instruction and remediation, that adjusts to a learner’s interactions and demonstrated performance level, and subsequently anticipates types of content and resources learners need at a specific point in time to progress (Ifenthaler, 2015; Pirnay-Dummer, Ifenthaler, & Spector, 2010). MathSpace, an adaptive learning system for mathematics that was piloted at Swinburne University of Technology showed great results in supporting students in their mathematical skills (Alao, Lee, O’Kane, & Jackson, G, 2016). This system enables hand writing recognition with all mathematics problems, allowing learners to undertake different mathematical questions with immediate feedback at each stage of the solution. As Hattie and Timperley (2007) argued, ongoing or frequent feedback, especially feedback that was visible to both learners and teachers, has a significant impact on learning. In the MathSpace system, the students are guided with feedback for each step in mathematical problems, facilitating the development of students’ understanding on how to improve. Intelligent systems are now capable of dynamically providing questions that will support the students’ learning based on the learner’s capabilities (Malekzadeh, Mustafa, & Lahsasna, 2015). Smart Sparrow is another authoring tool that several international and Australian higher education institutes use to create a dynamic learning environment to support different learning styles. This also facilitates real-time dynamic mapping and sequencing of instruction to match individual learner characteristics.

**Massive online formative assessment**

Providing an engaging learning environment with timely feedback to the learners plays a major role in the development of MOOCs course design (Gregory & Lokuge Dona, 2015; Lokuge Dona & Gregory, 2015a). MOOCs have successfully employed digital badges to support and encourage formative assessments as a method to engage and motivate learners to develop skills, acquire knowledge and provide feedback to others (Lokuge Dona, Gregory, & Pechenkina, 2016). There are particular challenges involved in providing large numbers of online learners with timely and meaningful feedback as they progress through a series of learning activities. However, it was evident that the use of constructivist approaches such as the five stage model explained by Salmon (2011) with e-moderating skills for e-tivities in an online environment (Salmon, 2013), enables ways to manage a large number of online learners effectively with useful and engaging feedback mechanisms (Salmon et al., 2015). Competency-badges are also proposed by Baker (2007) and are consistent with the notion of a Mini-MOOC proposed by Spector (2014b) that focuses on a tightly defined well-structured domain for which automated formative feedback can be generated.

There is much pressure in Asian countries to facilitate learning for large cohorts, especially taking distance courses. The lesson from the experience in Indonesia is that ICT needs to be integrated into formative assessment to reap better benefits. The distributed basic education project in Indonesia required thousands of in-service teachers without degrees to complete a baccalaureate or lose their jobs. In-service teachers had to do this while working. Being a full-time student while working full-time is a challenge. Those involved were constantly seeking ways to minimize such a heavy load. Getting immediate feedback on learning tasks was essential, but extremely challenging to provide. Without the support of the Internet in remote areas, the only alternative was to send tutors to the countryside to help. The lesson from the experience in Indonesia is that ICT needs to be integrated into formative assessment. The distributed basic education effort in Indonesia, funded in large part by USAID, was successful, although not all objectives were achieved. An important outcome of the effort was the impetus to make Internet access widely
available to students and teachers throughout Indonesia. Universitas Terbuka has been instrumental in realizing the goals and objective of the effort (see http://www.oeconsortium.org/members/view/568/).

Teacher support in technology-enhanced formative assessment

As is the case with nearly every change in learning and instruction, proper training and professional development in support of teachers is required. Lack of teacher support has been cited as one reason that very promising technologies have failed to scale up and achieve sustained success (Spector & Anderson, 2000). Yang et al. (2014) discussed how important it is to provide professional development to teachers to change the culture of feedback provision in classrooms and to build positive instrumental attitude towards formative assessment. In a similar vein, To and Carless (2015) also discussed the importance of providing exemplars and guiding teachers towards building strategies to implement formative feedback methods. What has not been explored is how these findings might change when formative feedback is being automatically generated by an intelligent assessment engine.

While many have discussed the need to reform teacher training and better support teacher professional development, especially in the area of technology-enhanced learning and formative assessment, there are only a few instances where this has happened on a large and sustainable scale. One example cited by OECD (Organisation for Economic Co-operation and Development) is Finland, which has steadily improved teacher training, curricula, and technologies to support learning, instruction and assessment (see https://www.oecd.org/pisa/pisaproducts/46581035.pdf). Other countries in which large-scale reforms have occurred in concert with teacher support and the integration of technologies to support learning, instruction and assessment include, among others, China, Ireland and South Korea. The fact that more progress has not occurred with regard to the adoption of advanced technologies in support of formative assessment and innovative learning environments is in part due to need to train and re-train teachers and gain societal consensus on the value and importance of committing the resources needed to support 21st century skills. We conclude the paper with some synthesizing remarks about conclusions and recommendations for future efforts.

Conclusions and recommendations

In order for the full potential of formative assessments to be realised in the context of supporting critical thinking, inquiry learning and 21st century skills, it is important for decision makers at all levels to be aware of the significance of formative assessments as well as formative evaluations (Spector & Yuen, 2016). The different institutions and technological capabilities also need to be evaluated with the professional development capabilities to implement a sound strategy for formative assessment for learning. The professional development opportunities for staff to clarify and understand the role of formative assessments and evaluations in learning and instruction should be developed, widely emphasized and disseminated. Providing relevant theoretical and empirical grounding and short but poignant examples will assist teachers and decision makers in navigating their way through the process of implementing formative assessments.

Without emphasis on formative assessment, support for new assessment tools and technologies, and an open assessments repository, it is not easy to change current formative practice and current emphasis on summative assessment and high-stakes testing. Two forms of change are required: (a) more emphasis on formative assessment (and less on summative assessment) as formative assessment is linked directly to improved learning, and (b) formative assessment practices to address learning situations that present difficult challenges (e.g., large and multi-grade classrooms, inquiry- and problem-based learning). New technologies and technology enhanced delivery modes such as MOOCs, serious games, and gamification will be unable to realize their full potential and impact on learning if new methods, tools and resources are not used effectively to encourage students to gain knowledge and learn the necessary skills. According to Shute and Ventura (2013) and Webb et al. (2013), new methods include stealth assessments (i.e., embedded and continuous unobtrusive measuring of performance while learners are engaged in computerized tasks). This type of data can be used to make meaningful interpretations about learner habits. It also can be provided to the learners, to give them a meaningful representation of evidence and arguments about their achievements. These strategies could enable a wider range of measures to contribute to judgements of students’ achievements, thus supporting their learning. In addition, with the introduction of new technologies and course
delivery methods, it is also important to consider professional development of staff in order for them to be comfortable in using these models and tools.

Learning in the 21st century includes digital literacy (skills associated with searching, evaluating, using, modifying and creating digital artifacts) and reasoning literacy (critical thinking skills) among the other basic skills to be developed in primary and secondary education that need to be enhanced through formative assessment for learning. While technological developments have opened new methods of conducting formative assessments, it is important to analyse the suitability of such learning technologies for the purpose (Marope, Chakroun, & Holmes, 2015). Turning research tools such as HIMATT into tools usable by teachers in a variety of classrooms should be a priority for funding agencies and educational developers. Supporting learning analytics is yet another priority recommendation, although this is a challenge when much of the necessary data is held in non-open repositories. Another recommendation important with regard to personalized learner is to create accessible and robust learner profile that include preferences, interests, biases as well as performance data. Such databases have yet to be created on a large and sustainable scale but are important to fully realize the potential of personalized formative feedback. In order to move assessment into the 21st century, educational institutions need to invest in professional development and use of new tools and technologies especially well-suited for complex problem-solving domains, personalized learning and massive online learning situations. In addition, new assessment tools and technologies could be used for meaningful diagnostic and cross-cultural purposes to form the basis for informing and improving educational systems to enhance students learning. This type of supported learning environment can assist life-long learning, including developing effective problem solving and critical thinking skills.

Our basic recommendation for researchers, developers, educators and policy makes is to address the over-emphasis on summative assessments and evaluations and under-emphasis on formative assessments and evaluations. One form this could take would be in the form of an open assessment repository (OAR) or an education observatory (e.g., similar to the discontinued ERIC clearinghouses), which could be funded by governmental agencies and foundations interested in genuinely supporting 21st century skills and knowledge. A second but equally important recommendation is to fund and support the development and deployment of powerful formative assessment tools that align with 21st century skills especially in the areas of critical thinking and inquiry- and problem-based learning. A final recommendation is to properly train and support teachers in the important but somewhat neglected area of formative assessment.

References


A Multilevel System of Quality Technology-Enhanced Learning and Teaching Indicators

Nancy Law¹*, Dale S. Niederhauser², Rhonda Christensen³ and Linda Shear⁴

¹Faculty of Education, University of Hong Kong, Hong Kong // ²Department of Curriculum and Instruction, College of Education and Human Services, West Virginia University, USA // ³Institute for the Integration of Technology into Teaching and Learning, College of Information, University of North Texas, USA // ⁴Center for Technology in Learning, SRI International, USA // nlaw@hku.hk // dsniederhauser@mail.wvu.edu // Rhonda.christensen@gmail.com // linda.shear@sri.com

*Corresponding author

ABSTRACT

In this paper we elaborate and extend the work of the EDUsummIT 2015 Thematic Working Group 7 (TWG7) by proposing a set of indicators on quality Technology-Enhanced Learning and Teaching (TEL&T). These indicators are intended as one component of a set of global indicators that could be used to monitor implementation of the Education 2030 agenda, which aims to “Ensure inclusive and equitable quality education and promote life-long learning opportunities for all.” The proposed indicators address conditions at the student, teacher, school and system levels, and are organized in a systemic framework to help foreground interactions and interdependencies within and across the different levels. This framework highlights the need for longitudinal, multilevel designs in evaluation studies of TEL&T implementations, which will also contribute to a better understanding of the links between policy, policy implementation, outputs, and outcomes. Sample indicators are presented to illustrate the framework, and suggestions are made for use of the framework in evaluation studies. Further, the proposed framework could be used to underpin the development of an open, worldwide collaborative of educational evaluation researchers, practitioners and policymakers, thereby adopting a crowdsourcing approach to systematically address the complex challenges in evaluating quality TEL&T.

Keywords

Technology-Enhanced Learning, Indicators, Quality learning, Multilevel, Systemic change

Introduction

Improving quality of education is one of the six goals of the Education for All (EFA) Framework for Action that was slated to be achieved by 2015 (UNESCO, 2000). Representatives from 164 governments met in April 2000 in Dakar, Senegal and committed to:

*Improving all aspects of the quality of education and ensuring excellence of all so that recognized and measurable learning outcomes are achieved by all, especially in literacy, numeracy and essential life skills.*

According to the EFA Global Monitoring Report (UNESCO, 2015a), there have been impressive improvements across the world in access to education during this time, but quality of educational experiences and attainment of measurable learning outcomes both within and between countries have been extremely variable. There have been numerous cross-national assessments of education quality during this fifteen-year period, including four TIMSS and three PIRLS studies (http://timssandpirls.bc.edu/), six PISA studies (https://www.oecd.org/pisa/aboutpisa/), and the ICILS 2013 study (http://www.iea.org/ictl/2013.html), as well as multiple regional studies (e.g., the studies conducted by the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ, http://www.sacmeq.org/?q=about-sacmeq) and the Latin American Laboratory for Assessment of the Quality of Education (LLECE) (http://www.unesco.org/new/en/santiago/education/education-assessment-llece/). One major critique of these studies is their narrow focus on cognitive learning outcomes. Conceptions of quality education should be broadened to include outcomes such as attitudes and dispositions (Knezek & Christensen, 2008; Williams & Engel, 2013), and social and emotional outcomes (Learning Metrics Task Force, 2013).

Quality of teachers, faculty training opportunities, quality of learning materials, the nature of the learning environment, and school leadership all contribute to student learning outcomes in important ways (Burns & Darling-Hammond, 2014; OECD, 2013). Studies of quality learning and teaching should include the conditions for learning at multiple levels of the education system, including conditions that support teachers and school leaders, as well as students (Jaquith et al., 2015; Law, Kampylis & Punie, 2015). A shortcoming of existing large scale studies of student achievement is that these are cross-sectional studies, lacking longitudinal data to shed light on the effects of
different conditions and interventions such as regional or national education reforms (Goldstein, 2004; Schmidt & Burroughs, 2013; Christensen, 2015) and discount the multiple levels that exist in education.

The education agenda for the next fifteen years was set at the World Education Forum held in Incheon, South Korea in May of 2015. The Forum culminated in the Incheon Declaration report, *Education 2030: Towards inclusive and equitable quality lifelong learning for all* (World Education Forum, 2015), which defined high quality education as education that:

... fosters creativity and knowledge, and ensures the acquisition of the foundational skills of literacy and numeracy as well as analytical, problem-solving and other high-level cognitive, interpersonal and social skills. It also develops the skills, values and attitudes that enable citizens to lead healthy and fulfilled lives, make informed decisions, and respond to local and global challenges through Education for Sustainable Development (ESD) and Global Citizenship Education (GCED). (p. 7-8)

Although the overarching goal of the declaration report was broad and holistic in scope, authors recognized the central role that technology should play in addressing these goals, stipulating that: “Information and communication technologies (ICTs) must be harnessed to strengthen education systems, knowledge dissemination, information access, quality and effective learning, and more effective service provision (p. 2);” and calling for comprehensive systems to monitor and evaluate the implementation of the EFA goals at national and international levels.

An appropriate set of indicators lies at the core of any system of evaluation (Bryk & Hermanson, 1993). For those countries that might choose to adopt such indicators, they could be used to measure progress and inform decision-making for continuous improvement at the national policy-making level, and provide global monitoring data for the UNESCO Institute for Statistics as called for in the Incheon declaration. Designing a systemic assessment approach, and identifying an appropriate set of indicators, was the focus of the EDUsummIT 2015 Thematic Working Group (TWG) for Indicators of Quality Technology-enhanced Learning and Teaching, with specific objectives to:

1. Develop a literature-based conceptual model of the impacts of ICT implementation strategies on multiple levels of the education system: including individual, classroom, school, district and system levels.
2. Identify a core set of quality indicators for technology-enhanced teaching and learning at the different levels based on the above conceptual framework.
3. Develop a preliminary mapping of the indicators to existing cross-national studies, and common national/institutional monitoring and evaluation mechanisms.

Objectives were addressed within the TWG7 working group, and shared across the other working groups, during intensive meetings over a two-day period at EDUsummIT 2015. This article elaborates on a brief summary of the outcomes of those discussions (Law, Niederhauser, Shear & Christensen, 2015) with regard to developing a multilevel conceptual framework for the design of quality Technology Enhanced Learning and Teaching (TEL&T) indicators (Objective 1), and providing a preliminary sketch of the kinds of quality indicators that would contribute to such a framework (Objective 2). Objective 3 is beyond the scope of the present paper.

**Indicators for quality TEL&T: Meaning and scope**

Integrating the use of digital technology into the learning and teaching process to improve the quality of learning outcomes has become an important strategy for improving educational quality, and is often referred to as Technology-enhanced Learning and Teaching (TEL&T). Technology-Enhanced Learning (TEL) is an increasing focus for educational policy makers, school leaders and teachers around the world as technological literacy becomes progressively more important in the global society. Many countries have launched national policies that promote integrating TEL into the school curriculum (Plomp et al., 2009), with the US launching its first national technology plan in 1996; Singapore developing four IT Masterplans since 1997; the UK introducing its National Grid for Learning in 1997; and Malaysia launching its Smart Schools program in 1996. The policy level rationale for emphasizing TEL is not only to improve learning outcomes, but to also transform the learning process and foster new capabilities necessary for leading healthy and fulfilled lives, make informed decisions, and respond to local and global challenges in the 21st century (World Education Forum, 2015). UNESCO (2008) proposed a policy framework that aligned national educational goals and curriculum, and the role of information and Communications Technology (ICT) in teaching and learning, with the state of economic development. While the extent to which the UNESCO
document influences the formulation of national TEL master plans and strategies in different countries is not clear, it is evident that policy makers generally recognize the increasing importance of ICT in social and economic development, and the need to prepare citizens (particularly the school age population) for life in a knowledge-based society in which basic literacy and numeracy skills are no longer sufficient. Education in a digital age must prepare learners with information literacy skills as they use technology to communicate, collaborate, and problem-solve (Partnership for 21st Century Skills, 2003)—all necessary skills for successful participation in developed economies.

What kinds of indicators would be needed to investigate the extent to which the vision of providing quality TEL&T has been achieved? At a basic level, we would need indicators for measuring students’ learning outcomes. Since 21st century learning outcomes (e.g., problem solving and collaboration) go beyond specific knowledge or skills that can be taught and assessed directly, indicators are needed that will inform us of the students’ conceptual understanding, actions and interactions as they engage in the learning process. In fact, literature on TEL&T over the past two decades shows a clear need for implementing more effective teaching and learning practices if we are to achieve more ambitious metacognitive, social emotional, attitudinal, and learning outcomes (Law, Pelgrum & Plomp, 2008; OECD, 2013). Hence indicators for quality TEL&T should include indicators for learning processes and practices. We also know that learning processes and interactions do not take place in a vacuum, but are embedded in conditions for learning, which include the physical, social, cultural and digital aspects of the learning environment. Another crucial set of indicators should relate to the digital tools used and the roles the tools play in learning interactions, or e-learning use.

The four italicized categories of indicators described above refer to indicators for student learning. However, studying these aspects of student learning in isolation would not provide us with a holistic picture to enable a comprehensive understanding of the conditions that contribute to student learning. In the literature, both in TEL&T and in educational reform and innovation, there is wide recognition that changing pedagogical practices requires a deep level of teacher learning (Looi & Teh, 2015), which involves the acquisition of new knowledge, skills and competencies as well as changes in values and beliefs about learning, assessment and appropriate ways to use technology with their students. Hence, we must include indicators related to teachers’ knowledge, skills and beliefs when determining indicators for quality TEL&T.

In turn, the teaching and learning practices that occur in schools are very much influenced by leadership and contextual factors at school and system levels. Thus, there is a need for a multilevel set of indicators that will help us understand and monitor the extent to which the goals of quality TEL&T have been achieved and better understand the factors contributing to student achievement.

**Challenges to developing a set of quality TEL&T indicators Fit for national and global comparisons**

Many countries have made heavy investments in infrastructure, teacher professional development and digital learning resources in the implementation of their TEL masterplans. Most of these countries have their own evaluation mechanisms, but unlike the case of scientific or medical research, where there is typically a set of commonly agreed upon indicator standards and conventions (as well as established requirements for data collection and analysis), educational researchers tend to develop their own instruments. One reason for this is that different indicators are needed to serve different research contexts and purposes. However, one often encounters very similar indicators and instruments in the TEL&T literature that share the same broad conceptual framework and methodological approach, but cannot be easily linked or compared because of differences in methodology and nature of the indicators. There are also differences in the expertise of those conducting the research, and varying levels of support for TEL&T research in different countries. Reflecting on the monitoring of the Millennium Development Goals (MDG; UNESCO, 2000), the UN System Task Team (2013) pointed to the need for well-coordinated and professionally staffed statistical services at national and international levels in order for global monitoring to be realized. Adequate monitoring is a resource intensive process, requiring both funds and expertise—which is particularly challenging for developing countries. Further, any system of monitoring at a system-wide level is necessarily limited in its scope and the scale of data collection.
A systemic multilevel framework to conceptualize quality indicators for TEL&T

To address the challenges discussed above, we elaborate in this paper a conceptual framework for the development of a set of indicators for evaluating TEL&T, and provide illustrations on the types of indicators that will be needed. Based on the definition of quality learning within the framework of Education 2030, the focus for TEL&T is not simply to help students learn more effectively what they have traditionally been learning. Rather, it is to develop lifelong learning capabilities by providing learning opportunities that would not be otherwise accessible to a given group of students (World Education Forum, 2015). A rich body of literature shows the need for innovations in pedagogy, curriculum and assessment to achieve such outcomes (e.g., Voogt & Knezek, 2008). Consequently, TEL&T initiatives need to be supported by strategies and mechanisms for teacher learning, leadership learning and organizational learning. Further, the biggest challenge to the implementation of ICT-enabled learning innovations is their scalability (Kampylis, Law & Punie, 2013). Studies on change, innovation and sustainability all point to the need for the change to be multilevel (Blamire & Gerhard, 2009; Law, Kampylis & Punie, 2015). Student learning takes place at the core of the education ecosystem, situated within classroom contexts, which in turn are dependent on hierarchically nested conditions such as teacher learning, leadership practices and infrastructure at the school level, and policies and practices at the system level (Davis, 2008). Changes that takes place at each level impact other levels as interactions occur within and across levels (Davis, Eickelmann & Zaka, 2013).

![Figure 1. A diagrammatic representation of the interrelationship among the different levels of indicators (adapted from Law, Niederhauser, Shear & Christensen, 2015)](image)

While there are multilevel models that analyze the influence of pedagogy and school level factors on student learning (e.g., Fraillon et al., 2014), there is no model of interaction among the factors within each level, nor is there an explicit model that differentiates the mechanisms through which factors at different levels influence students’ learning outcomes. This results in the common practice of lumping all indicators from the different levels collected at the same cross-sectional time into a multilevel model. The lack of significant relationships among system-level factors, school factors, teacher factors, and student learning outcomes in large scale international comparative studies discussed earlier is an indication of the shortcomings of the current approach to multilevel modeling. Law (2015) proposed a parsimonious multilevel learning model to underpin the conceptualization of the many indicators.
involved in TEL&T. This model (see Figure 1) considers the changes happening at student, teacher, school and system levels in the process of TEL&T implementation, building on the literature that highlights the interdependencies of the changes taking place across the different levels (Davis, Eickelmann & Zaka, 2013; Law, Kampylis & Punie, 2015).

The indicators listed in Figure 1 are not meant to be exhaustive, but serve to illustrate the interactions and interdependencies across the categories and levels of indicators. At the student level, examples of 21st century skills include critical thinking, communication, creativity, collaboration and digital literacy. The opportunity to learn these skills depends on whether students had opportunities to engage in the types of learning interactions, supported by appropriate e-Learning use such as those presented in the figure (this relationship is illustrated by the arrows #1 and #2 in the figure). These opportunities to learn are in turn influenced by the conditions for learning shown at the student level: the school ICT infrastructure, home access to ICT, the pedagogy adopted by the teacher, etc. (illustrated by the arrows #3 and #4). As the literature on TEL&T demonstrates, such conditions do not present themselves as a direct outcome of a policy stipulation, but involve complex emerging and interacting processes (Kozma, 2003; Kampylis et al., 2013). Taking the condition of pedagogy as an example, this requires that the teacher possesses Technological Pedagogical Content Knowledge (Mishra & Koehler, 2006), and learning and assessment design expertise (Laurillard, 2013). These professional capacities can be conceptualized as learning outcomes at the teacher level (arrow #5 in Figure 1), which are in turn dependent upon the opportunities to learn (learning interactions and e-Learning use) and the conditions for learning available to the teachers concerned.

The conditions for teacher learning in turn depend on school level factors such as the vision, professional development opportunities and staff appraisal system at the school level (arrow #6 in Figure 1). Similarly, the emergence of these school level factors depends on the system level factors such as national education policies, e-Learning masterplans and school inspection criteria (arrow #7 in Figure 1). These “factors” are not static, and changes require a process of interaction and decision-making among stakeholders. By categorizing these factors as learning outcomes at the school and system levels, we wish to highlight the fact that changes in these factors may be transient (as in superficial learning) and may not be sustainable unless there is deep engagement by the relevant stakeholders at the institutional/system level in the deliberation of these changes. In addition, stakeholder interactions can be designed and scaffolded to facilitate decisions that are aligned with the overall desired direction of pedagogical change.

There are two important benefits of adopting this framework in the development of TEL&T indicators. First, the factors to be considered at each level can be categorized into the same four groupings: learning outcomes, conditions for learning, learning interactions, and e-learning use. As indicated by the red arrows in Figure 1, students’ learning outcomes are influenced by the learning interactions and e-learning use of students, which in turn are influenced by the conditions of learning available, such as school ICT infrastructure, pedagogy and assessment practices. Similar relationships exist across the four groups of factors at each of the other three levels.

A second benefit of grounding the indicators framework in a multilevel learning model is that it provides a theoretical guide to hypothesizing the causal connections among the factors across different levels. Students’ learning outcomes will feedback on the conditions for learning at the other three levels as these will strengthen or challenge the assumptions about learning, pedagogy, assessment and the role of ICT in supporting learning that underpin conditions such as the curriculum, school vision, national e-Learning plans, school inspection criteria, and national digital learning resources. The learning outcomes at the three higher levels affect the conditions for learning, learning interactions and e-learning use at other levels. For example, the teachers’ learning outcomes (TPCK, learning and assessment design expertise) influences the pedagogy and assessment practice as experienced by their students as conditions for learning; national strategies to support joint school/e-learning innovation projects (as system-level learning outcomes) provide opportunities for learning interactions at the teacher level; and staff appraisal criteria (as school-level learning outcomes) is one of the conditions influencing teacher learning.

In the following sections, we will provide an initial elaboration of the indicators under the four groupings at each level, as discussed in TWG7 at the EDUsummIT and supported by appropriate literature.
In this section, we describe the kind of student level indicators suggested for inclusion during the EDUsummitIT 2015 Workshop. Student learning is the focal goal of the educational system, which requires support and aligned learning interactions from many levels including teacher, school and system/policy. Around the world, preparation of students for constructive participation in our global society will require use of technology to effectively communicate ideas, and to collaborate locally and internationally to solve common problems.

Technology readiness involves more than just technological literacy skills. “Students will spend their adult lives in a multi-tasking, multifaceted, technology-driven, diverse, vibrant world – and they must arrive equipped to do so” (Partnership for 21st Century Skills (P21), 2003, p.4). Students must not only learn to use information technology (IT) but also to be ready to use technology for learning. Technology plays a major role in the definition of 21st Century skills, critical thinking, problem-solving, communication, and collaboration. This set of skills is commonly referred to as digital literacy (Resta et al., 2011, p. 3). According to ISTE Standards for Students, “students use critical thinking skills to plan and conduct research, manage projects, solve problems, and make informed decisions using appropriate digital tools and resources” (ISTE, 2007, p. 1).

Based on the above, indicators of TEL&T learning outcomes include:
- Students’ ability to demonstrate digital literacy by selecting the appropriate technology tool for a required task
- The ability of students to take responsibility for their appropriate uses of technology including safety in interacting with others online, appropriate sharing of information on social media sites, limiting the amount of time spent on technology
- Students’ demonstration of the skills necessary to organize, analyze, evaluate and communicate their ideas with others using digital resource
- The ability of students to demonstrate critical thinking in using digital resources to tackle unfamiliar, authentic and open-ended problems
- The extent to which students use digital tools to work productively with peers and subject matter experts

While systemic changes may be taking place in schools, changes in pedagogy and the learning environment may take more than a year or two before their impact on student learning becomes measurable. Indicators could be statements on a continuum to measure the levels of progress towards good practice. Based on the literature in this area (Fraillon et al., 2014; Sankey et al., 2014; ISTE, 2014) and discussions at EDUsummitIT 2015, conditions for learning for students should include indicators to answer the following types of questions:
- To what extent do students have personal access to a variety of digital resources to complete school assignments?
- Are students provided with appropriate skills and access to communication tools for enhanced learning opportunities?
- To what extent are students given choices in the selection of topics to study and the application of personal learning strategies?

Once the basic conditions for learning are in place, how students use technology for enhanced learning is important. The indicators for learning interactions include:
- How does technology support the students’ ability to revise their work based on peer, teacher and expert feedback?
- In what way does technology support the learner to be in charge of collaborative exploration of authentic problems?
- Does technology allow students to better self-reflect and self-monitor their learning?

Student indicators discussed in this section imply corollary teacher, school and system level indicators that need to be in place to monitor and evaluate holistically the technology enhanced learning ecosystem.
Teacher level indicators for TEL&T

Teacher-level factors, like access to technology, frequency of use and amount of teacher professional development, have frequently been assessed when examining TEL&T in schools. However, TWG7 teacher-level indicator subgroup participants felt it was more important to examine HOW teachers use ICT with their students to support the kind of learning outcomes that will prepare students to lead healthy and fulfilled lives, make informed decisions, and respond to local and global challenges as advocated in the Incheon Declaration (World Education Forum, 2015) and to ensure they develop the 21st century skills that will allow them to do so. How teachers choose to integrate ICT use into their practice is intrinsically linked to their personal belief systems about teaching and learning, which exert powerful influences on teachers’ curricular decision-making and their pedagogical practice (Niederhauser & Stoddart, 2001).

It is relatively easy to track factors like the number of Internet-connected computers to which teachers have access, the number of hours teachers used technology with students, and time teachers spent in professional development activities. Examining intrapersonal factors like teachers’ technological knowledge and skills, confidence in their ability to use technology and the value they perceive in using technology to aid in the teaching/learning process, the depth of knowledge they have in the content they are teaching, and their beliefs about how students learn, is a much more challenging endeavor. We will need to move beyond the kinds of self-report surveys and questionnaires that have been the norm up to this point and include research methods that draw on ethnographic research traditions like observation and interviews that take place over extended periods of time.

Based on the literature in this area and TWG7 teacher-level indicator subgroup discussions at EDUsummIT 2015, indicators of teachers’ learning outcomes could include personal attributes and observable behavior, as exemplified in questions 1-3 and questions 4-5 below respectively:

1) To what extent do teachers have positive self-efficacy and outcome expectations that will motivate them to effectively integrate technology into their practice?
2) To what extent have teachers developed knowledge and expertise that allows them to engage in pedagogy that support the kinds of learning outcomes associated with helping students develop 21st century skills?
3) To what extent do teachers have deep knowledge of the content they are teaching that would allow them to teach using the pedagogical practices advocated by current reform efforts?
4) To what extent do teachers integrate technology into the curriculum in ways that promote meaningful learning?
5) To what extent do teachers provide activities that require the learner to use appropriate technologies while engaged in collaborative exploration of authentic problems?

Extensive research has shown that professional development opportunities that are hands-on, sustained, and focused on student learning, have a significantly more positive effect on teaching practices than more traditional workshops (e.g., Darling-Hammond & Richardson, 2009), as do professional communities that provide collaborative ongoing supports (Vescio, Ross, & Adams, 2008). Indicators for learning interactions of teachers should include:

- Do teachers have opportunities to undertake collaborative lesson planning and assessment design for TEL&T?
- Do teachers have opportunities to peer-observe TEL&T lessons and to receive support for reflective practice such as action research?
- Do teachers have opportunities to engage with peers and school leaders in joint identification and exploration of drivers & obstacles to TEL&T?

Teachers need to be provided with the necessary conditions for learning in order that the above learning interactions can take place, and indicators for these include:

- Do teachers have access to technology (including learning analytics) and technological pedagogical support to design, implement, assess and provide feedback to students’ collaborative inquiry?
- Do teachers have, in their schools, organizational routines and structures to support collaborative co-design, peer learning, and an open, trusting and collaborative culture for sharing and risk taking?
- Do teachers have, in their schools, a formal role and organizational mechanisms for them to participate in decision-making on curriculum, assessment and TEL&T developments?

The successful implementation of TEL&T in the classroom implies support from school and system levels.
School level indicators for TEL&T

As described above, successful implementation of TEL&T in the classroom entails a complex transition for many teachers, requiring not only the use of new tools but the adoption of new models of teaching. School-level indicators, then, focus on the types of support that school-level leadership, cultures, and resources can provide to support teachers through the transition and to sustain that support over time. This section focuses on four commonly-recognized elements (e.g., Kennisnet, 2014): opportunities for professional learning, access to technology, curriculum resources, and leadership and vision.

The most common indicators tracked for professional development often relate to access and dosage, such as the number of hours teachers spend in professional development courses related to TEL&T. On the other hand, teachers’ practices will not be changed by increasing access and dosage per se, but rather by professional learning opportunities that support professional collaboration and reflective practice. A similar principle applies for the relationship between the level of digital infrastructure/technical support available and the quality of TEL&T implemented in a school. Common basic measures (e.g., UNESCO, 2015b) include pupil-to-computer ratios and access to electricity, connectivity, and other elements of infrastructure that allow their use. Further indicators such as technical support to teachers and learners, the bandwidth and reliability of Internet access during the school day, and digital learning resources are also important in mature implementations to support strong integration of ICT into teaching and learning without disruption. Research findings indicate that the pedagogical potentials of digital infrastructure and resources are often not realized (Cuban, 2009). Availability alone will not have impact on teachers’ practices unless teachers perceive these to be in response to their professional aspirations and needs. This requires the conditions of learning for teachers to be school level decisions made as a consequence of teachers’ professional explorations in TEL&T (i.e., teacher learning) and the participatory decision-making process involving different stakeholders (i.e., school level learning). Conceptualizing the indicators of conditions for teacher learning as indicators of school level learning outcomes helps to highlight the need to take account of the dynamic, interactive history of these indicators in predicting how these may contribute to a school’s capacities to implement TEL&T for quality learning in students.

The conditions for learning and learning interactions at the school level depend greatly on the availability of school routines such as co-planning and peer observations, timetabling, and staffing arrangements to support a productive architecture for learning (Stein & Coburn, 2008). Hence indicators of conditions for learning and learning interactions at the school level are closely related conceptually to the literature on instructional leadership (Robinson, Lloyd, & Rowe, 2008; Leithwood et al., 2004), including a clear vision for teaching and learning. Indicators related to instructional leadership and vision may include the frequency of various leadership activities, both central and distributed, and more qualitative indicators of the leadership supports perceived by teachers (e.g., Bryk et al., 2010). Additional measures examine not just the existence of vision statements, but the focus and the degree of guidance given to support 21st century educational outcomes (Twining, 2014).

An additional set of indicators for conditions of learning at the school level relate to system level factors. Vision and goals for TEL&T, priorities for staff and infrastructure development, as well as strategies and leadership practices at the school level are heavily influenced by educational policies and strategies at the system level, including:

- Whether TEL&T is a national educational priority, and if so, the nature of its vision;
- Whether the national curriculum, assessment, school inspection/accountability and teacher accreditation systems are aligned with an emphasis on nurturing 21st century skills such as learner self-direction and collaborative inquiry;
- Whether there are policy level structures and mechanisms to support TEL&T innovations, such as joint-school innovation projects, quality circles, sharing of good practices, etc.

System level indicators for TEL&T

The system level factors that constitute conditions of learning at the school level described in the previous section are in themselves system level learning outcomes. This is because national educational priorities, TEL&T goals and strategies and associated implementation mechanisms are in themselves products of negotiations and decision-
making processes at the system level; and such processes constitute learning interactions at the system level. Indicators of learning interactions at the system level include:

- The extent to which different stakeholders are involved in the policy discussions;
- The extent to which decision-making is informed by authentic explorations of TEL&T implementation in different school and community contexts;
- Whether there are mechanisms to promote sharing of experiences across schools, districts and regions in the TEL&T implementation process;
- Whether there are monitoring, evaluation and feedback mechanisms to fine-tune the system level factors (i.e., improve the system level learning outcomes) as the implementation progresses.

The conditions for learning at the system level are macro conditions outside of the arena specific to the education system. Examples of these indicators include:

- GDP of the country;
- National digital infrastructure such as broadband penetration and home ownership of computational devices;
- Extent to which the political system allows/encourages democratic discussion and community participation in policy decisions.

**Discussion and conclusion**

In this paper we elaborate and extend the work of the EDUsummIT 2015 Thematic Working Group 7 (TWG7) by proposing a set of indicators on quality Technology-Enhanced Learning and Teaching (TEL&T) as one component of a set of global indicators that could be used to monitor implementation of the Education 2030 agenda. The core contribution of our work is to propose a multilevel system of quality indicators that can be used for the monitoring and evaluation of TEL&T implementations. Quality learning that fosters 21st century outcomes requires student-centered, collaborative, inquiry-oriented learning interactions that are vastly different from the mainstream pedagogical practices currently found in classrooms. Achieving quality TEL&T student outcomes will require interdependent changes at all levels of the educational system. The proposed model conceptualizes factors influencing TEL&T implementation that result from school and system level decisions as learning outcomes at these respective levels. This model of indicators highlights that evaluations should not only focus on the magnitude of these learning outcomes factors, as their impact on TEL&T implementation depends very much on how far the stakeholders within and across the different levels had opportunities to engage in learning interactions (i.e., interaction that promote the sharing and alignment of vision, understanding and concerns) during the decision-making process. We have also argued for the need to go develop more informative indicators by drawing on ethnographic research traditions that take place over extended periods of time.

An important implication of this paper is the need for a systemic approach to the design of indicators in evaluation studies of TEL&T. This indicators framework highlights the need for all stakeholders (policy makers, funders, educators and researchers) to take account of the different timespans typically involved for the learning interactions to lead to observed learning outcomes. It provides a framework to identify which sets of indicators need to be included for the purpose of a specific study, and to avoid overlooking the key interdependent indicators. It also calls for multilevel, longitudinal evaluation designs in evaluation (Goldstein, 2004).

For policy makers and funders, this framework can be used to guide the selection of indicators that matter most in the context of the policy goals and initiatives. In particular, it is not adequate to simply measure policy input and student learning outcomes, and should also include indicators of learning processes and learning outcomes at all levels. By making the interdependencies of the different indicators explicit, mechanisms can also be established for these indicators to influence policy and decision-making.

For educators, institutional leaders and practitioners, the conceptual framework can be used to guide the selection and use of indicators to provide feedback and ensure alignment across contexts, processes and outcomes within a level. The appropriate choice of indicators can also provide feedback on whether the changes taking place at the focal level (e.g., teacher or school level) is aligned with what is needed and what is happening at other levels.
For researchers, this multilevel system of indicators can also serve as a framework to facilitate an open “crowdsourcing” approach to collaboration within the educational evaluation research community, similar to the open-source software development communities. Currently this is just a bold idea, which clearly requires support from the global research community as well as resources and expertise support from international research agencies to establish the necessary statistical infrastructure, quality criteria and guidelines, and protocols for the sharing of indicators and data. Such a development will potentially stimulate collaboration among researchers within and across national boundaries at an unprecedented scale, and help to address the global evaluation challenge for Education 2030. It will also allow us to build more holistic theories of connected learning: learning of students, teachers, schools, leaders and policy-makers as a connected system.

Acknowledgements

We are grateful for the contributions of the working group members of the EDUsummIT 2015 Thematic Working Group 7, who include, in addition to the authors, the following: Linda Shear, Esther Care, David Smith, Jonghwi Park, Bent Andresen, Hans van Bergen, Deirdre Butler, Allan Christie, Jill Downie, Tania Broadley and Louise Starkey. Nancy Law acknowledges that her contribution to this work is supported by the Hong Kong Research Grants Committee Humanities and Social Sciences Prestigious Fellowship Scheme # 37600414.

References


Smart Partnerships to Increase Equity in Education

Margaret Leahy¹, Niki Davis²*, Cathy Lewin³, Amina Charania⁴, Hasniza Nordin⁵, Davor Orlič⁶, Deirdre Butler¹ and Olatz Lopez-Fernandez⁷

¹Dublin City University, Ireland // ²University of Canterbury, New Zealand // ³Metropolitan University, UK // ⁴Tata Trusts and Tata Institute of Social Sciences, India // ⁵Universiti Utara Malaysia, Malaysia // ⁶Jožef Stefan Institute, Slovenia // ⁷Catholic University of Louvain, Belgium // Margaret.Leahy@dcu.ie // niki.davis@canterbury.ac.nz // c.lewin@mmu.ac.uk // acharania@tatatrusts.org // nizadin@uum.edu.my // davor.orlic@ijs.si // Deirdre.Butler@dcu.ie // olatz.lopez@uclouvain.be

*Corresponding author

ABSTRACT

This exploratory analysis of smart partnerships identifies the risk of increasing the digital divide with the deployment of data analytics. Smart partnerships in education appear to include a process of evolution into a synergy of strategic and holistic approaches that enhance the quality of education with digital technologies, harnessing ICT “smartly” both in relation to learning and support of the partnership itself. To guide strategic development as data analytics start to emerge in the schooling sector, two cases of large multi-stakeholder partnership initiatives aiming to increase access to education with ICT nationwide in India and Malaysia are analyzed. Mapping the partners’ collaborative activities in Davis’ Arena of change with digital technologies enabled the identification of both local and global influences in that ecological framework, which inform the choice of partners and their roles to increase equitable access to education. Research and development is recommended so that multi-stakeholder partnerships leverage data analytics alongside technology enhanced learning in the schooling sector with strategies that proactively increase equity.

Keywords

Smart partnerships, Multi-stakeholder partnerships, Digital equity, Technology enhanced learning

Introduction

Developments in learning technologies are enabling researchers to envision “smart” learning environments that supplement the role of teachers and learners using data analytics and other tools to enhance learning and administrative processes as they happen (Kinshuk, Chen, Cheng & Chew, 2016). Smart multi-stakeholder partnerships will be necessary to leverage these technological tools effectively as many schools and related educational agencies no not have the required expertise or infrastructure. However, the notion of a smart partnership is emergent and there is very little research to draw upon. We therefore present an exploratory analysis of the evidence to date, both in relation to literature and our own case studies, in order to identify areas for future research and development. The main goal of this paper is to define, illustrate and discuss Smart Partnerships in education, a theme that was assigned to thematic working group 1 for the 2015 EDUsummitIT (see http://www.curtin.edu.au/edusummit/).

The foundations of smart partnerships

While the concept of “smart partnerships” in education is still emergent, there is growing interest in “smartness,” generally associated with the use of technology, and particularly the collection and analyses of data, to improve aspects of day-to-day life. We now consider a selection of the literatures on “smart” and partnerships pertinent to the development of schooling and digital technologies in education in order to conceptualize and define smart partnerships in education.

The increasing focus on “smartness”

The term “smart” suggests innovative and transformative changes driven by new technologies. It brings up notions of data-driven decisions and technology-enabled data-sharing, plus communications and collaboration, all leading to continuous improvement (Kitchin, 2016). Currently, smart cities are of great interest globally with the common aim to use ICTs to transform systems and services, so that the quality of life and work is improved (e.g., Albino, Berardi
As such, the concept of smart cities is not limited to the diffusion of technologies, but also focuses on people and community needs. Smart cities are presented as large organic systems that connect various sub-systems including for example, transportation, energy, education, health care and buildings (e.g., Dirks & Keeling, 2009; Kanter & Litow, 2009).

Through reviews of the literature, various scholars have attempted to identify the core components or dimensions of a smart city. We focus here on a useful categorization of a smart city as three interconnected categories (Nam & Pardo, 2011; Meijer & Rodríguez Bolívar, 2015). Technology is a key attribute of a smart city and central to all definitions through an expectation that data-driven, networked technologies can be used to reconfigure aspects of daily life. This is argued to make smart cities more interconnected and efficient through the use of technology to engage in complex analytics, modelling, optimization and visualization (Harrison et al., 2010; Washburn et al., 2009; Meijer & Rodríguez Bolívar, 2015). The second critical factor in smart cities is its people and how they interact – i.e., “smart people” can be an important component in smart cities. This factor comprises a range of aspects including “lifelong learning, social and ethnic plurality, flexibility, creativity, cosmopolitanism, open-mindedness, and participation in public life” (Nam & Pardo, 2011, p. 287). Creativity is recognized as a key driver of the smart city and thus, education, learning and knowledge have central roles to play (Nam & Pardo, 2011; Thuzar, 2011). Finally a “smart community” is also recognized as critical, emphasizing the need for members and institutions to work in partnership to transform their environment (Berardi, 2013). Thus, collaboration is central and the involvement of all stakeholders, including the citizens of the city, is essential (Meijer & Rodríguez Bolívar, 2015). It also highlights that the strong support of key leaders is fundamental for success.

In some countries schools that make substantial use of technology to improve teaching, learning, management and administration have been described as “smart schools.” In Malaysia for example, the concept of smart schools has been utilized for around 20 years and is central to current education policy. The smart school term denotes “a learning institution that has been systematically reinvented in terms of teaching/learning practices and school management in order to prepare children for the Information Age” (Choi, Lee & Lee, 2015, p. 344). Thus, as well as technology improving education, the young people will be equipped for the workplace and for contributing to society. In common with people being a critical factor in smart cities, Choi and colleagues (2015) argue that the most important elements of a smart school are the teachers and students.

Smart learning environments and smart classrooms use technology to tailor learning to the needs of individual students, providing immediate feedback and support (e.g., Kinshuk et al., 2016). More recently, with the increasing focus on big data, data analytics has been advocated as a means of empowering both teachers and learners to improve learning experiences and outcomes (Clow, 2013). Learning analytics systems record numerous details about learners’ online activities providing information about engagement and progress, as well as predictive modelling such as attainment and dropout/retention (Clow, 2013; Papamitsiou & Economides, 2014). Thus, learning analytics can facilitate meaningful feedback, timely interventions and data-driven decision making in teaching and learning. The use of learning analytics in online learning environments is set to become increasingly used in higher education, where it is currently more developed, and also in school contexts (Johnson, Adams Becker, Estrada, & Freeman, 2015; Johnson et al., 2016).

There are many issues, however, in the use of learning analytics in educational contexts (Clow, 2013; Papamitsiou & Economides, 2014). Educators may not find it easy to ask the right questions of the data or interpret the results. The complexity of such systems could act as a deterrent to uptake as could the need to change practices. There are also concerns about the ethical issues surrounding data collection from individuals at such scale and potential for surveillance (Slade & Prinsloo, 2013). Finally, its usefulness relies on the embedding of online learning into teaching and learning practices, which itself requires sufficient architecture.

Smartness can lead to improved services through timely feedback to inform decision-making and interventions. Therefore, through the engagement of technology, people and communities, smartness is data-driven with analytics available in real-time and requires collaboration of stakeholder groups or at least their representatives. Technologies are essential but the knowledge and creativity of individuals working with partners at multiple levels within and across institutions are equally important. This leads to consideration of the literature on partnerships in education.
Partnerships in educational contexts

Partnerships in education can be summarized as “a mutually satisfying relationship, which typically involves the free sharing and exchange of knowledge and ideas to the benefit of [all] parties” (Falloon, 2015, p. 216). Partnerships should be viewed as a process rather than an event, concern relationships between both institutions and people, and are dynamic (constantly evolving) (Grobe, 1990). Grobe (1990) suggests that levels of involvement can vary from support (e.g., one-directional provision of resources) to co-operation (greater levels of participation with shared decision-making and short-term goals) and ultimately collaboration (greater levels of commitment and involvement, complex, with longer-term goals).

Although the literature has been drawn together over decades, it mainly relates to partnerships such as those between schools and parents, or those partnerships across and between educational organizations. While this body of literature provides useful information about the impact and range of activities the partnership supports, evidence about their functioning and success factors is limited. Instead we turn to literature on public-private partnerships and multi-stakeholder partnerships in education with a view to identifying the enabling factors/success indicators of partnerships, including those involving ICT in education initiatives. Partnerships that bring together a range of stakeholders are frequently advocated as a way of improving or expanding education (Draxler, 2008; Sarvi, Balaji & Pillay, 2015) as well as enhancing the capability and capacity of its members (Davies & Hentschke, 2006; Draxler, 2008). Such partnerships pool knowledge and expertise can provide additional funding and/or other resources, share responsibility and risk, and generally focus on solving large-scale, complex problems in response to needs (Davies & Hentschke, 2006; Draxler, 2008; LaRocque & Lee, 2011).

Public-private partnerships (PPP) are generally understood as joint government and industry initiatives whereas multi-stakeholder partnerships generally refer to partnerships that bring together a wider range of public, private and civil society stakeholders. In education, they are seen as a way of obtaining greater involvement of the private sector, improving and strengthening education systems by making education relevant for the economy, developing programs targeted to specific groups, and improving infrastructure (Draxler, 2008; Education International, 2009). For example, as cited in Pillay and Hearn (2009), the experience of the Organisation for Economic Co-operation and Development shows that PPPs can play a vital role in mobilizing the scale of resources required for developing ICT infrastructure, applications and locally relevant content, as well as developing the human capacity required for harnessing the full capacity of ICT (Ichiro & McNamara, 2003).

Such partnerships are not without their challenges. PPPs for example can weaken government control and activity, through a shift away from education as a public good towards the privatization of education with a subsequent loss of public accountability (Draxler, 2008; LaRocque & Lee, 2011). In addition, partnerships can face difficulties due to the complexity of underlying contracts between partners and insufficient capacity in the non-state sector (LaRocque & Lee, 2011). Moreover, despite intentions to improve equitable access to education, partnerships targeted with improving educational services can decrease equity (Patrinos, Barrera-Osorio & Guáqueta, 2009). For various reasons, there may also be resistance from some stakeholders (Patrinos et al., 2009).

Although the research evidence of the impact of PPPs and multi-stakeholder partnerships in education is limited, some success indicators were identified:

- The involvement of a broad set of stakeholders including the local community (Cassidy, 2007; Geldof, Grimshaw, Klein, & Unwin, 2011)
- Clear and shared aims and objectives (Cassidy, 2007; Draxler, 2008; Geldof et al., 2011; Falloon, 2015; Sarvi et al., 2015)
- Appropriate risk allocation and risk sharing (Osei-Kyei & Chan, 2015)
- Strong commitment from high-level leadership (Ministries of Education, companies) to ensure buy-in from all participants (Grobe, 1990; Osei-Kyei & Chan, 2015; Sarvi et al., 2015)
- Effective and open communication mechanisms (Grobe, 1990; Osei-Kyei & Chan, 2015)
- Shared decision-making and ownership (Grobe, 1990; Draxler, 2008)
- A clear finance plan (Sarvi et al., 2015)
• Effective management through change management strategies, strong regulation, joint accountability, monitoring and evaluation (Cassidy, 2007; Grobe, 1990; Draxler, 2008; Falloon, 2015; Osei-Kyei & Chan, 2015)
• Commitment to capability and capacity building (Cassidy, 2007)
• Mechanisms for sustainability (Draxler, 2008; Geldof et al., 2011)

Thus, success of partnerships depends on the effect of their actions on education, improved teaching and learning, better infrastructure and management of education systems, greater participation and engagement by the community and broader commitment by all sectors of society to enhancing education quality. It also depends on partnership size which affects sustainability and resilience. For example, Eickelmann (2011) found that schools sustaining innovative practice with ICT continued to engage with more partners whereas schools that had become less innovative had dropped partners. However, the relative importance of critical success factors can vary from country to country (Chou & Pramudawardhani, 2015) and, by implication, some factors may be unique to a specific context.

Drawing together the selected literature reviewed above, we argue that Smart Partnerships in education involve the use of technology to facilitate and support the partnership but also have a focus on improving teaching and learning through smart schooling. Smartness leads to improvement, in our case in learning, through data-driven decision-making support systems that increase efficiency in the day-to-day management of teaching and learning.

Towards a definition of Smart Partnerships

In addition to an ongoing review of relevant literature, the conceptual development of Smart Partnerships the work for this paper entailed interconnected stages, forming a spiral of activities leading towards the development of a definition of smart partnerships.

In the year preceding the 2015 EDUSummit, potential smart partnerships were identified and explored by members of the working group. Some were initially described in a report produced to support UNESCO Institute for Statistics (Twining, et al., 2015). The illustrations were purposefully developed to be wide ranging, including provision for learners, teachers, administrators, service providers, agencies in both school based and out of school provision. Initial analyses suggested that Smart Partnerships include the following key characteristics:

• Synergetic complementary contributions from partners, with clearly defined roles and responsibilities
• A vision that embraces a shared understanding across partners of how educational change can be brought forward by the partnership and is future focused
• Alignment between the defined goals and the paths undertaken to reach them
• Metrics that are used by the partners to keep the initiative on track and build sustainability
• Resilience so that the partnership is sustained
• Technology access is improved for equitable purposes, including reduction of the digital divide.

At the 2015 EDUSummit thematic working group 1 (TWG1) discussed these characteristics. Discussion centered on issues such as the stakeholders involved in partnerships, ways to ensure the participation of all stakeholders, the tension in developing a shared vision, trust and respect across and between partners, the need to address scalability and to ensure sustainability, and how to apply ICT “smartly.”

To deepen understanding of smart partnerships, three illustrations were selected for further analyses each of which was identified as having many or all of characteristics of a Smart Partnership. While the sampling approach was convenient, the illustrations were also selected to be as different from each other as possible. Integrated ICT in schools and communities in rural and remote regions, in India, is large-scale, organic approach with evidence of sustained growth addressing needs of underserved children and youth. The MARA SmartEdu Partnership, in Malaysia, is linked with the Government’s vision of smart schools that adopts a largely top-down approach. OpeningupSlovenia’s multi-stakeholder partnership takes the form of intersecting projects among which there is analysis of big data (Urbančič & Orlič, 2016).

Davis’ (in press) Arena is an ecological framework can enable perception of influences from the local to the global scale. This ecological perspective has assisted the development of inclusive worldviews that promote more equitable
approaches to education, supporting UNESCO’s work and recommendations that began with the World Summit on the Information Society (ITU, 2015). This framework, conceptualized as an Arena of Change with Digital Technologies in Education (see Figure 1), was deployed to inform the updating of UNESCO’s global survey of ICT infrastructure schooling (Twining et al., 2015). The Arena of Change with Digital Technologies in Education (Davis, in press) was used as a tool to critically analyze each of the illustrations and it helped to clarify the interaction between global and local educational ecologies on four main axes: political, bureaucratic, professional, and commercial (including community and open educational resources).

An Arena map of each illustration was constructed (Figures 2-4) to identify the partners and organizations involved and the range in which they operate, and also to clarify the interaction between global and local educational ecologies on the four main axes. The mapping process was useful for identifying the remit of partners in the partnership, as well as demonstrating the synergistic relationships across partners. A large number of overlapping and nested ecologies indicated complex local, regional and global interactions.

Revisiting the literature after the mapping exercise enabled TWG1 to identify seven characteristics of Smart Partnerships which were subsequently refined to form a definition. Each illustration was analyzed for evidence of these seven characteristics to establish whether or not, each could be considered a smart partnership. Preliminary analyses demonstrated that only one could be described as a smart partnership. The illustrations and literature review were developed further after the EDUsummit.

Characteristics of Smart Partnerships

Grounded in our work to date, we propose that Smart Partnerships in education have most or all of these seven characteristics:

1. Include partners within and across education (including teachers, their organizations, and researchers), government (of education, commerce & law enforcement etc.), industry, communities, and civil society (e.g., non-governmental organizations, NGOs).
Have a shared purpose (values, concept vision) that evolves into a synergy (more than a sum of the parts).

Have a strategic and holistic approach.

Enhance the quality of education with digital technologies (ICT).

Harness ICT smartly (e.g., evidence immediately deployed to improve performance).

Recognize their role in the emergent process(es).

Facilitate their own organizations to change.

While it is possible to have a smart partnership with a limited scope, a large initiative to enhance the quality of education with digital technologies (ICT) for a region is more likely to be sustained with a Smart Partnership that encompasses all seven characteristics listed above. It is also noted that a Smart Partnership may include one or more smaller Smart Partnership(s) within it. However, where few of characteristics are evident, the “smart” nature of the partnership is very limited.

Two illustrations of potential smart partnerships

Two of the illustrations of potential smart partnerships are elaborated and discussed in this paper, both of which aim to promote equity; and each of which has been developed with one author as a key informant. The Slovenian illustration, OpeningupSlovenia, was considered to be at too early a stage to inform our current work and so, it was moved to the appendix.

India: Integrated ICT for underserved children and youth

The Integrated approach to Technology in Education (ITE) is an initiative of the Tata Trusts in twelve mostly rural locations in Eastern and Northern India reaching over 17,000 underserved children by 2015 and continuing to grow. Initiated in 2012, the central concept of ITE is to integrate digital technologies into the curriculum and pedagogy of participating schools and learning centers (Charania, 2012-2014). Adopting a constructivist pedagogical framework, ITE seeks to improve teaching and learning, and foster authentic and project-based learning for young people in some of the most underprivileged Indian regions. The students, mostly first time computer users, create learning artifacts such as weather charts, jute production in India, or population density in different cities to deepen their learning of curriculum content. In this way, technology is embedded within curriculum design and pedagogy (Charania & Davis, 2016).

The Arena for the ITE smart partnership in India is depicted in Figure 2. At the center is a learning center for adolescents in a public school in which trained coordinators and teachers implement ITE. These are nested in clusters of villages at the district level in which parents, committees and their leaders are also central stakeholders for implementation. At the state level is the non-governmental organization (NGO) who hosts partnerships between district authorities and the Tata Trusts. In relation to global and local ecologies, the political axis includes the state party and national party while the bureaucratic axis identifies the provision of textbooks at the state level plus the Ministries of Education and Human Resource Development (MHRD) at the national level. UNESCO lies at the intersection of these axes. On the professional axis are Tata’s Institute of Social Sciences and the external evaluators of ITE at state level. At the global level are the Massachusetts Institute of Technology (MIT), International Society for Technology in Education (ISTE) standards, and interns from the University of California at Berkeley. The Tata Trusts are positioned at the intersection of the professional/commercial axes and the NGOs (e.g., Suchana NGO) at the commercial/political axes intersection. On the commercial (community/OER) axis some companies such as Google and the NIT (Next Infocommunications Technology) service provide tools and infrastructure used by students and teachers.

More detail of the development, including an analysis of Smart Partnership characteristics, is provided in Charania and Davis (2016) which identifies evidence of all characteristics authenticated through external evaluations (Myers & Zhao, 2013). Key points are provided below with numbers of the seven characteristics previously listed, indicating the characteristics identified:

- The scaling up of the partnership across four phases has involved many partners including NGOs, schools, communities and service agencies - even police in one community. (1)
The visibility of students’ artifacts online facilitates dissemination across India, supporting others to adopt and adapt the approach to meet their learners’ needs. The digital artefacts smartly serve multiple purposes, even though analytics have yet to be deployed. (2, 3, 4, 5)

The central role played by the Tata Trusts and their commitment is acknowledged as a strength. This enabled partners to scale their impact in collaboration with many other stakeholders who may become less peripheral in future as they facilitate their own organizations to change. This is likely to be most challenging for such influential stakeholders who have conflicting interests. (6, 7)

Figure 2. Arena of the ITE initiative in India with learners in schools and madrassas at the center

Malaysia: MARA “Learning Powered by Technology”

The Malaysian education system is highly centralized with each school following the same curriculum, policies and teaching programs. Two different agencies manage schools: the Ministry of Education and the MARA agency (translated as Council of Trust for the People) under the Ministry of Rural and Regional Development. National policy in Malaysia articulates a vision to transform schooling, and reduce the digital divide between urban and rural students by providing quality Internet-enabled education to all schools throughout Malaysia (Ministry of Education Malaysia, 1998; 2012).

In 2014, 12 MARA Junior Science Colleges (similar to high schools) in rural areas were selected for an innovative approach to schooling called “Learning Powered by Technology”. As part of the initiative, all 12 colleges were provided with technology infrastructure for administration, learning and teaching including high speed wireless connectivity, architecture design; content repository and cloud computing through the nationwide online platform, SmartEdu (see Nordin & Davis, 2015). A program of professional development for two or three teachers from each college (“change ambassadors”) was implemented and coordinated by teacher educators at Universiti Utara Malaysia (UUM), primarily focusing on pedagogy and content. These change ambassadors subsequently had responsibility for professional development and supporting teachers’ use of the SmartEdu platform at their college. Each teacher involved created a course for their subject area, uploading learning materials, activities, quizzes and assignments.
onto the SmartEdu platform. Representatives from each college showcased their school’s outcomes at the MARA cation Summit in November 2015. Five participating teachers were subsequently selected as Microsoft Innovative Educator Experts.

The Arena for this potential Smart Partnership in Malaysia is presented in Figure 3, placing a teacher with class of students at the center. The interaction between global and local ecologies on the political axis are nodes with politicians who sit on the MARA governance board while the Ministry and MARA agency is positioned on bureaucratic axis. As with the previous illustration from India, UNESCO lies at the intersection of the political/bureaucratic axes. The Universiti Utara Malaysia as coordinator of the professional development program is placed on the professional axis. On the commercial axis is range of providers that support SmartEdu. Telecom (™) provide the wireless network and Content Capital providing the learning platform are present at national level; Microsoft software (MS) is placed at the global level plus its MS Educator Network and Innovator Educator program.

Analysis of the initiative using the characteristics of Smart Partnerships suggests that only the final characteristic was missing (#7 in the list of smart partnership characteristics). Key points are (followed by the number(s) of the characteristic(s) covered):

- Partners across government, education, business and industry are stakeholders in the initiative. (1)
- Partners designed and implemented the initiative which was underpinned by four clear and shared aims, including ICT and equity. These aims are operationalized through the provision of the online platform and centralized professional development. Each partner has a clearly defined role. (2, 3, 4, 6)
- The analytics linked with the digital learning platform, SmartEdu is not accessible to inform the partnership. However, like the ITE initiative, artifacts are emerging online to inform development. (5)
Discussion of smart partnership illustrations

Of the two cases described, only one, the ITE initiative in India, can be recognized as a Smart Partnership with all seven characteristics. The characteristics are discussed below.

Smart Partnerships have partners within and across education government, industry, communities, and civil society

As stated, each illustration has partners within and across education, government, industry, communities, and civil society. For ITE in India (Figure 2), the multi-stakeholder partnership evolved strategically and holistically over time as the intervention matured and mainstreamed in the public system. In the first year of the initiative when it was implemented across four learning centers the lead stakeholders were an NGO and the Tata Trusts (Charania & Davis, 2016). The initiative was scaled up in the second and third years of the program, increasing the number of the NGOs adopting the program from 1 to 18 and the number of young people participating from 1000 to 22000. Partnerships were also established with education departments in districts and their states that spread across India as interest grew in adopting the ITE program in public schools. This led to the Tata Trusts working with teachers, schools and district education authorities. Local communities comprising of parents, village committees and heads, and youth forums were also involved and continue to be important stakeholders for implementation on ground. Higher education and related professional networks are also key partners, forming global networks which have expanded over time. Partnerships have been established with U.S. Universities, whilst industry partners include Google and telecom providers. Involving a wide-range of stakeholders can improve educational outcomes (Draxler, 2008; Sarvi et al., 2015) and is a characteristic of a successful partnership that increases equity (Cassidy, 2007; Geldof et al., 2011).

In Malaysia, the SmartEdu initiative (Figure 3) comprises a partnership across government, education, business and industry. The Ministry of Education is the lead partner with overall responsibility with other partners providing the services and resources to the schools. Telecom (™) and Microsoft provide the technology facilities in the colleges while an e-learning company and a university provide the professional development. The participating schools are also active partners in that the “change ambassadors” and other teachers work together to implement the initiative on the ground with colleagues, students and parents. Commitment from a high-level leader such as a Ministry of Education has assisted buy-in from participants (Osei-Kyei & Chan, 2015).

These illustrations thus provide ample evidence of partnerships within and across education government, industry, communities, and civil society in each of the partnerships. They also highlight the complexity of the structure of these partnerships. Drawing on Grobe (1990), these partnerships are complex because they have multiple partners, often with multiple partners from each sector, and each partner has substantive responsibility within the initiative. There are also two or more levels of partnerships in each of the projects. It is also noticeable is that the structure of the partnerships continues to change. The ITE partnership in India for example, has continued to evolve as the partnership expands and matures and this results in complex dynamic structures.

Smart partnerships have a shared purpose that has evolved into a synergy with a strategic and holistic approach

There is a shared purpose across partners in both initiatives as partners work towards a common goal. The purpose of the Indian initiative is articulated as a set of four key objectives and strategies which broadly relate to enhancing educational outcomes and bridging the digital divide among students in disadvantaged settings. These objectives and strategies are the key focus of initial workshops and events in which new partners participate. Similarly, the Malaysian initiative is grounded in a set of four broad aims which relate to enabling teachers to use digital technologies for learning and teaching in their subject areas. These aims are operationalized through the provision of the online platform and centralized program of professional development. In both, this approach has not only led to a shared purpose among the partners but also supports implementation and most importantly, preserves the integrity and objectives of the initiative. Moreover, partners within each of these initiatives fulfill a specific role or responsibility; complementing each other’s capabilities and resources (human and technical), thus making the partnership stronger through the sum of its parts. Sharing aims and objectives is one of the key characteristics of successful partnerships (Falloon, 2015; Sarvi et al., 2015).
Smart partnerships enhance the quality of education with digital technologies that harness the technology smartly

Given that SmartEdu as part of the Learning Powered by Technology initiative was established in 2014, it is likely that it is too early for evidence of enhanced quality of education to emerge. Despite this, in Malaysia the emergence of a more creative student-centered approach in which both students and teachers use technology for learning and teaching has been noted by the selection of five teachers as Microsoft Innovation Experts.

There is evidence of enhanced quality of education that harnesses technology smartly from the ITE initiative in India. Prior to participating in the ITE initiative, most of the students had seen computers but did not use them at school, and teachers primarily enforced rote learning of content to pass school exams (Charania, 2015). Evaluation of the ITE program (Charania & Myers, 2014) highlights the main impacts as (i) improving student-teacher relationships, (ii) increased student interest in many subjects including languages, sciences and mathematics (iii) use of constructivist pedagogical processes and (iv) increased student confidence in using digital tools and the Internet.

In addition, a “smart” use of ICT has evolved over time as ITE made more efficient use of digital tools to facilitate feedback (and feed forward to new partners). In ITE one of the most useful indicators to drive forward planning is the evaluation of projects created by students. Initial arrangements entailed the sharing of student work by CD and e-mail. This was followed by the creation of a blog through which student projects could be uploaded to the Internet and most recently, a website has been created for this purpose. The inclusion of a discussion forum on this website further extends the feedback process by enabling teachers to share experiences, challenges and best practices. These uses of ICT improve efficiency and support data-driven decisions (Kitchin, 2016) although there is no evidence to date of the adoption of data analytics. However, this does meet with the definition of a “smart school” being a systematic re-invention of teaching and learning practices (Choi et al., 2015).

In smart partnerships, partners recognize their role in the emergent process and facilitate their own organization to change in ways that promote sustainability and scalability

There is no evidence to suggest that partners in the Malaysian initiative have facilitated their own organization to change in ways that may sustain and scale the Smart Partnership. However, participation in the ITE initiative has enabled some of the partners to facilitate change in their own organization. This was particularly evident in a district of Assam. Prior to joining the ITE initiative, the mission of the NGO working in Assam was sustainable resource management. Therefore on joining ITE, staff and volunteers at this NGO had a very limited range of educational strategies. Through adoption and implementation of ITE, they began to work firstly in a small number of adolescent learning centers in villages and later in 50 district schools. This led to implementation of ITE in the state model schools, with the NGO being the implementation organization. The NGO has thus evolved from a grassroots level organization in livelihood and child protection, to become a recognized education resource organization with the credentials to work with the state education department. This partnership between the Tata Trust and the NGO has become the largest scale implementation for ITE.

Conclusion and recommendation

It is striking that our analyses of both illustrations provide evidence of the lack of adoption of data analytics in Smart Partnerships in education and in classroom contexts that are targeting disadvantaged young people, despite increasing adoption in school contexts (Johnson et al., 2016). Some “smart” uses of technology were observed in the most developed partnership (ITE) leading to increased efficiency and changing practices. We are pleased to note that, unlike the findings of Patrinos et al. (2009), there was no evidence of these partnerships increasing inequality or the digital divide; rather, they improved education equality and improved student access to technology through additional infrastructure.

This paper has elucidated the emergent concept of Smart Partnerships and provided a definition. Of the two illustrations discussed here, only one displays all seven characteristics. It is possible that the other example will evolve into a Smart Partnership and we believe that such evolution could increase its impact in education while addressing the digital divide. These illustrations may also prompt smart partnerships to emerge to serve children and youth in other parts of the world.
We conclude that further research into Smart Partnerships is essential to inform the development of scalable innovative funding mechanisms that are likely to secure the resources to unleash the full potential of digital technologies in education. It would be beneficial to conduct more in-depth research into the success factors underpinning the illustrations from India and Malaysia. In turn, findings should be used to refine the definition of Smart Partnerships proposed in this paper. Identifying further examples of full or emerging Smart Partnerships would further strengthen understanding in this field. Moreover, specific research is required to fully understand the role of data analytics in such partnerships and other smart uses of technology to maximize opportunities for data-driven decision making and further improvements to teaching and learning through real-time monitoring and feedback that do not increase the digital divide.

Acknowledgements

Thanks to all of those who have been involved in the work of TWG1 before, during and after the 2015 EDUsummit, including the EDUsummit’s leaders and UNESCO Bangkok, who have supported our work. We also acknowledge all of those who have contributed to the illustrations in this paper.

References


Appendix 1. Slovenia: OpeningupSlovenia

In Slovenia, OpeningupSlovenia (OuS) was established in 2014 as a way to promote learning and teaching through ICT and digital content, particularly Open Educational Resources (OER) and Open Technologies for Open Learning. The initiative is designed to provide a case study for the European Commission’s policies on open education (European Commission, 2014). Supported by national policy (Republic of Slovenia, Ministry of Education, Science and Sport, 2014) and driven by research, this top-down/bottom-up initiative is a partnership of 15 institutions across all levels of the Slovenian education system i.e., schools, higher education, vocational education and training, lifelong learning institutions and companies (see OpeningupSlovenia http://www.ouslovenia.net/). As part of the initiative, partners work together to design and implement innovative cross-sectoral projects including the application of theory and cutting-edge technologies such as Artificial Intelligence and Big Data (Urbančič & Orlič, 2016). The projects are intended to create opportunities to innovate for organizations, teachers and learners to help participants acquire digital skills and develop new ways of learning as well as to support development and availability of OER and open data. By 2015, 27 projects were running, each with a different set of partners. Examples include:

- “My Machine” (http://mymachine.si/) in which children from kindergarten and primary schools design their dream machine, solutions are proposed by higher education students with the best ones selected by the children before a prototype is built by students in Technical Oriented Secondary Schools. In addition to schools partners include the Institut Jozef Stefan, Ljubljana and OS Savska naselje. The project is linked to the My Machine Global project (http://www.mymachineglobal.org).

- TraMOOC (http://www.ouslovenia.net/project/tramooc/) is a project, involving universities, research organisations and industry that started in 2015 to develop high-quality translation of all types of text included in MOOCs, such as assignments, tests, presentations, lecture subtitles, and blog text. MOOC resources are likely to be valuable to schools so they are users, rather than full partners. The online translation platform aims to provide translation between English and the languages of sixteen countries.

- A UNESCO Chair on Open Technologies for Open Educational Resources and Open Learning (http://unesco.ijs.si/) at the Institute in Ljubljana, Slovenia was established in 2015 to link these developments with the global network of OER chairs in other countries. The Chair has operated with existing EU FP7 research projects and new projects funded by the Government of Slovenia and European Commission.

The Arena for this potential Smart Partnership in Slovenia sketched in Figure 4 indicates the pivotal role of research projects involving students (S), teachers (T), researchers (R) and policymakers (P) who are placed at its center. In relation to the interaction between global and local ecologies, on the political axis the partners include research institutions together with UNESCO, the European Commission, and a federal agency in the USA, all of whom feed into policy adoption in Slovenia. The Slovenian Ministry of Education and the European Commission are identified on the bureaucratic axis. The professional axis includes selected partners who are currently involved in OuS that are fed research data by students and their teachers: Knowledge 4 All Foundation Ltd., Massachusetts Institute of Technology (MIT), Stanford, the European Organization for Nuclear Research (CERN), and 60 Artificial Intelligence labs. On the commercial/OER axis are research projects that are developing tools for use in Slovenian schools and other educational contexts. Open research methodologies are placed on the professional/OER intersection.

Analysis of OuS as a smart partnership reveals few of the characteristics of a smart partnership to be present. Although the initiative involves a range of partners across government, education and industry and is underpinned by a set of four broad objectives, it comprises of a number of projects which are essentially stand-alone units. In addition, although it is planned that some of the projects will make use of cutting-edge technologies for data analytics, there is of yet no evidence of this.
Figure 4. Arena map of OpeningupSlovenia and its partners with students and their teachers in school at the centre who are engaged with researchers.
A Smart Partnership: Integrating Educational Technology for Underserved Children in India

Amina Charania1* and Niki Davis2
1Tata Trusts, India // 2University of Canterbury, New Zealand // aminasachin@gmail.com // niki.davis@canterbury.ac.nz

*Corresponding author

ABSTRACT
This paper explores the evolution of a large multi-stakeholder partnership that has grown since 2011 to scale deep engagement with learning through technology and decrease the digital divide for thousands of underserved school children in India. Using as its basis a case study of an initiative called integrated approach to technology in education (ITE) spearheaded by Tata Trusts in India working in partnership with an increasing number of other organizations, this paper seeks to illustrate what a large multi-stakeholder partnership in education is and how it can work to serve education equitably. In addition to tracing the growth and development of this multi-stakeholder partnership, the paper examines the ITE-related work partnership against seven characteristics of education-industry/community partnerships re-interpreted in terms of smart partnerships at EDUsummIT 2015.

Keywords
Smart partnership, Education industry partnership, Change with digital technologies, Integrating ICT, Digital equity

Introduction

Today, education-industry partnerships can assist schools to expedite the integration of digital technologies in their pedagogy and administration (Eickelmann, 2011). Multi-stakeholder partnerships known as smart partnerships (SPs) may be particularly valuable in terms of supporting the development of more equitable educational infrastructure and reducing the digital divide, which is the gulf between those who can readily and effectively access information and communication technologies (ICT) and those who, for various reasons, cannot.

According to Snow (2011), “Smart partnerships are collaborations linking the assets and initiatives of institutions with community assets and interests for powerful long-term impact.” UNESCO is just one large-scale educational stakeholder that promotes the multi-stakeholder partnerships as a means to “create equitable, dynamic, accountable and sustainable learner-centered digital learning ecosystems” (UNESCO, 2015, point 19). However, as Leahy et al. (2016) point out, scholarly literature contains little information on the development, utility and effectiveness of education-related multi-stakeholder partnerships that have “smart” learning environments (in relation to smart learning environments see Kinshuk et al., 2016).

At the EDUsummIT 2015, a discussion of multi-stakeholder educational partnerships recognized that some may be characterized as smart partnerships (SP) when they meet the following seven criteria (Davis et al., 2015; Leahy et al., 2016):
- draw partners from within and across a wide range of educational enterprises and stakeholders
- have a shared purpose (values, concept, vision) that evolves synergistically
- have a strategic and holistic approach
- enhance the quality of education via digital technologies (ICT)
- harness ICT smartly in order to monitor educational outcomes and provide feedback aimed at improving performance
- recognize their role in the emergent process(es)
- facilitate change within their own organizations.

In this paper, we explore the growth of one partnership that is developing large-scale integration of technology in teaching and learning through the initiative called Integrating Technology in Education (ITE) for school age children in India, which began its work in a remote area of the Eastern part of the country. The central stakeholder in this initiative from the start is the Tata Trusts, a philanthropic organization committed to the betterment of India and its people, and the purpose of the ITE initiative is to enhance deep learning experiences using technology of the upper-primary and secondary school aged children and adolescents which would also bridge the digital divide. In many
parts of India, the digital divide is all too apparent and is mediated by the variables of gender, age and socioeconomic status (see, e.g., Eamon, 2004). Socioeconomic status is particularly important with respect to the initiative discussed in this paper because it is a primary reason for the high dropout rate from school in many parts of India (see, e.g., Pankaj & Poornima, 2008). ITE objectives focus on ameliorating the effects of this type of inequality.

We begin our account with a brief explanation of how we conducted the study. We then trace the growth and development of the Tata Trust’s ITE-focused partnership, after which we illustrate and discuss how and to what extent the partnership exhibits the seven SP characteristics listed above.

![Map of India with sites marked in red, green, and purple](image)

**Figure 1.** The spread of the ITE smart partnership in India from 2011-2015

**Method**

We used case study methodology to generate our data, which came from more than 20 documentary sources associated with the SP and its ITE initiative and from interviews, conducted by the second author of this paper, of the first author (an initiating stakeholder in the development and implementation of the initiative). Both of us reflected on and analyzed the data. The seven characteristic of a SP were applied deductively to seek out, identify and then describe each aspect. The study built on our collaboration with Twining to inform UNESCO Institute of Statistics development of new indicators for ITE infrastructure in education (Twining, Davis, & Charania, 2015) and work described in Leahy et al. (2016) to map the ITE SP on Davis’s Arena of change with digital technologies in education, which was also used in Twining et al. (2015). Like other scholars who study practice in which they are
immersed, we were aware of the potential for bias. However, because the second author is not involved in the SP or the ITE, we consider that she was able to bring an independent perspective to her questioning and data analysis, especially in regard to probing for the complex and organic processes occurring in the SP’s development to date. We also gave all other stakeholders involved in the SP opportunity to check and adjust, as they saw fit, the findings emerging from our analysis.

**Growth of the ITE multi-stakeholder partnership**

Our examination of the information that we collected allowed us to identify four clear development phases. The map in Figure 1 illustrated the growth of impact of the ITE SP starting with one site in phase one (red), the five sites added in the establishment phase two (green), and expanding to all other sites indicated by the circles on the map during scaling up phase three (purple).

The phases are described below and the evidence is summarized in Table 1.

- **Phase 1. Creation of vision and concept:** The ITE initiative began in 2011 when the first author presented the ITE concept at North East Hill University conference. The first author’s doctoral studies and research at Iowa State University Center for Technology in Learning and Teaching and with the second author heavily influenced the development of the concept. The concept was further developed as student centered use of technology for meaningful learning. This concept was then tried with a grant from Tata Trusts to the already supported non-governmental organization (NGO) Street Survivors so that they implemented ITE in their four coaching centers in a remote and challenging part of the Murshidabad district of West Bengal abutting the India-Bangladesh border (Charania & Meyers, 2014). The ITE approach calls on teachers to design lesson plans focused on encouraging students to create learning projects based on the curriculum using digital technologies and the Internet. The pilot was successful in terms of gaining students’ interest in learning school subjects and teachers’ ability to integrate technology-enabled student projects within curriculum, and it was at this point that the multi-stakeholder partnership really got underway. The approach also worked strategically to improve access to these technologies for underserved children in Murshidabad by generating expertise and resources that could be reused in increasingly “smart” ways to support the spread of the ITE in other Trust supported education projects. This experience at the pilot and the concept paper formed the basis of a presentation to the Tata Trusts Board, which agreed to spearhead expansion of the ITE initiative in other projects supported by the Tata Trusts in India.

- **Phase 2. SP establishment:** During this phase in 2012/2013, the ITE initiative spread to other locations in West Bengal and Uttar Pradesh, mostly to education projects already supported by the Trust. For example, the Vikramshila Education Resource Society (another West Bengali NGO) joined with the ITE initiative at its learning centres in partnership with police and other community members; it was able to draw these partners into the SP. Similarly, the Trusts’ partnership with the Suchana NGO (also in West Bengal) enabled ITE to be layered into this NGO’s work with tribal communities in West Bengal (see Charania, 2015 for a description). Meanwhile, NGOs working in madrasas in Uttar Pradesh reallocated their existing budgets so they could integrate the ITE into their work. Each new partnering expanded the number of children served, broadened expertise across language and cultural parameters and drew in and made more cost-effective use of resources. These outcomes, along with external evaluations in the form of interns’ and consultant’s reports, papers and video recordings, confirmed the ITE concept as highly relevant to the partners and the children they were serving. By the end of its first two years, the partnership had reached out to and decreased the digital divide for 10,000 young people in the state of West Bengal and Uttar Pradesh. The ITE initiative had also begun spreading via existing Trust partners to other states.

- **Phase 3. SP scaling up:** Over 2013-2014 the Tata Trusts’ decision to scale up the ITE became a strategic decision. Trust leaders sought partners’ advice on securing support and collaboration from government bodies and also agreed to bring in an additional project called CLIx developed with the Massachusetts Institute of Technology and Tata Institute of Social Sciences (Tata Trusts, 2016b), which would later facilitate the scaling up. In addition to “layering” ITE into partners’ work, Tata staff deliberately sought government involvement so as to gain the leverage of its infrastructure and powers of authorization. The scaling up jumped significantly when new partners in the state of Assam enabled outreach to 74 schools and about 12,000 more students. In addition at this time, Vikramshila extended the ITE to about 35 government schools and madrasas in West Bengal and an increasing number of NGO partners made it possible to reach children in additional states and邦(cols:604-1014)}
districts, including those tending to have very low rates of literacy and retention in upper primary schools (see http://www.dise.in/drc.htm). Additional partners, such as Digital Empowerment Foundation NGO, were recruited to support innovations that increased access to ICT infrastructure with community-led wireless towers. When a private school (Samaritans Help Mission) joined as a partner, it brought more flexibility in timetabling and resources and the partnership became better able to trial innovations, such as students using Skype conferences to share their project work with neighbouring government schools and madrasas.

- **Phase 4. SP consolidation:** Mainstreaming of the ITE in government schools augurs well for sustainability in the future. The integration with the CLix project is likely to increase the partnership between the initiative and Massachusetts Institute of Technology. The partnership has also increased its attention to address risks, such as that of cyber safety. Although Google-derived resources supported cyber safety in the earlier phases, the need to more firmly integrate cyber wellness into the initiative is imperative, especially for more vulnerable populations where there is little history and experience of ICT in society (Park, 2015). A proposal for the Tata Institute of Social Sciences to convert ITE training into certified courses for government-appointed and NGO teachers is also under consideration. This will call on more direct partnership with Tata Institute of Social Sciences and other experts from international universities. Additional challenges will also be addressed, such as increasing capacity for leadership within the partnership and for maintaining the position of ITE among larger projects that are supported by the Tata Trusts.

**Table 1. ITE evidence summarized over the four phases**

<table>
<thead>
<tr>
<th>Phase</th>
<th>Year(s)</th>
<th>Partners (place, state or region)</th>
<th>Key processes</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Creation of ITE vision and concept (2011/12)</td>
<td>2011–12</td>
<td>Street Survivors NGO (West Bengal)</td>
<td>Field visit and further development of ITE concept with SS executive; Tata trustees at SS give approval for ITE pilot project in the form of computer education</td>
<td>SS integrates ITE component at four learning centers</td>
</tr>
<tr>
<td>2012</td>
<td>Sir Dorabji Tata Trust/Tata Trusts NGO (nationwide)</td>
<td>Concept creation, discussion of pilot implementation at SS. Approval of the concept at the trustee meeting</td>
<td>Permission gained to develop ITE proposals for Tata funding with existing partners</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Four SS learning centers + related communities and schools (West Bengal)</td>
<td>Two trainings in Murshidabad, with enthusiastic uptake</td>
<td>Exemplary products: students’ projects, teachers’ lesson plans; some teachers become first ITE “heroes” keen to spread the practice.</td>
<td></td>
</tr>
<tr>
<td>(2) SP Establishment (2012/13)</td>
<td>2013</td>
<td>University of California at Berkeley at Davis (USA)</td>
<td>External evaluation documented interns etc.</td>
<td>Outputs (Myers &amp; Zhao, 2013; Charania &amp; Myers, 2014)</td>
</tr>
<tr>
<td>2013–</td>
<td>Vikramshila Education Resource Society NGO (West Bengal)</td>
<td>Introduction of ITE including training for volunteers and partner schools</td>
<td>ITE expands in madrasas Online portal for student projects launched</td>
<td></td>
</tr>
<tr>
<td>2013–</td>
<td>Nalanda and PVCHR (Uttar Pradesh)</td>
<td>Online teacher portal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013–</td>
<td>Suchana NGO working with tribal children</td>
<td>ITE volunteers training other organizations Innovative tribal ITE</td>
<td>Expansion of ITE to range of languages and cultures ITE teachers leading</td>
<td></td>
</tr>
</tbody>
</table>
and schools in remote region (West Bengal) implementation including teacher trained to drive ITE van bringing ICT access to remote areas innovative strategies enable implementation in remote areas

(3) SP Scaling up (2014/15) 2014 Massachusetts Institute of Technology (USA) Study visit to USA for key ITE staff facilitated by Tata Trustee Negotiation of CLiX connected learning initiative State and district authorities agree to leverage infrastructure for ITE in schools and validate participation Ministry-provided list of Assam and West Bengal schools promotes ITE engagement

2014 Ministry of Education state and district (West Bengal & Assam) Three additional NGOs adopt ITE approach in their work

2014– South Odisha Initiative, Samaritans Help Mission, Society for Awareness Harmony and Equal Rights Introduction of wireless access

2014– Digital Empowerment Foundation NGOs (New Delhi) Multiple meetings with state mission director including ITE teacher training

2015– Assam Ministry for secondary education (RMSA) (Assam) Implementation of ITE in more schools, including model schools in Assam

(4) SP Consolidation (2016– ) 2016– Tata Institute of Social Sciences (Mumbai) Proposal to institutionalize ITE in the Centre for Education of Innovation and Action Research Improved alignment between potentially competing projects University accredited teacher training

Extent to which ITE aligns with the seven characteristics of a smart partnership

SP include partners within and across educational organizations

ITE clearly meets the first SP characteristic. Tata Trusts rapidly reached out to other partners, especially between 2013 and 2015. The partnership, which began with Tata Trusts as the key partner in 2011, quickly brought in other partners. After the pilot in 2011/12 involving one NGO providing four coaching classes for 1000 children, the partnership expanded to 18 partners reaching about 17,000 children by 2015. Partners by this time included not only NGO members, but also researchers (education, technology and development studies), government agencies (education), civil society such as a police agency, higher education and related professional networks, industry-based organizations, and local communities made up of parents, village committees and their leaders, as well as youth forums. Local communities have been part of the partnership from the start and they, supported by the NGOs, who actively engage with them on the ground, have continued to be central stakeholders in the implementation and sustainability of the initiative. One of the NGO partners brought a strong connection with local police by setting up learning centers in the premise of police stations. This is an initiative of Kolkata police with the NGO to work closely with the children of migrant and slum dwelling communities. Partners also brought with them additional partners, such as particular types of schools and their communities like madrasas and global networks via the higher education and professional agencies. Since 2013, the Tata Social Internship program has organized internships on the ITE initiative. The first two interns, Masters students from the University of California at Berkeley, contributed
essential evaluation analyses at the end of the first year of implementation (Myers & Zhao, 2013). In 2015, another intern used video and photo-voice to document the impact of ITE at five organizations, which was uploaded onto internet as a YouTube clip in February 2016 (unpublished document specifying Tata Trusts, 2016a [video]). In addition, two scholars from Trinity College Dublin provided valuable feedback to ITE during extensive field visits in 2015.

**Partners have a shared purpose that evolves into a synergy**

Partners have a shared multi-faceted purpose clearly stated in ITE objectives that are to
- bridge the digital divide and foster digital citizenship
- create interest in learning that also improves attendance and retention in schools
- enhance learning achievement
- improve learning processes and pedagogy.

These objectives are informed by four key strategies enacted through workshops, exemplars, and extensive partnership negotiations and support with the partner NGOs. After the first conceptualization of the ITE concept at a national conference in India, the shared vision sharpened through implementation of the ITE pilot. After the pilot, those NGOs involved by March 2013 jointly confirmed the initiative’s objectives. This shared purpose was then articulated within the grant proposals for the next batch of NGOs that wanted to implement ITE. The formal arrangements between and among existing and new partners also provide evidence and dissemination of a clearly articulated and shared understanding of the ITE. The shared purpose continues to be the key focus of the initial workshops and events that propagate the ITE initiative across and among new partners.

**A SP has a strategic and holistic approach**

Experience shows that the ITE approach is readily disseminated to and understood by new partners. However, implementation of the initiative has been problematic in some instances because of challenges that are mentioned below and the partners find ways to work around the challenges where possible. In doing so, they are illustrating the advantages (flexibility being one of them) that a strong partnership brings to educational change.

Workshops and other means through which the ITE initiative is propagated has kept the design, purpose and overarching approach of the ITE consistent across the participating organizations in different parts of the country and with partners who continue their work overseas. After the pilot, establishing and sharing the ITE approach with other potential partners (NGOs and other agencies) was not difficult because the student-made artefacts continue to become available online to clearly convey how the initiative works in practice. Showcasing these artefacts during meetings with potential partners led to the conclusion that, if students in such a remote area could create thought-provoking and creative learning artefacts using technology, then surely the approach could be more easily implemented in less challenging contexts. During meetings potential partners have also had hands-on experience of developing lesson plans and student projects because some participants role played as students. Furthermore, participants were asked to collectively engage in evolving the objectives of the ITE, and that process has led to a further shared elaboration of the ITE approach. Tata Trusts has continued to firmly articulate and thus legitimized this development when processing applications for funding grants by those organizations that wished to implement ITE.

Establishing a shared purpose between the Trusts and government departments has proved less straightforward. It is likely to be an ongoing challenge for several reasons, including staff changes in the departments particularly at top managerial level. For example, in 2015 Tata Trusts and the Assam Secondary Education Mission (RMSA) signed a memorandum of understanding in which the scope of the ITE was extended beyond its usual parameters to accommodate the professional development needs of the teachers, particularly those in the state’s 14 model schools. Although the state officials were interested in the purpose of ITE, their most pressing motive for adoption was that ITE could serve the state’s modernization agenda, which the model schools led. Regular communication between the ITE partners and witnessing of students engagement by the state officials is likely to influence these officials to support the roll out the initiative to more government-funded general schools in future because, as these officials see
the outcomes of ITE on the ground, they will appreciate the ITE’s wider purpose and deeper approach to learning (see Fullan & Langworthy, 2014, in relation to deeper learning).

Establishing a share purpose for ITE can also occur indirectly. In West Bengal, Tata Trusts did not have a direct role in formalizing the partnership with district officials; this was done by the NGO Vikramshila. This NGO had already been working on quality improvement in education and had built rapport with the district departments before it layered the ITE into the schools and government madrasas. It is useful to note that Vikramshila, having gained pedagogic expertise in the ITE initiative through its partnership with the Trust, became sufficiently confident to showcase ITE to the government. The exhibits of the students’ work and the teachers’ growing confidence in integrating ICT in their lesson plans were impressive. In Assam, although the NGO Gramya Vikas Manch has less experience in pedagogy and understanding of the state’s educational processes, it has strong connections with the district administrator and good reach on the ground and in the community. These advantages made it possible for the NGO to invite Tata Trusts to engage directly with the district to share the ITE concept. From there, the first author was able to conduct the Trusts’ ITE orientation and training sessions directly with government teachers, an approach that helped them and their leaders to understand the concept and that increased their level of readiness to adopt ITE.

**SP works to enhance the quality of education with the aid of digital technologies (ICT)**

The continuing flow of student-produced artefacts complemented by logbooks maintained by teachers (see ITE handbook, Charania, 2012–2014) provides one means of assessing the impact of the ITE partnership on the quality of education. Although more evaluation work is required, progress to date suggests that the initiative is having the desired outcomes. Saxena (2014), for example, evaluated the ITE implementation of over 300 students in four private madrasas supported by the Vikramshila NGO. Analyses showed significant improvement of scores on tasks that measured about 150 students’ ability to connect concepts with real life situation using technology as well as their research and analytical skills and this external consultant’s report concluded,

> [The] ITE component has made affirmative changes in the learning approach and environment within madrasas. Unlike the majority of IT interventions being carried, which are restricted to learning basics of computers and skills development around using certain software/applications, ITE intervention has focused on integration of technology in curriculum transaction and teaching learning processes ... ITE has fitted well with the curriculum, as it is woven into teachers’ lesson plans ... With guidance from teachers, children have made multimedia presentations on lessons related to vitamins, water cycle, fractions, etc. They also have access to internet ... it has also been challenging for the teachers, as children are learning faster and better than them! According to the management of madrasas where ITE is being implemented, the presence of computers in the madrasas has generated interest amongst students; they have become more regular and attentive [and] it has also got parents interested ... (Saxena, 2014, p. 15)

In addition, USA university intern Kato in notes accompanying his video of the implementation and ongoing ITE practice at five different learning sites stated that the ITE had,

> ... enabled students to achieve real-life experiences incorporating higher order thinking. ... For example, in the community project about the river in Khanjanpur, students in Suchana were able to learn and discover how the lifestyle and culture of the villages around the river evolved as the river channel changed over time. (Tata Trusts, 2016a)

**SP harness ICT smartly**

ITE partners focused on encouraging teachers and students to use digital technologies to further learning outcomes and the quality of education. One of the most important uses of ICT amongst the ITE partners is that of monitoring the engagement of teachers and students with the initiative and providing teachers and students with feedback on the quality of that engagement. These mechanisms, essentially feedback loops, are an intrinsic part of the initiative every time it is rolled out in a new setting. The stakeholders can efficiently monitor the spread of the initiative, collect information (including teacher and student feedback), and analyze it. Further analysis is also used to provide advice and ideas for ITE-related improvements in the schools and learning centers.
As ITE has expanded, so too has the volume of feedback, and ICT has become more important to expedite this process. An example of this developing practice is evident in one of the most useful indicators — the increasing number and quality of student artefacts. The first author and the master trainers initially shared these projects via CDs or email. In 2013, a volunteer from one of the partner organizations, Child in Need India, established a blog in BlogSpot (http://sdtite.blogspot.com/p/contact.htm) so that student projects could be uploaded to the Internet. As the blog grew and became unwieldy, another organization and later a partner, the Digital Empowerment Foundation (2009) was recruited to create an ITE website and upload student projects onto it. This portal is now being maintained by the web master at the Digital Empowerment Foundation, who has also become an ITE associate at Tata Trusts working with a number of resource volunteers from other partners. The portal invited teachers to provide feedback, engage in discussion to share challenges and best practice. Discussions are led by a volunteer at Vikramshila.

In 2014, after the master trainer resource group had been formed, its members received training to identify the projects that best fit the ITE ethos. This training has enabled them to recognize quality lessons and projects, and they have developed assessment rubrics that teachers and other stakeholders can use. That information has been collected electronically and evaluated by the first author and resource group as a result there have been several modifications to the scope and practice of the ITE, as follows,

- **Reduction of the number accompanied by an increase in the quality of the learning projects selected for exhibition on the portal:** Too many projects were produced in the initial phase of the initiative, which resulted in shallow learning. The first author introduced a logbook so that teachers and students could keep an account of the types of lesson plans and projects that meet the quality objectives in the ITE manual (Charania, 2012-2014). After consistent emphasis during all the training, some of the partners also began to map lesson plans with the level of learning achieved against Bloom’s taxonomy.

- **The inclusion of community projects:** Although the initiative did not emphasize community projects, some very good student-initiated projects appeared. As school and learning centers have become more adept at integrating ICT use into their curricula, community projects have become better established as part of ITE. These projects are also showcased on the ITE portal.

- **Installation of local language type fonts:** When it became evident that some children were struggling to express themselves because the interfaces they were using did not include fonts for their local language, volunteers installed these fonts for them. As a result, the quality of these children’s projects increased almost immediately.

- **Spreadsheet training:** The logbook was introduced to monitor and showcase the variety of ways teachers could bring ICT use into their lesson plans. Reports indicated that by 2016, only about a quarter of the student projects had included spreadsheets, probably because teachers are not proficient in spreadsheet use, so further professional development is needed.

**In SP Partners recognize their role in the emergent process(es)**

SPs provide for networking within and across the partner organizations. These networks allow partners to compare the activities, successes and challenges evident in their own organizations against those in the partner organizations. These insights not only allow organizations to effect improvements or bring in new ways of doing things in their spheres of activity but also to actively contribute their ideas for improvement and new strategies in partner organizations. Each year all partners of the ITE initiative meet to discuss their success, challenges, and solutions. For example, a partner with expertise in a particular area such as power backup-up and connectivity solutions will share their expertise that address another partner’s challenges and subsequent continue to support such partners to overcome the challenge. In this way, the partners gain an applied understanding of ways in which they can contribute to the partnership to assist in realizing its vision. Partners can also identify other organizations that they think may benefit from resource sharing. For example, the first partners in ITE lent their most experienced ITE coordinators to participate in an ITE exchange program, wherein the coordinators spent two weeks with newer ITE partners helping teachers in the classroom and leading some demonstration ITE classes. By engaging in these ways, all partners gain first-hand appreciation of their role in scaling up the partnering process and activities.
Another example of the emergent processes occurred in 2012 when Tata Trusts invited an NGO called Gramya Vikas Mancha (GVM) to work with its signature program in adolescent education. At the time, GVM was actively working in Assam’s flood-prone areas and was the Trusts partner for diversion-based irrigation. In addition, GVM was working on a child safety initiative with another prominent organization in India. Because GVM has a huge volunteer force and strong connections with communities and local administrations, they were highly effective in implementing the adolescent program, which sought in part to encourage out of school adolescents to return to school. In the process, the NGO developed good ties with the education departments in the districts where they operated. These ties, in turn, made it possible for Tata Trusts to encourage GVM to bring the ITE initiative into a number of village-based adolescent learning centers. Schools in neighboring vicinities became interested in what was happening and asked how they could participate. Tata Trusts provided some training, a process that facilitated interest from schools, learning centers and organizations in other districts and so further spread the ITE concept. So far GVM has brought an additional 50 schools into the ITE initiative so that ITE is now used in their classrooms by subject teachers. In addition, a memorandum of understanding between the state department and Tata Trusts to implement ITE in the model schools was initiated with GVM being the implementation organization. GVM’s role in the ITE partnership and their engagement with the districts working on the adolescent education project, along with their strong connections with the community and local schools, assured the state education department that this NGO had the identity and credibility to work effectively with the state. These emergent processes have scale up the ITE initiative effectively.

**Partners facilitate their own organizations to change**

GVM provides the strongest illustration to date of how the organizations involved in a SP can facilitate change in their own organizations, so that they are better able to improve educational outcomes for teachers and students, which is central to all educational SPs. Since becoming involved in the ITE, GVM has evolved from a grassroots organization involved in child protection to become an education resource organization in the state of Assam. This development occurred as the scope of the ITE increased with each new partnering between the GVM and district and state agencies. Initially, GVM’s typical community volunteers involved in the NGO’s child safety and adolescent education activities had no professional educational qualifications. The Tata Trusts’ ITE funding enabled GVM to hire ITE coordinators who had credentials in teaching, and special preference was given to those candidates who passed the teacher entrance test. In addition, GVM prepared its volunteers, program coordinators and credentialed teachers to liaise with state-government teachers and schools, a development that prompted requests from other district education departments to help them implement ITE in their districts. In promoting change within its own organization, GVM has been able to demonstrate its ability to become a leading NGO working with the state and the districts in the area of education.

**Conclusion**

This paper has provided an account of a large multi-stakeholder partnership that has co-evolved with educational technologies since 2011 to increase interest and engagement with learning through technology and bridge the digital divide for underserved students in India, including those in remote regions of the country. Over time, Tata Trusts, the founding partner which spearheaded the ITE initiative, has focused on scaling up the initiative by working with an increasing number of partners representing many different agencies, including NGOs and district and state government providers of education. This approach has been very different from the past where Tata Trusts had partnered with single NGOs on a project by project basis to fulfill its objectives. Tata Trusts has consistently moved away from its previous approach by drawing into the partnership (or encouraging its existing partners to do this) multiple agencies engaged in multiple interventions directed towards improving the lives of India’s most marginalized people. Although partners and their stakeholders come to the partnership with different motivations, the vision of the ITE partnership has worked organically to enable alignments of those motivations within the vision of ITE. This has been evident when partners witnessed the student engagement and teacher confidence in the showcases and elsewhere. However, the strength of the partnership has not yet been tested by the loss of any one partner, although that is likely in the future.
The integrated approach of the ITE SP described in this paper appears to be the key to its success because it has normalized collaboration through promotion of a vision, in this case using ICT to increase equity of educational access and provision, and build capacity for this work in education and related services. Importantly, that activity is underpinned by an organizational structure of collaborative partners that has resilience, not least because it is promoted and strategically funded by the core partner, Tata Trusts. However there are challenges that include building capacity for leadership, and spreading that leadership among the partners, rather than centralizing it within Tata Trusts. Although partners have led implementation on the ground, more is needed to extend leadership in research and ongoing strategic development.

Scrutiny of our case study evidence reveals that the ITE has all seven characteristics of SP partnership identified at the 2015 EDUsummIT, including embedding collaboration into the organizational cultures of the partnered organizations. However, some of these characteristics are more strongly in evidence than others. For example, while the ITE SP has been robust enough to encourage teachers to extensively integrate technology with curriculum, it has yet to find the expertise (whether within existing partner organizations or newly identified and invited partners) to develop and implement assessment practices. Similarly, although the ITE SP is providing some effective professional development for teachers, the need to ratchet up provision has been recognized. And although orientation and showcase events have included school leaders in hands-on activities that enable them to address misconceptions about ICT use in education, more support will be required to enhance their ability to integrate ICT into school curriculum (see, in this regard Mackey et al., 2015; Price, 2015). Given that both assessment and school leadership development remain common issues with ICT in educational contexts less challenging than those described for ITE in India (see, for example, Gibson, Downie, & Broadley, 2015; Fullan & Langworthy, 2014), it will be a measure of the utility and strength of the ITE SP if it were to effectively address these challenges in the future.

The SP characteristic (within characteristic 6) that may present ITE partners with the greatest challenge is to find ways to harness ICT to increase the rapid use of “smart” feedback from ongoing ITE interactions. For example, ITE has yet to design automated use of data analytics in its portal and there is little if any expertise in this area within the partnerships, including the implication for data security when working with children. This has also been hampered by limited digital access of many partners, particularly poor access for teachers and their leaders. However, this situation is likely to improve as the partnership grows and becomes more sustainable, because that will also facilitate improved access. Therefore it is recommended that the ITE initiative plan strategically to implement data analytics in the future so that it can inform their work with students as well as partnership processes so that this SP will become very smart indeed.

In conclusion, we recognize and applaud the way that the ITE multi-stakeholder partnership is building the capacity of teachers and students to make effective educational use of available digital technologies and to spread that capability among many languages and cultures to teachers and students that have been very much on the wrong side of the digital divide. This example in India provides an illustration of possibilities for other marginalized populations, including those in Africa and the Pacific. As more dynamic partners join this partnership and other SPs emerge to serve other regions, the divide could be reduced, which gives us hope that the World Education Forum (2015) may become a reality.

Acknowledgements

The authors wish to thank all of those mentioned in this article for their support and encouragement, including EDUsummIT colleagues. We also thank our reviewers for their advice and critique that we have used to improve the paper. We are happy to include the following statement at the request of The Tata Trusts. Education has been one of the key focus areas for Tata Trusts. For decades, the Trusts have worked towards enabling access of quality education for rural India. With the role of technology in the contemporary global context, Tata Trusts have shifted focus on bridging the digital divide and building 21st century skills for improved learning experiences.

References


Responding to Challenges in Teacher Professional Development for ICT Integration in Education

Jo Tondeur1*, Alona Forkosh-Baruch2, Sarah Prestridge3, Peter Albion4 and Shiyama Edirisinghe5
1Ghent University, Belgium // 2Levinsky College of Education, Israel // 3Griffith University, Australia // 4University of Southern Queensland, Australia // 5National Institute of Education, Sri Lanka // jo.tondeur@ugent.be // alonabar@levinsky.ac.il // s.prestridge@griffith.edu.au // peter.albion@usq.edu.au // shiyamaj@gmail.com

ABSTRACT
There is wide agreement that teacher professional development (TPD) is a necessary element in educational change, especially for the more effective application of technology to enhance learning. The research literature reports many examples of successful TPD but there remain many challenges to wider and deeper success in the variety of different contexts. Discussion by the thematic working group on TPD at EDUsummIT 2015 in Bangkok identified several challenges to successful TPD. This paper discusses those challenges, describes four cases of successful TPD from varied contexts, and derives a model for TPD based on observed commonalities in the cases.

Keywords
Teacher professional development, ICT integration, Case study

Introduction
Teacher learning for ICT integration was an important topic in the International Handbook of Information Technology in Primary and Secondary Education (Voogt & Knezek, 2008), which initiated the EDUsummIT meetings that included working groups on teacher professional development (TPD). In 2011, Thematic Working Group 3 (TWG3) on TPD at EDUsummIT highlighted the necessity of aligning action at multiple levels in order to ensure a shared vision of what is desirable and possible (Twining, Raffaghelli, Albion, & Knezek, 2013). Working from the recommendations of EDUsummIT 2011, the working group on TPD at EDUsummIT 2013 began with three foci for discussion, namely, engaging all stakeholders in developing a shared vision for ICT in education, engaging more teachers in communities and networks for PD, and reducing the gap between educational research and practice. Their discussions reaffirmed the importance of engaging all stakeholders in decisions, promoting networks and communities for TPD, and integrating ICT in authentic TPD. A paper developed from those discussions presented a set of illustrative cases and a conceptual model for linking research with practice (Albion, Tondeur, Forkosh-Baruch, & Peeraer, 2015).

The working group on TPD (TWG3) at EDUsummIT 2015 in Bangkok had the deliberations from previous meetings as background and was conscious of the challenges of ensuring that TPD for ICT integration in education meets the needs of teachers across a wide variety of contexts and cultures. Lack of suitable TPD may exacerbate the digital divide that already exists at multiple levels between and within countries and even within individual schools (Anderson, 2010). Hence, merely providing ICT does not inevitably improve learning, but beyond access, it is how teachers use ICT that makes a difference, and TPD is critical to achieving valued outcomes. This paper addresses the challenges to effective TPD identified during those discussions by presenting a set of cases that have highlighted the challenges in different contexts and abstracting lessons learned to suggest a model for effective TPD.

Challenges for TPD

ICT implementation in education systems should be accomplished systematically, following evidence-based diffusion models (Kozma & Vota, 2014). Technology utilization de facto, however, is far from being a means for systemic change: It rather facilitates “islands of innovation,” based on stellar cases carried out by excellent teachers, who practice innovative pedagogy using technology unrelated to TPD (Forkosh-Baruch, Nachmias, Mioduser, & Tubin, 2005). Extending and diffusing these individual excellent practices may occur by constructing professional learning communities within schools, in a model of networked communities of practice that facilitate sharing of
experiences (Twining et al., 2013). Currently, allocation of resources for the improvement of the quality of teacher professional knowledge is insufficient in scope and in quality (Leask & Younie, 2013). At the same time, ICT is expected to transform education, thereby promoting 21st century skills. Hence, the need for effective TPD is crucial, but we must ask what TPD would be most beneficial and how should it be most effectively delivered, so that the digital divide is overcome, and the gaps identified in usage and outcomes are addressed.

Determining teacher effectiveness, specifically when using ICT, is challenging. Ministries of Education worldwide expect ICT to be used to enhance the quality of education; nevertheless, there is insufficient ICT-utilization to assist teachers in gaining professional knowledge to improve their practice (Leask & Younie, 2013). Modeling technology utilization for professional development may serve as a powerful means of facilitating teachers’ use of technology innovatively in class (Ertmer et al., 2012). These are among the numerous challenges to TPD for ICT in education. In their discussions, the EDUsummIT thematic working group identified the five challenges described below as a focus for activity.

**Challenge 1: Contextualization with sociocultural awareness**

Technology enables us to create, collect, store and use information; to connect with people and resources all over the world; to collaborate in creating knowledge; and to distribute and benefit from knowledge products (OECD, 2015). However, many people lack access to technology, resulting in a new form of exclusion often described as the “digital divide.” Lack of access to the Internet is one of the most damaging forms of exclusion (Tondeur, Sinnaeve, van Houtte, & van Braak, 2011). However, there are many variations and levels of access. They suggested that those concerned should be thinking of a “gradation” instead of a divide between those who can use ICT to access, adapt and create knowledge and those who cannot (Warshauer, 2003; Gorski, 2005). Hence, globalization should not imply homogenization of culture. TPD in support of technology use in education should be both sensitive to, and enabling of, differences in historic, social, cultural, economic, and political contexts. Technology integration is also influenced by specific school cultures that require careful alignment of content and pedagogical knowledge. These differences should be seen as assets within TPD.

**Challenge 2: Sustainability and scalability of PD**

Providing continuing TPD about technology integration in education is challenging because of the large numbers of teachers to be reached and the need for frequent updates in response to continuing developments in ICT. Sustainability, meaning regular and long-lasting renewal and efficient use of available resources, and scalability, meaning capacity to reach all and disseminate ideas, are key characteristics for success (Albion et al., 2015; Voogt et al., 2015). Barriers for scaling and sustaining PD include social and cultural factors, lack of teachers’ technological, pedagogical, and content knowledge (TPACK) (Mishra & Koehler, 2006), inadequate infrastructure, limitations of Internet diffusion, linguistic differences, and geographical separation (Edirisinghe, 2015).

Working with teachers to develop their knowledge, beliefs, and attitudes can build a sustainable culture that supports ICT as integral to learning and teaching (Ertmer & Ottenbreit-Leftwich, 2010). Enabling teachers to share their ideas and provide examples of their good practices, facilitates understanding, sharing and negotiating, and consequently transference into local settings (Prestridge & Tondeur, 2015). Opportunities for TPD available through online communities of practice, social networking and online environments can provide both sustainable and scalable outcomes across geographical and cultural contexts.

**Challenge 3: Empowering pedagogy through ICT**

Supporting the effective application of technology to enhance learning and teaching in novel ways may serve as a foundation for successful TPD. Utilizing ICT in novel ways within TPD may facilitate innovative pedagogical practices that will, in turn, bring into practice innovative teachers who may affect the education system as a whole, thereby leveraging efforts in the field and establishing Professional Development 2.0 (Prestridge & Tondeur, 2015). Education evolves in parallel with innovative pedagogical practices using technology so that novel ICT-empowered pedagogies are emerging constantly. These include new approaches to content delivery and merging of content from
different disciplines, which may in turn create a new curriculum. In short, what is expected of the TPD process (effective application of ICT to enhance learning and teaching), is not simply to be a process of transformation and/or innovation but a process of social change in the transaction of pedagogy and content. It is assumed that “history is on the side of change” (Cobb, 2007, p. 14).

Challenge 4: Technology discernment

Educational decision makers—whether teachers, principals or policy-makers—have to make wise decisions about the selection and deployment of ICT and about the content and delivery of TPD to support applications of ICT in order to ensure that the outcomes are enhanced education for all. Critical analysis of available data may not necessarily lead to a clear decision and may need an additional perceptive judgment, “psychological or moral in nature,” (Albion et al., 2015a, pp. 25) called discernment.

Trauffer (2008) asserts that “discernment represents a multidimensional concept of decision making by logic and reason, by empathy gained through understanding, and by moral ethics” (p. 13) and “the ability to regulate one’s thinking in the acquisition and application of knowledge to make decisions that are right, fair, and just (p. 90).” This notion is more than simple critical thinking required of an educational leader when deciding about the technological inputs, and can be described as technology discernment. The power of discernment when applied to the choice of technology in the form of products, services or processes involved in the TPD, can ensure the real working needs of the TPD participants are met and thus keep them engaged and motivated in a sustainable manner.

Challenge 5: Systemic and systematic TPD

Professional development of teachers requires a lifelong learning approach, beginning with pre-service teacher education programs, and continuing throughout their professional lifespan. During the discussions among TWG3, several projects were showcased addressing the importance of systemic approaches to change. In this respect, effective preparation of pre-service teachers for technology integration requires attention to (1) all the stakeholders at different levels in the education system and (2) local factors (cultural and structural), but also demands similar attention toward the relationships between the themes (Kay, 2006; Tondeur et al., 2012).

At the same time several TWG3 members stressed the importance of systematic (gradual and evolving) change efforts. This aligns with the results of Seels et al. (2003) who concluded that it should take a long period with constant reiterations to see substantial change in technology integration (see also Albion et al., 2015a; Tondeur et al., 2016). Underpinning this conclusion is the understanding that teacher participation in the learning “process” and the development of learner autonomy (and self-regulation, especially online) are considered outcomes of professional development (Prestridge & Tondeur, 2015). Systematic PD also refers to the need for lifelong professional processes.

Four cases with different challenges and solutions

The four cases described below are drawn from contexts with varied geographic and economic characteristics, and refer to different emphases, e.g., target population (teachers or teacher educators), scope (regional, national), focus (ICT tools, educational initiatives) and diffusion patterns (face to face, online or mixed, synchronous or asynchronous). They may not be representative of all possible contexts; however they provide insights applicable to other contexts. Each case responds to the particularities of its context, and yet there are commonalities, both in the challenges to which the cases respond and the methods embodied in those responses. By considering the similarities and differences it may be possible to identify patterns that will point to an integrated framework to guide future work in teacher professional development for ICT in education.

Case 1: TPD for technology integration in Kenyan schools

The main objective of TPD was to enhance the capacity of Kenyan secondary schools to achieve change associated with effective application of ICT into the curriculum. This was primarily done through peer learning and sharing of
ideas and experiences between schools in combination with intra-school learning. In the development phase a team was appointed to set up an intervention designed by bringing together the Ministries’ experiences from previous ICT integration initiatives, lessons learnt from benchmarking, and literature relevant to ICT integration in education. In the second phase, school-level sensitization workshops were organized where representatives from the Ministry visited each school with the aim of demystifying ICT integration and getting teachers informed about the possibilities ICT would bring. To manage the process of ICT integration, schools were encouraged to form small teams comprised of teaching and non-teaching staff to oversee planning and implementation in school. Phase 3 started with a three-day workshop on development of ICT school policies. Each school delegated a representative and the interactive workshop was facilitated to encourage peer learning, inter and intra school learning. The teams returned to their schools to brief their colleagues and to prepare for more comprehensive capacity building workshops. Each school was expected to use the knowledge gained to develop an ICT-policy plan.

To facilitate the start of the implementation of their ICT school policy, each school was provided with 14,500 euros. Each school made informed choices on how they would use the funds to invest in ICT equipment. The first activity in the school-based training (Phase 4) was a five-day training on the development of ICT integrated lessons, held in the participating schools for all teachers. In the training sessions, teachers organized themselves in subject working groups for the duration of the training, to encourage peer learning. The trainers in the workshop would help the teachers brainstorm and then train them on the ICT skill(s) required to accomplish the task. Next, each team presented ICT integrated lessons they had developed to the plenum and received feedback. Finally, monitoring visits were conducted to establish the progress the schools were making and to inform a second training in schools to increase the understanding and making of ICT integrated lessons. The teachers once again worked in Teacher Design Teams in consultation with the trainers who followed them to class to observe how they managed the lesson with students. All teams would then convene for a feedback session and improvements were made. Findings of a longitudinal study (Tondeur, Krug, Bill, Smulders, & Chang, 2015) identified a range of challenges such as the importance of the context (theme 1: contextualization). This Kenyan case illustrates how schools are influenced by multiple historic, social, cultural, economic, and political contexts (Krug & Arntzen, 2010), e.g., school leadership faced daily challenges produced partially by the Kenyan Ministry’s new curriculum policy on ICT integration, but also because of the specific social, physical and cultural conditions of each school’s context (e.g., lack of electricity or power breakdowns; lack of time, the number of pupils).

The starting point of this project was to explore the context specific processes of ICT integration within secondary schools in Kenya, and to identify various conditions that influence the success and/or failure of technology integration in these schools. This perspective was informed by researchers who have argued for a more holistic approach (Challenge 4: systemic) (e.g., Krug & Arntzen, 2010). Clearly, the process of ICT-integration should not be facilitated as stand-alone events. Rather, PD should be part of a cycle of inquiry that supports teachers learning, try out and receive feedback. One of the main failures of many past programs, not only in Kenya, was that schools were provided with expensive equipment, but with little or no support for teachers’ PD (Challenge 2).

**Case 2: Social networking for self-generating TPD in Australian schools**

A TPD program focused on enhancing technology-enabled practices was implemented across geographically removed schools in Australia. The underlying premise was to enable teachers from any location in Australia an opportunity to form and contribute to a networked learning community. No face-to-face contact occurred but teachers could access and work within the online space at any point in their professional working day. An inquiry-based professional development model was adopted to shift the approach of professional development from content delivery to content generation, such that the teachers were collaborating, reflecting and analyzing self-practices, in conjunction with independent inquiries to build greater understandings about what good technology-enabled teaching looks like.

Four phases were aligned with the four school terms in Australia. Term 1 involved the professional development provider-mentor setting up the online space, welcoming teachers, initiating “getting to know you” activities and dealing with administrative requirements. The mentor’s was to support the teachers in making conscious their current beliefs about the use of technologies, their current practices and how these related to one another. Term 2 involved teachers in planning for an innovation in their teaching described as the teacher’s “Action Learning Project.” With the help of the mentor and the community of teachers, each teacher developed a plan for implementing a new
technology into their classroom. Term 3 became the implementation phase where each teacher enacted their plan and shared, monitored and reflected on it through the online community. Term 4 involved the redesign of their Action Learning Plans with clearer connection to their pedagogical beliefs.

Multiple communications and web 2.0 tools were used to support the PD program: online collaborative discussion space, individual teacher profile pages and a repository for shared resources – any technologies that enabled teacher engagement, e.g., email, text messaging, real time chat software, blogs and Twitter. The mentor played a key role in facilitating professional learning in response to individual teachers’ needs, e.g., support and nurture engagement; provide ideas, opportunities, and activities; and direct teachers to specific prospects but there was no requirement for teachers to engage. In this way the professional development was considered self-directed by the teachers, what could be called “demand-driven,” with free-flowing pathways rather than staged tasks that were sequential and controlled.

The approach illustrated in this case responds in some degree to each of the challenges identified by TWG3 during its discussions in Bangkok. Taking the TPD online addressed issues of geographical equity for teachers separated by the long distances in Australia and that recognition of diversity was reinforced in the structure of the program that enabled teachers to tailor their learning to meet their own professional needs in their local school context. Perhaps the greatest contribution of this case in relation to the identified challenges is in the area of sustainability and scalability. Using technologies to enable teachers working in varied locations and socio-economic environments provided opportunities for cross-boundary, cross-cultural national and potentially international collaborations. This model of PD can leverage opportunities for teachers from different countries to learn from one another. At the heart of the PD is a virtual learning community that offers a place for teachers to engage and learn professionally. Teachers sharing, challenging and co-creating a deeper understanding of the issues, barriers, enablers and, importantly, varied pedagogical beliefs and practices that inform the use of technologies in classrooms is the lifeblood of long-lasting pedagogical renewal. A networked learning community can also grow and evolve over time with the inculturation of new members and the movement of participants in and out. Considering that teacher professional development can be episodic, with teachers moving in and out, focused or on the periphery as their needs change and grow, the responsive opportunities provided by networked learning communities are eminently well suited.

At the same time, TPD that uses ICT in ways that model the pedagogical possibilities in their classrooms and expose teachers to thoughts and practices of diverse colleagues can, at the one time, challenge them and empower them to address the unique needs of their own context, thereby addressing the third challenge of empowering pedagogy through ICT. The role of the mentor is critical in this TPD model for assisting each participant to navigate a personal path to learning. In the best cases the guidance from the mentor will depend upon knowledge of the participant and available opportunities allied with the sensitivity to encourage engagement without alienation. This balance goes beyond mere technical knowledge to constitute the discernment identified in the fourth challenge. Finally, this case of TPD also addresses the systemic and systematic challenge through its attention to the local context and development of teachers’ capacity for autonomy and self-regulation of professional learning.

**Case 3: Professional development through teacher initiatives using ICT in Israeli schools**

Israeli teachers are requested to engage in PD in order to receive teaching degrees for becoming expert teachers. ICT is defined as a pillar theme for this advancement in their professional degree. To this day, ICT implementation was achieved mostly by ICT coordinators, which were assigned to promote mostly students’ ICT skills, but also to implement ICT in several subject areas learned within schools at across the elementary, lower-secondary and higher-secondary levels. PD focused on ICT coordinators’ sense of empowerment and viewpoints of the change elements that the National Computerization Program creates within the schools (Avidov-Ungar & Magen-Nagar, 2015). Since empowered educators assume high motivation for action and change, and may be driven by a need for constant lifelong learning. As a result, recommendations that were relevant to ICT coordinators were scaled to all educators, encouraging them to expand their personal knowledge on developing ICT-based instruction that uses digital tools in a wise decision-making process that promotes educational processes altogether (i.e., ICT discernment). This initiative was also aimed at ensuring sustainability of PD in the field of ICT-based practices in schools.
As part of a general reform of the Israeli Ministry of Education, lifelong learning paradigms were implemented in TPD in general, as it is part of the professional agenda for Israeli teachers, in which PD is considered an integral component of teachers’ identity and job requirements, allowing development of autonomy and self-regulation. Hence, PD reform according to this 2-year program is obtained via a process in which teachers initiate a project which is to be implemented in their school. The PD 2-year course is carried out by academic teacher education institutes (Avidov-Unger, 2013). The process is accompanied by mentoring the participants, from the design phase to the implementation phase. The mode of PD is mostly using the project-based learning method which allows teachers to better link content, pedagogy and technology (i.e., discernment). The project may be carried out with small groups of teachers, which allows collaboration between teachers within the same institute or between institutes. The face-to-face PD is accompanied by a Moodle course website, in which all projects are uploaded and open to other PD participants. This creates a community of practice, which in turn has an impact on improving the quality of each project.

ICT skills are practiced as part of the ongoing activities within the planning of the project itself. Teachers are required to explore the issue of focus in their project, and following the project design – to explore technological tools which may assist them in achieving their goal. For example, in a project for children with special needs, the project managers (i.e., the teachers) aimed to assist the students in their orientation and mobility outside the school; hence, they searched for applications to utilize for this project. Teachers were challenged via this mode of PD to empower their pedagogy using ICT and to explore novel pedagogy, and vice versa – to attain the required ICT skills and to study new digital environments suitable for their initiatives.

The projects are managed and documented using a blog throughout the professional development, with the goals being twofold: (a) to enable precise documentation of the process, not only of the final product which is handed in for assessment of the PD, and (b) to allow reflection of the learning process, in terms of ICT skills and competencies as well as disciplinary issues and implementation recommendations.

The end of the PD is celebrated by presentation of all projects. Outstanding projects are awarded prizes for excellence, and are disseminated widely throughout the district. Accessibility to all initiatives was enhanced by constructing a Google interactive map connected to data which points to locations and essence of each project, allowing participants to learn from successful cases country-wide, and to search for relevant initiatives according to subject matters, target populations and pedagogical principles. This map is located within a college portal and is accessible also to the general public, including educators, policymakers and other interested parties. All projects require a research component, to be administered when implementing the initiatives, in order to ensure conclusions and recommendation, thereby creating a body of knowledge that is useful in turn for further systemic and systematic PD. Research includes a research question which is stated within the project proposal, as well as benchmarks for assessment of the success of the project, in terms of the pedagogical impact, teachers’ and students’ roles, the novelty of the initiative, as well as the modes of assimilation and dissemination.

To summarize, engaging teachers in projects that have immediate value in their workplaces as well as engendering learning for the future supports sustainability. This maximizes the value obtained from teachers’ initiatives, making their work scalable by sharing these projects with the wider educational community. Teachers are enabled and encouraged to explore different approaches to applying ICT both in their own projects and through being able to see the results of work by others. Deciding which ideas to adopt, either self-generated or inspired by the work of others, develops technological discernment required to select the most suitable matches between content, pedagogy and technology.

Case 4: Professional knowledge sharing in a Sri Lankan pre-service internship

In Sri Lanka, the focus on ICT in teacher education has shifted from preparation for teaching specific courses about hardware and software toward promoting the use of ICT as a tool for learning and teaching across the school system (Edirisinghe, 2007). That change has been challenging because, although Sri Lanka is geographically compact, there are ten thousand schools in the state sector, with almost 250,000 teachers and over 4 million children. Balancing provision of ICT equipment to schools and professional development of teachers presented logistical challenges; without TPD, equipment would not be used effectively, but without equipment teachers could not apply what they learned in TPD.
Commencing in 2004, a needs analysis was undertaken in 122 schools and a TPD was developed to assist teachers in adopting ICT to create new learning environments (Edirisinghe, 2007). Recognizing the complexity of changing teachers’ practice, the research team developed the TPD around a KASP model, addressing teachers’ Knowledge (K) about ICT and pedagogy, Attitudes (A) toward the application of ICT in pedagogy, and practical Skills (S) for working with ICT as a foundation for reshaping Practice (P) and life-long learning for professional growth.

Although the TPD had promoted constructivist approaches to learning and teaching, instructivist practices were deeply rooted in societal traditions and the most common uses of ICT were for digital presentations of lesson content rather than as a stimulus to learner activity and engagement. In cases where equipment issues prevented digital presentation, printed copies of the presentation material were distributed. Such uses were supported by authorities through awards for good digital production. Although there was some sharing of digital presentation resources by teachers, there was little movement toward the professional knowledge sharing (PKS) envisaged by the developers of the TPD. Hence a new approach was needed to achieve educational change supported by ICT.

In 2012 the Sri Lankan government commenced a new national program of transforming schools to ensure the achievement of knowledge-based development. The program was designed to provide schools with ICT for use across the curriculum, beginning with 405 Phase 1 schools. As described above, TPD for pedagogy with ICT had not been successful in promoting the desired changes in pedagogy and PKS. Hence a new approach was planned for trial in a pilot program with provision for a research team to validate the model and inform improvements ahead of wider adoption. The study has adopted a knowledge management model based on socialization, externalization, combination, and internalization (SECI) in which the four modes of knowledge conversion form a spiral (Nonaka & Toyama, 2003). The pilot group comprised 20 final-year students preparing to be English teachers working with mentor teachers during an intensive “internship” in schools. The initial stage of socialization enables interns and mentors to build relationships through tacit-tacit exchange of knowledge. In the second stage, the mentor externalizes (tacit-explicit) knowledge of pedagogy based on experience enabling combination (explicit-explicit) with the knowledge brought by the intern as they plan together. Finally, both are able to internalize (explicit-tacit) knowledge shared and developed. In this way the contextual and pedagogical knowledge of the mentor is blended with the new knowledge the intern brings about ICT and with new approaches to teaching. Early indications are that the approach is effective and it will be developed for wider implementation among the 3000 teachers who graduate each year.

This case addresses each of the challenges identified above to varying degrees. By facilitating collaboration and knowledge exchange between interns and mentors, the model respects both and supports the melding of the intern’s new knowledge about ICT in pedagogy with the mentor’s experiential knowledge of teaching in the local context. If the approach is successful and can be replicated with all graduating teachers, it offers a sustainable and scalable method of disseminating new approaches as they become available while remaining sensitive to the accumulated knowledge among teachers in schools. By promoting the intergenerational exchange of knowledge the project offers opportunities to blend the best of old and new to generate novel pedagogical approaches while exercising technology discernment to make choices that are attuned to local contexts. The approach is also systemic and systematic, having the potential over time to reach all corners of the education system and, by engaging interns with mentors in knowledge exchange, establish the importance of lifelong professional learning for all.

Discussion

Based on comparisons across cases, it is possible to identify some common characteristics that appear to contribute to their success in achieving enhanced application of ICT in education through coupling provision of resources with effective TPD. Teacher inquiry leading to development of projects that are sensitive to local context can be effective for promoting teachers’ professional learning at the same time as enhancing learning in their classrooms. Collegial sharing of those projects with professional networks and communities propagates successful innovation to other teachers and informs future adaptation by the originators of those projects. Such sharing needs to value the contributions of all participants by making space for each to both give and receive, even if that is sometimes the only comment that may contribute to future developments. Respecting the professionalism of teachers by allowing them agency to respond to local needs and valuing their contributions to the profession motivates them to engage more fully. Systemic inputs including hardware, software and training can be directed into the system by way of the teacher projects with increased likelihood that the resources will be applied to enhancing learning in classrooms.
Engaging teachers in learning through inquiry, especially using ICT, is a powerful way to move them from a traditional view of education as transmission of knowledge from the expert teacher to the learner and toward understanding of education as a lifelong process of learning with and from each other.

Each of the cases presented above attended to the specific conditions in its schools or other locations by engaging participants in development of teaching plans and resources tailored to their own needs (cf. Kozma & Vota, 2014). Such approaches are respectful of the participating teachers as professionals and give them agency in the process rather than delivering solutions from external experts that devalue the insights that teachers have developed through experience. In the case of Sri Lanka, the connection between pre-service interns and experienced mentors recognizes that teachers at all stages of their careers can contribute something of value but are also able to benefit from professional learning.

Although it is implemented in different ways across the cases, encouraging sharing among teachers is a powerful strategy for addressing the challenge of making TPD sustainable and scalable. The focus in the Kenyan schools on developing a shared vision for ICT on which to build new practice together (see also Tondeur, van Keer, van Braak, & Valcke, 2008), the adoption of social networking in the Australian case to facilitate sharing across distance, the wide sharing of individual projects in the Israeli case, and the involvement of teachers at varying career stages in Sri Lanka case all engage teachers in exchanging knowledge and experience. By creating environments and processes to facilitate that exchange and validating it through official support and recognition, these cases are developing the conditions to enable sustainable and scalable change through networks and communities (Prestridge & Tondeur, 2015). Both the creation of the means for sharing and validation by school and system authorities are important for encouraging teachers to engage. This creates a mixed bottom-up and top-down diffusion process, in a twofold manner: first, in creating a climate of sharing ideas and initiatives, thereby allowing the scaling of ideas and initiatives, and secondly, in exercising ICT-skills and competencies, thereby practicing 21st century literacies which are crucial in facilitating meaningful learning processes.

Engaging teachers in a form of inquiry learning is another common characteristic across these cases. The nature of the projects varies but typically teachers are engaged in inquiry leading to planning and developing resources for teaching with ICT and subsequently sharing what they produce with colleagues (cf. Agyei & Voogt, 2012). Project-based learning, problem-based learning, and challenge based learning are among the variety of learner-centered constructivist pedagogies based on inquiry (Larmer, 2014) and which are widely recommended for adoption for teaching with ICT. When teachers participate in inquiry learning around ICT they experience firsthand the style of...
pedagogy that they are being encouraged to adopt in their classes. In this way TPD can assist teachers to adopt pedagogical forms that are empowered by ICT to engage learners in authentic activity rather than replicating more traditional instructivist practices.

Education is inevitably about more than technical processes, invoking questions of value about what should be learned and how. For change in education to take root and persist, teachers must be persuaded that it aligns with values and will be beneficial for learners (Sang et al., 2010). By granting teachers agency in the change process, each of these cases has gone beyond simply disseminating technical knowledge and skills to exercising the wisdom of discernment that enables negotiation of change that responds to the exigencies of each context. The message from these cases emphasizes the importance of those charged with promoting change exercising sensitivity to the context of change and the people within it.

The challenge of making PD systemic and systematic has also been addressed by these cases. Engaging teachers in developing and sharing workplace projects ensures systemic effect through consideration of local conditions and initiation of ripple effects that intersect with teachers at all career stages (Lim et al., 2014). The inquiry learning approach allows teachers agency in determining the specifics of their professional learning and opportunity for collegial feedback that informs future iterations of their projects. In this way the professional learning meets the challenge of being systematic by evolving gradually as part of a lifelong professional process.

**Conclusion**

The model presented in Figure 1 has been developed through reflection on how the diverse cases presented above responded to the challenges for TPD identified by thematic working group 3 at EDUsummIT 2015. The discussions of TWG3 were informed by the deliberations of previous EDUsummIT meetings (Twining et al., 2013; Albion et al., 2015b), continued immersion in the literature, and the experiences of members drawn from a variety of contexts around the globe. Each of the cases responded in its own way to the challenges identified by TWG3 in developing TPD for ICT-integration. Despite the wide variations in details of the TPD it was possible to discern common threads that are embodied in the representation of Figure 1.

As noted above, it is not the mere presence of, or access to, ICT-resources that will transform education. Transformation will be effected through what teachers do with the technology that is available to them. Because technology and the local context are changing it is necessary for teachers to engage in an ongoing process of inquiry through which they appropriate the affordances of technology to develop creative responses to local needs. The cyclical process of teacher inquiry is central to the model and is informed by systemic inputs, contextual factors including the needs and responses of learners, and the wider professional networks and communities within which teachers share their experiences and from which they draw inspiration for their own continuing learning. Teacher agency is a critical element for motivation through the sense that their contribution is valued and their professionalism trusted. The cycle of inquiry at the heart of this model is consistent with the recommendation for design-based research that arose from EDUsummIT 2103 (Albion et al., 2015b). Teacher-researchers adopting inquiry learning and working close to the context of learning are well placed to discern what works well for learning and to adopt or adapt practices as appropriate to their context. There is no claim that this model is the solution to all the challenges of TPD for ICT in education but it does encapsulate principals from successful cases and points the way to future developments.

**References**


119


Educational Digital Technologies in Developing Countries Challenge Third Party Providers

Don Passey1*, Thérèse Laferrière2, Manal Yazbak-Abu Ahmad3, Miron Bhowmik4, Diana Gross5, Janet Price6, Paul Resta7 and Miri Shonfeld8

1Lancaster University, Lancaster, UK // 2Université Laval, Quebec City, Canada // 3Sakhnin College, Sakhnin, Israel // 4UNESCO Bangkok Office, Bangkok, Thailand // 5Global Citizen Educate, Myanmar // 6University of Tasmania, Hobart, Australia // 7University of Texas at Austin, Austin, USA // 8Kibbutzim College of Education and MOFET, Ramat Aviv, Israel // d.passey@lancaster.ac.uk // Therese.Laferriere@fse.ulaval.ca // manalyazbak@gmail.com // m.k.bhowmik@unesco.org // Diana@GlobalCitizenEducate.org // Janet.Price@utas.edu.au // resta@austin.utexas.edu // mirish@macam.ac.il

*Corresponding author

ABSTRACT

In this conceptual paper, we consider issues and challenges of third party and governmental organisations in planning and implementing access to and uses of digital technologies for learning and teaching in developing countries. We consider failures and weaknesses in the planning and implementation processes highlighted by research in developed countries (as well as successes supporting implementation). We problematise these issues and challenges, conceptualise them in order to focus on longer-term rather than shorter-term ones, and offer new alternative models and ways of conceiving these practices for future sustainability.

Keywords

Educational equity, Digital sustainability, Long-term implementation, Developing countries, Gaps and challenges

Introduction

This paper focuses on educational change and provision of opportunity, on changes involving digital technologies or information and communication technologies (ICT). While it is recognised that introducing ICT can affect teaching and learning, it is more specific implications for third party and governmental organisations (typically philanthropic foundations, consultancy, hardware or software company groups, or agencies with specific education and ICT remits) which we focus on here. In this paper, digital technologies or ICT are defined as the entirety of hardware and software systems accessed by learners or teachers, which might include desktop, laptop or handheld technologies, mobile technologies, networking tools, and online resources and software. Regarding change processes, terms used in this paper include innovation (a method, idea or practice developed by and new to a user), implementation (a process putting a plan into effect), and integration (combining something new into an existing system or practice). Some national education systems (the United Kingdom, Denmark, the United States of America, and Australia, for example) have implemented ICT into teaching and learning over a period of 25 or more years (see Tatnall & Davey, 2014). In spite of such a long period for integration of ICT into educational practices, concerns continue to be raised; ICT has neither been accepted for teaching or learning on a wide scale (see OECD, 2015, for example), nor brought about the range of benefits expected from the investment (e.g., Selwyn, 2010). However, research has shown that ICT can bring about educational benefits (e.g., Passey, 2014; Tamim et al., 2011). A key difference between these apparently contrasting findings, identified by the latter authors, concerns approaches and roles of teachers, tutors, counsellors, policy makers, or parents supporting educational practices. ICT does not necessarily bring about change or benefit without appropriate support. In developed countries, even where ICT has become part of teaching and learning environments, sustainability can be identified as an issue, nevertheless. An absence of sustained classroom-based innovation with digital technology in both developed and developing countries is reported (Attewell, 2001). Digital tools and resources can be underused, given the pressures of curricular demands in developed countries (Cuban, 2015). For developing countries, these are important messages to consider; but contexts in which implementations might be undertaken may not be identical, so sustainability factors might not be the same.

In many developing countries, national policy makers and third party organisations currently focus on a number of desires: To introduce ICT to enhance teaching and learning; to promote educational opportunity (or equity); to learn from past experience so that implementation might be more effectively handled; and to generate capacity building in the use of ICT. To succeed, ICT-related educational programmes should be designed, adopted and implemented by
government and third party organisations to accommodate a number of recognised issues (detailed further later in the paper). Importantly, technology continues to change rapidly and is often repurposed, and time is needed to implement and recognise agreed outcome benefits (what we refer to here, and will detail later, as “the U-challenge”). Additionally, there are differences and complexities within the contexts in different countries (political, social, technological, linguistic, cultural, economic, local and religious). All three of these issues have significant implications for teaching and learning. If long-term integration is to be achieved, these issues must be considered appropriately. Long-term integration in this sense can be thought of as providing sustainability; however, even this concept needs to be considered appropriately within implementation contexts. Initially in this paper, we provide a background to this field; we then conceptualise issues and features in order to develop a starting framework for those working in the field, and then identify possible future approaches involving longer-term research and development agendas.

A key question we explore in this paper is, how can educational change involving digital technologies be problematised and conceived for governmental and third party organisations seeking to implement long-term sustainable practices. Problematisation of change issues is an essential process if third party organisations and others involved in implementation are to avoid key problems arising, by putting appropriate processes in place ahead of identifiable critical periods. From the conceptions we present, we offer new alternative models and ways of conceiving practices for future sustainability. We will consider the context, identify key issues and challenges, problematise underlying factors, derive an appropriate model of implementation, provide a planning framework, offer recommendations, and discuss implications for research.

Digital technology implementation and the context in developing countries

Integration of digital technology into teaching and learning is a double-edged challenge. While online distance education increases access without borders to a variety of subject and topic contents (Bakia, Shear, Toyama, & Lasseter, 2012), onsite formal education is facing rising expectations regarding the practices and nature of methodology (Griffin et al., 2012; UNESCO, 2011; Voogt & Knezek, 2008). Previous papers from Thematic Working Group 4 (TWG 4) of the EDUsummIT community (Resta, 2009; Resta, 2011; Resta & Laferrière, 2013), portrayed: the state of infusion of ICT in the world; aspects of digital equity that researchers have pointed to; initiatives taken; and persisting issues and challenges.

Figure 1. ICT revolution and remaining gaps (ITU, 2015)
At the latest EDUsummIT conference 2015, TWG4 focused on onsite sustainable implementation with digital technology, primarily in classrooms. Absence of sustainable innovation and implementation with digital technology was identified as a “new situation” requiring noteworthy attention. This critical gap needs to be overcome before substantive progress can be made in supporting educational equity concerns through digital technology approaches.

While digital equity inside classrooms during mandatory schooling years is a key concern of third party initiatives, digital equity outside the classroom is also a factor to consider. As shown in Figure 1, the United Nations specialised agency for ICT (the International Telecommunications Union) indicated in a recent report (ITU, 2015) the impressive global progress in penetration of Internet-based information and communication; but the penetration rate is only 9.5% in the least developed countries.

In terms of the efficacy of access for educational purposes, research indicates it is essential to access broadband to derive full benefits from the Internet. Figure 2 shows population proportions accessing the Internet through landline and mobile telephones. As noted in the World Economic Forum (Dutta et al., 2015), the widening divide in broadband access between the most developed countries and the least developed countries is a discouraging trend.

Seeking to address this digital divide for education, the World Summit on the Information Society (WSIS) set the target of 2015 for connecting all secondary and primary schools with ICTs (ITU, 2014). This target is an ideal but mammoth undertaking:

Evidence shows that LCRs [“learner-to-computer connected” ratios] are generally decreasing across many countries, while school Internet rates are increasing – both generally and for fixed broadband specifically. However, change is not uniform and occurs at different rates in different countries. Typically, countries that have strong policies and set targets for ICT in education with high-level government and sector-wide support show the most rapid change. (ITU, 2014, p. 75)

Where schools use technology, research (e.g., Becker & Riel, 2000; Tamim et al., 2011) continues to find that the pedagogy in use makes the difference: technology used as “support for cognition” has greater effect than technology used for “presentation of content.” Bringing the Internet to schools and classrooms – whatever money, time and energy it may require – is only part of the equation. As Kozma (2005) stated, reviewing impacts of ICT in schools in developing countries: “changes in curriculum, pedagogy, assessment, and teacher training is likely to result in widespread use and learning” (p. 18). Technology must do more than reinforce the “teacher effect” (an expression used in quantitative studies measuring variables affecting student outcomes). It is here defined as “the teacher’s unique contribution to student learning” (Kupermintz et al., 2001).
Issues and challenges

We need to initially consider how forms of implementation (such as pilots, projects, sequential adoption models, or whole-school developments) are being approached, how successful they are, identifying weaknesses and gaps (e.g., Wagner et al., 2005). With little or no concern for sustainability from day one, initiatives can promise results that will not be achieved (Gichoya, Hepworth, & Dawson, 2006). In developing countries where digital access is still rare, once a project is over, use of technology, if any, becomes more difficult: hardware and software become obsolete; connectivity can be too expensive; technical support and professional development are lacking (Trucano, 2015). So, more often than not, capacity building comes to a stop, and scalability does not occur (Breuleux et al., 2002 Looi & Teh, 2015).

Many innovative and implementation practices do take place, both inside and outside classrooms (Steyn et al., 2011; Voogt et al., 2015; see also the French website Adjectif, at http://www.adjectif.net). EDUsummIT 2015 TWG4 reviewed eleven examples of such case practices from five continents, using ISTE’s (2009) essential conditions to leverage technology effectively for learning. These essential conditions are: (1) shared vision; (2) empowered leaders; (3) implementation planning; (4) consistent and adequate funding; (5) equitable access; (6) skilled personnel; (7) ongoing professional development; (8) technical support; (9) curriculum framework; (10) student-centred learning; (11) assessment and evaluation; (12) engaged communities; (13) support policies; and (14) supportive external context. Case writers estimated all these conditions as being more than half-present, with levels of presence estimated from 62% to 85%. These conditions appear to be important, but how they fit into longer-term processes for sustainability, involving third party and governmental organisations, is an important concern, which we focus on in subsequent sections of this paper.

Potential users of an innovation must accurately determine how to adapt and implement an innovation to their setting and whether sufficient conditions to make adaptation successful exist (Looi & Teh, 2015, p. ix). But, short-term studies demonstrate initial successes in adoption of digital technology are not enough (Ahmad, 2015; Labonté-Hubert, 2013). Long-term implementation studies reveal that conditions change, and new challenges arise (Laferrière, Hamel, & Searson, 2013; Passey, 2011; Sandholtz et al., 1997). Therefore, innovation and implementation need to be monitored regularly, in ways that match user and technology adaptation over time. For sustainability and digital equity to be achieved, long-term adaptability from technological, pedagogical, cultural, social and learning perspectives all need to be considered and in place.

Third party providers and policy makers should give special attention to curriculum frameworks outlining or enabling uses of ICT, and their alignment with implementation (e.g., ICT uses supporting student-centred learning), assessment and evaluation. In spite of abundant access to technology in developed countries, curricula and testing measures are both perceived by teachers as limiting uses of ICT (Fu, 2013). Developing curricula and testing measures, so that ICT use in teaching-learning processes becomes increasingly perceived as necessary, is a long-term challenge facing all education systems.

Regarding digital content to support country curricula, educational software and applications (apps) are widely developed in western countries, the content often reflecting those cultures. As a result, users in non-western countries must often learn and understand western culture to gain educational value from the resources. Developing resources by, or in conjunction with, individuals from regions in which they will be used, could remove one barrier to achieving educational equity. Additionally, this practice would potentially increase employment within this industry in regions where software/apps would be implemented. Similarly, language can act as a technological barrier. Most resources are in English, with non-English websites usually appealing to stereotypes to achieve marketability (Kalyanpur & Kirmani, 2005). Thus, it is important to build local communities of developers who can take forward continuity of resource innovations.

Factors affecting long-term implementation and sustainability

When issues and challenges influencing unsuccessful implementation are known, these need to be problematised. Digital gaps evolve because of gaps in access, adaptability, literacy and concerns held by some communities (see, for example, Hilbert, 2014). Fundamentally, introducing an element such as ICT into practice, to support both teaching and learning, is effectively concerned with change and its management. Conceptualising that change is clearly
Patterns of change involving ICT need to consider fundamental elements or factors. These factors may concern the ICT itself. Major hardware changes occur about every 5 years (Passey, 1999). Most recently, a major change occurred around 1995 when Internet use initially increased, and other major hardware changes since then have included interactive whiteboard technologies from the year 2000, mobile technologies from 2005, robots from 2010, and 3-dimensional, video and peripheral equipment such as printers from 2015. Software changes occur about every 18 months (discussed also in Passey, 1999). These changes not only include software updates and upgrades, but the emergence of new software. More recent significant software changes have included wider access to video editing, simulation and virtual world software, and video game editing and creation software.

Other factors affecting change concern how teachers support educational practices. Initial implementation of ICT into teaching practice leads to a downturn in performance. Mevarech (1997) described this as a U-curve, and she considered the implications that this has for teaching, learning and the appropriate timing for the identification of learning benefits. Leung, Watters and Ginns (2005) studied teachers in a case study school, with younger teachers reporting perceptions of their ICT abilities and self-efficacy decreasing during the first year of a project, and lowering of uses of ICT during the second year. These study findings paralleled those of a study conducted by the lead author, exploring how mathematics teachers in 20 schools integrated ICT into their practices; the teachers reported challenges during the first year of the project, and their uses of ICT decreased during the second year when professional development support was reduced. Benefits accruing after a period of time, once the downturn has been overcome, will not arise if implementation is not maintained to that second year. However, identifying forms of benefit needs to match forms of technologies and their uses. As stated elsewhere (Higgins et al., 2012; Passey, 2014), there is a need to consider what a technology does (in terms of its affordances, uses and outcomes for teaching and learning) before trying to identify its benefits and impacts. Regular changes in technologies and their appropriation for teaching and learning require regular updating and professional development. The “U-challenge” in this context refers to those time periods that teachers are implementing uses of technologies when their performance decreases, due initially to the need to accommodate new practices (Mevarech, 1997), and finding the most appropriate ways to benefit from these practices, then later, having to grapple with technologies that become increasingly obsolete or incompatible as time goes on.

Yet other factors affecting change concern the wider range of stakeholders who both influence educational practice and are influenced by educational practice. Change occurs within a system, and all actors in that system need to be involved and to understand what is happening and its consequences. Parents, students, teachers, school managers, third party providers, regional and national policy makers, educational advisers and researchers, and politicians are all stakeholders in the system; their values and concerns need to be known if change is to be managed successfully. These are factors concerned with national and cultural context. Context affects change, and contextual factors can hinder or support that change. Hence, political, social, technological, linguistic, cultural, economic, local and religious factors can affect change by being what have been described in some research studies as “drivers” or “barriers.”

Long-term change requires a concern for sustainability. However, sustainability in this sense should not be considered as stability, or lack of change; it should be considered as a way to manage and handle ongoing and successive change. Long-term sustainability requires adaptability on the part of the actors involved, including third party providers. If technologies alone are the main agents that change over time, then even so, users require periods
of adaptation to those changes, to become familiar with differences, with additional benefits or disadvantages, or with practices that enable effective outcomes.

These factors mean that third party providers implementing ICT for educational purposes need to consider the future very carefully. Such “next-generation alignment” (aligning what is provided in a current context to also fit and match future needs) requires third party providers to take on board both desired future practices and estimated future changes in education and ICT. For regional and national developments, this suggests that governments should generate partnerships with universities, philanthropic foundations, teacher and parent associations, and businesses, in order to draw on their expertise and desires in this regard. In taking forward implementation of ICT informed by such expertise and desires, successive cycles of data-driven discussion and decision-making will then be necessary.

Third party organisations (e.g., foundations) that want to make a contribution (in terms of resources, funding, time, or energy) toward educational digital equity processes should be encouraged to work with policy guidelines that target sustainable innovation in settings of their choosing. They will need to manage their own expectations, as well as the expectations of those they want to support. They will have to choose between long-term commitments that in some settings keep improving the level of presence of the essential initial conditions identified by ISTE and effectively leveraging technology for shorter-term learning (termed “spikes” of innovation by Florida, 2005). Spikes of innovation refer to technology use for learning concentrated in some areas where teachers, administrators and learners come together, grow an understanding and develop skill regarding its uses for learning. It is important for third party organisations to understand the relevance of such spikes of innovation. Otherwise, short-term actions are likely to first seduce and later disappoint teachers and learners. In this respect, managing the “U-challenge” (periods when teachers are implementing uses of technologies, when their performance initially decreases due to the need to accommodate new practices and later when upgrading and review of practices occur) is vitally important.

Models of implementation and integration

How third party providers can model integration of ICT into educational practice is important, taking on board both the issues and challenges, and the factors that influence change. We consider initially the usefulness and applicability of existing and new models of implementation and integration. Traditionally, models of technology integration have focused on an implementation through stages or phases (such as the model developed by Hooper & Rieber, 1995), even when considering pedagogical implementation or change (see, for example, Puentedura, 2013). Conceptual approaches that focus more on the context of the change have also been developed (such as that of Corbett & Rossman, 1989), and uses of this form of approach have identified important factors beyond the technologies themselves (including political and cultural factors). Indeed, taking this point further, some authors (for example, Oliver, 2011; Pannabecker, 1991) have argued that implementation models need to move away from concerns with technological determinism. Other models have focused more on the factors that influence individual take-up of technology uses, such as the widely considered Rogers’s (2003) Diffusion of Innovations Model, and the Technology Acceptance Model (TAM) of Davis (1989). While these models consider the individual user and their initial acceptance, other models have taken a perspective that is more concerned with ongoing adoption, such as the Concerns-Based Adoption Model (CBAM) of Hall, Wallace and Dossett (1973) and the post-acceptance model of Bhattacherjee (2001).

In contrast to models based on the elicitation of factors or features, Todnem By (2005) discusses the crucial importance of conceptualising change when considering its management. Factors identified in the previous section in this paper (technological, teacher practice, stakeholder, and contextual factors affecting long-term integration and sustainability) suggest that approaches for educational change with ICT should adopt management of continuous and incremental change with ongoing changing ICT. These factors support a visualisation of the conception of change as one having periods of downturn followed by benefit, then short periods of time where performance again dips (explained further below) prior to successive cycles arising. Figure 3 shows this as a long-term process, flowing across a period of perhaps 10 years.

In this model, downturns occur regularly, likely to arise because of technological shifts (referred to earlier as happening perhaps every 5 years for hardware and 18 months for software). After a time, different technologies and resources, redundancy of technologies, and difficulties associated with backward and forward compatibility and processing capacity, all mitigate a reduction in the ability of the teacher or user with the facilities afforded. Whilst
those facilities would be entirely usable at the outset, so there is a need to update, upgrade or replace, if the user is to maintain their performance. At the same time, this updating, upgrading or replacement needs to be considered in terms of necessary associated professional development and the way that abilities to use at the time of change allow the user to focus on adaptability. In Figure 3, base performance does not increase over time. The reason for showing performance in this way is to indicate that individuals need to accommodate change and retain their base performance. In reality, performance might change and increase over time, but this is an aspect where we need more research, to establish what happens at individual and group levels.

Figure 3. Successive cycles of ICT-determined change

Planning for sustainability

With a suitable model identified, plans emerging from this model can be detailed. If long-term integration of ICT is an aim of third party and government organisations, then a key concern needs to revolve around the planning elements of the process (as suggested by Figure 3 above), including a timeline integrating the elements identified and considered above. For example, considering a 10-year plan (shown in Table 1) that also integrates ISTE’s (2009) essential conditions to effectively leverage technology for learning:

- implementation starts in year 0
- major hardware changes are shown every 5 years, even though the exact nature of them is unlikely to be known in advance
- software changes are shown every 3 years, even though some will happen every 18 months, and again their nature is not likely to be known in advance
- performance downturn is assumed to occur over a one-year period, every five years when major hardware changes occur
- benefits are assumed to occur after the one-year downturn has been overcome
- regular updating is determined by hardware changes, software changes, and challenges arising during periods following downturns being overcome
- systemic actors need to be identified at the outset, and there needs to be a regular check for changes that might occur and the implications these have in terms of those involved
- contextual factors need to be identified at the same intervals, to ensure that barriers and drivers are adequately considered
- adaptability concerns need to be integrated with initial and regular updating

When considering the ways in which features for planning (identified in Table 1) need to be integrated to enable sustainability as far as is possible, the feature of adaptability (a key factor identified by Rogers (2014), as more critical than intrinsic motivation itself), clearly needs to take account of previous experience in building new knowledge concerned with content, pedagogical and technological features for teaching and learning (Mishra & Koehler, 2006). In this case, the visualisation of the conception of change being considered here should perhaps be concerned more with a looping cyclical form of development (identified initially from case studies described in Passey, Capstick, & Poole, 1997), and shown in Figure 4, across years 0 to 11 from Table 1. Again, no increase in baseline performance is shown; the same reasoning as that in Figure 3 applies here.
Table 1. Planning for sustainability

<table>
<thead>
<tr>
<th>Factor</th>
<th>ISTE essential conditions to review</th>
<th>Year 0</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5</th>
<th>Year 6</th>
<th>Year 7</th>
<th>Year 8</th>
<th>Year 9</th>
<th>Year 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting phase</td>
<td>(1); (2); (3); (4); (5); (6); (7); (8); (9); (10); (11); (12); (13); (14)</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>Hardware changes</td>
<td>(4); (5); (6); (7); (8)</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>Software changes</td>
<td>(4); (5); (6); (7); (8)</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>Performance downturn</td>
<td>(6); (7)</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>Benefits arising</td>
<td>(11)</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>Regular updating needed</td>
<td>(6); (7); (8); (9); (10); (11)</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>Systemic actors involved</td>
<td>(1); (2); (12); (13); (14)</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>Contextual factors identified</td>
<td>(1); (2); (12); (13); (14)</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>Adaptability concerns supported</td>
<td>(6); (7)</td>
<td></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. (1) shared vision; (2) empowered leaders; (3) implementation planning; (4) consistent and adequate funding; (5) equitable access; (6) skilled personnel; (7) ongoing professional development; (8) technical support; (9) curriculum framework; (10) student-centred learning; (11) assessment and evaluation; (12) engaged communities; (13) support policies; (14) supportive external context.

Figure 4. Looping cyclical form of ICT-determined change

Recommendations to third party and governmental organisations

If plans to implement ICT for education are to be successful, governmental and third party organisations need to be aware of key issues. From this conceptualisation and visualisation of implementation, recommendations at a policy level are:

- Be aware that change is inevitable, and that sustainability has to embed adaptability (Rogers, 2014). Examine innovation on a sustainable path towards digital equity by considering long-term adaptability, as well as referring to the ISTE initial “essential conditions” for project conception, implementation, and evaluation.

- Design projects inclusive of adequate time to build a reflective process that anticipates the dynamics of the “U-challenge” (time periods when teachers are implementing uses of technologies when their performance decreases, often during the first year, due initially to the need to accommodate new practices). Establish and nurture “spikes” of innovation perhaps every two years (see Table 1). For innovation to be sustained, innovation must be adapted to local contexts (which may need to be considered at institutional, local, regional or national
levels), reviewed every three years. The initial implementation phase must be followed by a long-term monitoring phase that identifies key points every two years. The formation of implementation and monitoring practices, including committees, is essential.

- Ensure understanding of what it is within a context that can gain systemic commitments in various contexts. It is important to be aware that digital educational resources reflect the culture of the country/region in which they are conceptualised and produced, and thus may inadvertently create barriers for reaching global educational equity.

- Commit resources and partners to long-term professional development of educators. Building a local community of developers, that will continue the innovations when the support is gone, should be established from the outset. As a part of this concern, materials in local languages should be available. Build in systemic and synchronous top-down (such as national and regional policies and support) and bottom-up (such as teacher and parent interest and commitment) processes that will assure sustainability.

Implications for research

Research has a major role to play in supporting sustainable digital technology implementation. Long-term planning for educational integration of ICT using a form of cyclical visualisation such as that shown in Figure 4 is not fully researched. Many studies have looked at the uses and outcomes of ICT in the short-term, perhaps over a period of 1 year, or 2 years at the most (see Passey, 2014, for example). Studies that look to explore the planning over long periods need to be instigated, and monitoring needs to be introduced over a period of years. Design-based implementation research is especially suited for such implementation and monitoring (Penuel et al., 2011). Initially, such studies and implementations can look at the experiences of the past, and consider how these might be projected to the future. For example, from the planning shown in Table 1, it would be possible to consider undertaking a study in a country to explore and identify more exactly the nature and implications for users of:

- hardware changes in education since 1995
- software changes since 1995
- experiences of learners, teachers, trainers or employers in their performance following ICT introductions
- benefits they have experienced from particular uses
- how these actors’ uses have been updated
- who the systemic actors have been within their individual contexts and the roles they have played
- the contextual factors that have been present, and which favour or hinder changes with ICT
- how sustainability has been and is being considered
- how adaptability has been and is being introduced

This form of study could be undertaken initially through a Delphi or phenomenological approach, interviewing teachers, learners, parents, school managers, third party organisation managers, national policy makers, politicians, advisers and researchers, and employers. To achieve longer-term concepts of these past changes and experiences, perspectives could be gathered across the compulsory, post-compulsory, training and employment arena.

Taking forward this initial form of research, short-term studies can then be conducted and findings can be integrated into the context of the wider and longer-term picture that a longer-term study provides. It is not just implementation and planning that needs to be considered in the longer-term, and strategically; this also needs to be the case with research that is conducted in this field.

Conclusions

This paper has explored issues and needs of those third party and government organisations seeking to implement ICT into teaching and learning practices in developing countries. For all concerned, it is vitally important that the processes involved are as successful as possible in leading to long-term integration and sustainability. In considering patterns of support for those working in developing countries, viewed through outcomes from developed and developing countries, it should be recognised that long-standing integration of ICT into teaching and learning
practices in developed countries has not been achieved at any identifiable widespread level. This paper has considered fundamental reasons for this, problematising these in order to explore alternative conceptions and approaches. It is clear that context is important when looking at change, integration and sustainability; context needs to be fully understood and accommodated if integration is to be successful. Yet current models of integration are very largely based on research and practice arising from applications in developed countries, which do not take sustainability factors fully into account. In view of these limitations and their implied constraints, a new conceptual model is proposed in this paper.

Any new model needs to be trialed and researched, to identify the extent to which it might be applicable within one or more situations. While a single case study is likely to be useful as a starting point, other cases will also need to be considered, so that variations and commonalities can be understood much more fully. In this respect, the role of research in supporting third party and government organisations seeking to integrate ICT in teaching and learning practices is clear; it is important that a research agenda, to enable concepts and outcomes to be fed into the processes and practices of those undertaking integration, is established. This agenda needs to fundamentally explore how long-term rather than short-term needs can be fully identified, accommodated and aligned to support those undertaking and intending change in these areas.

Acknowledgements

A special thanks to: other participants who submitted cases, Mar Mbodj (Université Gaston Berger de St-Louis, Sénégal), Julie Hoffman (Curtin University, Australia), Assetou Kouraogo (Ministry of Education, Burkina Faso), and Allan Yuen (University of Hong Kong); and Jonghwi Park, UNESCO Bangkok Office, for her support.

References


The Impact of the Flipped Classroom on Mathematics Concept Learning in High School

Kaushal Kumar Bhagat1, Cheng-Nan Chang1 and Chun-Yen Chang1, 2*

1Graduate Institute of Science Education, National Taiwan Normal University, Taiwan // 2Science Education Center, National Taiwan Normal University, Taiwan // kaushal.dgp@gmail.com // 697450066@ntnu.edu.tw // changcy@ntnu.edu.tw

*Corresponding author

(Submitted July 1, 2015; Revised October 27, 2015; Accepted December 2, 2015)

ABSTRACT

The present study aimed to examine the effectiveness of the flipped classroom learning environment on learner’s learning achievement and motivation, as well as to investigate the effects of flipped classrooms on learners with different achievement levels in learning mathematics concepts. The learning achievement and motivation were measured by the Mathematics Achievement Test (MAT) and Course Interest Survey (CIS), respectively. A pretest posttest quasi-experimental design was employed for this study. A total of 82 high-school students participated in this study, divided into experimental and control groups. The experimental group (41) was taught trigonometry using the flipped classroom method, while the control group (41) was taught by traditional teaching methods. The researchers employed independent sample t-test, analysis of covariance (ANCOVA), and multivariate analysis of variance (MANOVA) to analyze the data obtained. Findings indicated a significant difference in the learning achievement and motivation between the two groups, with students using the flipped classroom performing better. Further analysis showed a significant difference in the performance of low achievers in the experimental and control groups.

Keywords

Flipped classroom, Mathematics, High school students, Achievement levels, Motivation

Introduction

Mathematics is one of the important subjects in our school curriculum. “In the present quantitatively complex society, a person needs a functional knowledge of mathematical content to make informed decisions as a citizen and as a worker” (Wilkins & Ma, 2003, p. 1). However, Manjul Bhargava, an ace number theorist, who has recently won the Fields Medal, has mentioned that mathematics is taught as a robotic subject (Rajghatta, 2014). This leads the students to struggle in understanding mathematics courses in high school (Offer & Bos, 2009). In fact, mathematics educators are facing one of the major challenges to improve the performance of the students in mathematics (Tan & Tan, 2015).

With the rapid advancement in educational technology, many researchers have recommended the use of technology across the mathematics curriculum as it produces positive results in learning and understanding the concepts (Lazakidou & Retalis, 2010; National Council of Teachers of Mathematics, 2006). Researchers and practitioners have been exploring alternative strategies and teaching methods to engage and motivate the students in learning process. The flipped classroom is one of those alternatives. Many researchers have described the flipped classroom as a model in which learners access the online video lectures uploaded by the instructor prior to the classroom sessions and use class time to participate in meaningful learning activities, instructor-guided problem solving, and discussions (Bergmann, Overmyer, & Wilie, 2012; Chen, Wang, Kinshuk, & Chen, 2014; Fauth, 2015; Hughes, 2012). In addition, the flipped classroom allows the learners to learn at their own pace (Davies, Dean, & Ball, 2013). This results in a paradigm shift from the teacher-centered approach to a student-centered approach (Kong, 2014). Research evidence, on the use of the flipped classroom in the teaching and learning of various disciplines, including statistics, chemistry, English, nursing, engineering, and pharmacy (Davies et al., 2013; Fauth, 2015; Hung, 2015; Mason, Shuman, & Cook, 2013; Missildine, Fountain, Summers, & Gosselin, 2013; Schultz, Duffield, Rasmuseen, & Wageman, 2014; Strayer, 2012; Wilson, 2013) is available, but research in high-school mathematics appears to be limited. To fill this gap, the present study examined the effectiveness of the flipped classroom model on mathematics concept learning in high school.
Theoretical background

The cognitive theory of multimedia learning (CTML) was proposed by Mayer (2001). It was based on dual code theory (Paivio, 1986) and cognitive load theory (Chandler & Sweller, 1991). CTML centered on three assumptions: dual channels, which considered human to possess separate channels for information processing, an auditory/verbal channel and a visual/pictorial channel (Baddeley, 1992; Paivio, 1986); limited capacity, which postulated that human working memory has limited capacity for information processing (Baddeley, 1992; Chandler & Sweller, 1991); and active processing, which means humans must actively select, organize, and store information for long-term memory (Wittrock, 1989). CTML explains why the flipped classroom may improve learning. The flipped classroom incorporated some of the designed principles of the CTML:

- the multimedia principle, which states that students learn better from words and pictures than words alone (Mayer, 2001). The flipped classroom can implement this principle by providing videos that contain text and pictures along with narration.
- the modality principle, which states that students learn better from narration rather than on-screen text (Mayer, 2001). The flipped classroom can implement this principle by providing the explanation of the problem solving in mathematics.
- the individual differences principle, which states that all design principles have a stronger effect on low-knowledge learners. Therefore, within the framework of the CTML, the flipped classroom environment was expected to provide better learning achievement than the traditional classroom.

Related work

In the recent years, the flipped classroom approach has recently gained prominence in the education system due to the rapid development in technology. Numerous studies have shown positive effects of the flipped classroom model in teaching and learning activities. For example, (Davies et al., 2013) compared the traditional, simulation-based, and flipped classrooms in a college-level introductory spreadsheet course. They found that students made better academic gains in the flipped classroom compared to the other two approaches. They were also more satisfied with the learning environment of the flipped model. Similarly, Mason et al.’s (2013) study also found that students performed better in the flipped classroom compared to students in the traditional classroom in an upper-division engineering course and also showed greater satisfaction.

Missildine et al. (2013) conducted a quasi-experimental study and compared three approaches to learning: traditional lecture only (LO), lecture and lecture-capture back-up (LLC), and the flipped classroom approach of lecture capture with innovative classroom activities (LCI) in a nursing course. They also concluded that students in the flipped classroom demonstrated better learning achievement than students in classrooms using other approaches. One of the best features of the flipped classroom that makes it more of a student-oriented approach is that students can stop, rewind, and watch the lecture again, actions which is not possible in the traditional classroom. This allows the students to learn at their own pace (Fautch, 2015; Hung, 2015; Schultz et al., 2014). In addition, the flipped classroom allows the instructor to use the class time in an effective manner. They can cover more topics and also give remedial assistance to the low achievers.

On the other hand, teachers can engage average and high achievers to solve more problems and participate in more class discussions (Davies et al., 2013; Fautch, 2015; Mason et al., 2013). Recently, Bidwell (2014) reported on a pilot program that involved flipped engineering courses from Villanova. The results showed that low achievers grades in the flipped classroom were more than 10% higher than in a traditional classroom. Research studies also found that students demonstrated greater satisfaction and a more positive attitude towards the flipped classroom, which increased their learning motivation (Davies et al., 2013; Hung, 2015; Kong, 2014; Mason et al., 2013). Abeysekera and Dawson (2015) also postulated that the flipped classroom environments are likely to satisfy students’ need for competence, autonomy, and relatedness, which may result in greater levels of both intrinsic and extrinsic motivation of the students. Therefore, a review of studies provides evidence that the flipped classroom has a great potential to enhance learning performance and academic gains among learners.
Issues and research questions

Two conclusions can be drawn from the above review of studies. First, although research in the flipped classroom has been growing significantly in recent years, there is still a lack of studies conducted in high-school mathematics classroom settings. Second, we have not found any study that investigated its effect on students’ mathematics performance based on different achievement levels. Therefore, we have made an attempt to examine the impact of the flipped classroom in learning mathematics at high school. Three research questions that guided this study are:

- Is there any significant difference in the learners’ achievement scores between the experimental and control groups?
- Is there any significant difference in learners’ motivation between the experimental and control groups?
- Are there any significant differences in learners’ performance for students of different achievement levels between the experimental and control groups?

Methodology

Research design and sample

The present study used a pretest/posttest quasi-experimental design. A total of 82 high-school students, aged 14 and 15 years old, participated in the study. The composition of the experimental group was 41 (28 male and 13 females) and control group was 41 (24 male, 17 female). Based on the participants’ previous summative scores in mathematics, they were categorized into high, average, and low achievers in each group. Students who obtained scores ranging from 75 to 100 were categorized as high achievers, 60 to 74 as average achievers, and below 60 as low achievers. The experimental group underwent a lesson on trigonometry with the flipped classroom, while the control group followed a similar lesson using the conventional learning method. Both groups were given a pretest and posttest.

Instruments

Mathematics Achievement Test (MAT)

MAT was administered to measure the students’ performance. The content of the pretest and posttest was same, but the order of the test items was changed to avoid the same-set-response effect. The test items comprised 15 multiple-choice questions on the concept of trigonometry. Thirty-five minutes was allotted for MAT. Content validity of the instrument was determined quantitatively by a panel of five experts. For this purpose, the necessity of the items was assessed using a three-point rating scale: E indicated essential; U, useful but not essential; and N, not necessary. Finally, the content validity ratio (CVR) and the content validity index (CVI) were calculated. The CVR was calculated for each item based on the formula given by Lawshe (1975):

\[
CVR = \frac{n_e - N/2}{N/2}
\]

where \(n_e\) = the number of experts who rated an item as “essential.”

\(N\) = the total number of experts.

The CVI is the mean CVR of all retained items. For a panel of five experts, the minimum value of CVR for an item needs to be at least 0.99 to be accepted (Lawshe, 1975). All the 15 items showed acceptable values of CVR. The calculated CVI was greater than 0.80, which is also acceptable (Polit, Beck, & Owen, 2007). Finally, the Cronbach’s \(\alpha\) of the instrument was 0.723, which is acceptable (Barrett, 2001). This showed the validity of the instrument.

Course Interest Survey (CIS)

Keller (2010) designed a course interest survey (CIS) to measure students’ motivational reactions to the instructor-led-instruction. Attention, relevance, confidence, and satisfaction are the four factors of CIS. It contains 34 items with five-point Likert-scale items. The CIS is considered a valid instrument with a documented reliability coefficient.
of 0.95 (Keller, 2010). The survey was translated into Chinese before administration by a language expert. Cronbach’s α coefficient was calculated to verify the instrument’s internal validity. The four factors of the questionnaire had an adequate reliability (Cronbach’s α = 0.80, 0.82, 0.79, 0.84 for attention, relevance, confidence, and satisfaction, respectively).

Procedure

The duration of this study was six weeks. Two weeks before the intervention, both the experimental and control groups underwent a pretest. The authors selected three teaching modules based on the concept of trigonometry for this study: introduction to trigonometry, trigonometric ratios, and trigonometric identities. In the control group, instruction was provided in the classroom. Students were asked to attend the classroom-based lectures and complete their homework before the next class on their own. Thirty to forty minutes of the total class duration (50 minutes) was devoted for lecture and discussion. The remaining time was for problem solving. Homework consisted of textbook problems. On the other hand, in the experimental group, pre-recorded video lessons were uploaded to the dropbox the week before class. The average duration for each lesson was 15 to 20 minutes. Students were asked to watch the video lesson before coming to the class. During classroom time, students were engaged in the activities based on video lessons. Students were divided into groups to discuss the textbook problems. In the meantime, the students who needed remedial assistance were given face-to-face support. At the end of the intervention, both groups were given the posttest and CIS to collect scores for learning achievement and motivation, respectively.

Data analysis

Frequency and percentage were used for descriptive statistics. The data were also presented graphically to highlights the similarities and differences in the results. With respect to the learning achievement, a one-way analysis of covariance (ANCOVA) was conducted with group as a between subject variable, pretest scores as a covariate, and posttest scores as a dependent variable. To prevent the effect of sampling error, we applied ANCOVA. Independent sample t-test was conducted to determine the effects of the flipped classroom on learners with different achievement levels. With respect to the motivation, one-way multivariate analysis of variance (MANOVA) was conducted with group as a between subject variable and attention, relevance, confidence, and satisfaction as dependent variables. All analyses were conducted using the Statistical Package for the Social Sciences, Version 21 (SPSS 21). The statistical significance level was set at $p < 0.05$.

Results

Table 1 shows the total number of students with different achievement levels in mathematics for both groups. We observed that the distributions of students for high, average, and low achievers were almost similar. Before intervention, independent sample $t$-test was conducted on pretest scores collected from the experimental group and the control group. Results indicated that there was no significant ($t$-value = -1.167, $p = .247$) difference between the two groups. This showed that both the groups were similar in abilities before the intervention was conducted.

<table>
<thead>
<tr>
<th>Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>17(41.4%)</td>
<td>16(39.02%)</td>
<td>8(19.5%)</td>
</tr>
<tr>
<td>Control</td>
<td>14(34.1%)</td>
<td>19(46.34%)</td>
<td>8(19.5%)</td>
</tr>
<tr>
<td>Total</td>
<td>31</td>
<td>35</td>
<td>16</td>
</tr>
</tbody>
</table>

Learning achievement on different teaching methods

Before employing ANCOVA, Levene’s homogeneity test was conducted. The result showed the $F$ value was equal to 3.28 ($p > .05$). This indicated that the homogeneity test has not achieved statistical significance; therefore, ANCOVA could be applied. Table 2 summarizes the ANCOVA results in comparing student’s achievement scores. Students in the experimental group exhibited better performance than the control group, $F(1, 79) = 8.001$, $p < .05$, $\eta^2 = .092$. 
which is considered to be a large effect (Cohen, 1988). Figure 1 displays the graphical representation of the mean scores, with confidence intervals for the both groups for pretest and posttest scores. There was overlapping between control and experimental group for pretest scores, but not for posttest scores. This provided sufficient evidence that population means were same for pretest scores but different for posttest scores.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Partial eta squared</th>
<th>Observed power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Test</td>
<td>.595</td>
<td>1</td>
<td>.595</td>
<td>.118</td>
<td>.733</td>
<td>.001</td>
<td>.063</td>
</tr>
<tr>
<td>Group</td>
<td>40.461</td>
<td>1</td>
<td>40.461</td>
<td>8.001</td>
<td>.00</td>
<td>.092</td>
<td>.798</td>
</tr>
<tr>
<td>Error</td>
<td>399.503</td>
<td>79</td>
<td>5.057</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05.

Learning achievement based on different achievement level

Table 3 shows the results of learning achievement of the participants based on the different achievement level for pretest and posttest for both groups. We found that there were no significant differences in the students’ learning achievement between the groups for all levels except for average achievers. The mean scores of the pretest of the high achievers \((M = 13.11, SD = 1.45)\) and low achievers \((M = 5.25, SD = 2.12)\) in the experimental group are not significantly different from the control group \((M = 12.57, SD = 1.55; M = 5.75, SD = 1.38)\), respectively. However, significant difference is found in the posttest for all levels except for high achievers. The results show the mean scores of the average achievers \((M = 10.75, SD = 2.40)\) and low achievers \((M = 9.18, SD = 2.71)\) in the experimental group are significantly different from those in the control group \((M = 8.94, SD = 1.71; M = 7.62, SD = 2.26)\).

Figure 2 and Figure 3 display the graphical representation of the mean scores for pretest and posttest, with confidence intervals for the high, average, and low achievers. There was overlapping between control and experimental group for high and low achievers, but there was no overlapping for average achievers for the pretest scores. This provided sufficient evidence that population means were the same for pretest scores for high and low achievers. However, there was no overlapping between the control and experimental groups for average and low achievers. This provided sufficient evidence that population means were different for posttest scores for average and low achievers.
Table 3. Independent sample t-test for pre-test and posttest between the groups based on the achievement level

<table>
<thead>
<tr>
<th>Achievement level</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>MD</th>
<th>t-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>High</td>
<td>13.11</td>
<td>1.45</td>
<td>0.54</td>
<td>1.62</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>12.57</td>
<td>1.55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>Experimental</td>
<td>10.12</td>
<td>2.15</td>
<td>1.07</td>
<td>2.96*</td>
<td>0.00*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>9.05</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>Experimental</td>
<td>5.25</td>
<td>2.12</td>
<td>-0.50</td>
<td>1.26</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>5.75</td>
<td>1.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td>High</td>
<td>10</td>
<td>2.54</td>
<td>0.13</td>
<td>0.25</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>9.87</td>
<td>2.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>Experimental</td>
<td>10.75</td>
<td>2.40</td>
<td>1.81</td>
<td>3.93*</td>
<td>0.00*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>8.94</td>
<td>1.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>Experimental</td>
<td>9.18</td>
<td>2.71</td>
<td>1.56</td>
<td>2.83*</td>
<td>0.00*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>7.62</td>
<td>2.26</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05. SD = Standard deviation, MD = mean difference.

Figure 2. Error bar chart for pre-test based on the achievement level

Figure 3. Error bar chart for posttest based on the achievement level
Learning motivation

Descriptive statistics (shown in Table 4), including means and standard deviation are provided for dependent variables attention, relevance, confidence, and satisfaction. The results of MANOVA revealed that there was significant difference for attention, relevance, confidence, and satisfaction between the experimental and control groups, \( \text{Wilk's } \Lambda = .68, F = 8.90, p < .05, \eta^2 = .31 \). Therefore, univariate \( F \) tests were conducted for attention, relevance, confidence, and satisfaction. As shown in Table 5, the results of univariate \( F \) tests indicated a significant difference between the groups for attention \( (p < .05, \eta^2 = .17) \), relevance \( (p < .05, \eta^2 = .15) \), confidence \( (p < .05, \eta^2 = .10) \), and satisfaction \( (p < .05, \eta^2 = .18) \).

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG</td>
<td>2.80</td>
<td>.77</td>
<td>41</td>
</tr>
<tr>
<td>EG</td>
<td>3.55</td>
<td>.86</td>
<td>41</td>
</tr>
<tr>
<td>CG</td>
<td>3.16</td>
<td>.83</td>
<td>41</td>
</tr>
<tr>
<td>EG</td>
<td>3.84</td>
<td>.73</td>
<td>41</td>
</tr>
<tr>
<td>CG</td>
<td>3.16</td>
<td>.69</td>
<td>41</td>
</tr>
<tr>
<td>EG</td>
<td>3.64</td>
<td>.72</td>
<td>41</td>
</tr>
<tr>
<td>CG</td>
<td>2.96</td>
<td>.74</td>
<td>41</td>
</tr>
<tr>
<td>EG</td>
<td>3.66</td>
<td>.74</td>
<td>41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DV</th>
<th>SV</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Partial ( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>Group</td>
<td>11.53</td>
<td>1</td>
<td>11.53</td>
<td>17.16*</td>
<td>.17</td>
</tr>
<tr>
<td>Relevance</td>
<td></td>
<td>9.39</td>
<td>1</td>
<td>9.39</td>
<td>15.12*</td>
<td>.15</td>
</tr>
<tr>
<td>Confidence</td>
<td></td>
<td>4.63</td>
<td>1</td>
<td>4.63</td>
<td>9.20*</td>
<td>.10</td>
</tr>
<tr>
<td>Satisfaction</td>
<td></td>
<td>9.90</td>
<td>1</td>
<td>9.90</td>
<td>17.76*</td>
<td>.18</td>
</tr>
</tbody>
</table>

Note. *\( p < .05 \).

Discussion and conclusions

The present study compared the flipped classroom with the conventional method of teaching a mathematical concept, trigonometry, in order to examine its learning effectiveness. Both interventions were designed to deliver the same learning content by the same instructor. The statistical results of this study indicated that students in the experimental group outperformed in the posttest than the control group. This suggests that the flipped classroom environment improved learning achievement of the students in the experimental group. This result is consistent with previous studies (Davies et al., 2013; Mason et al., 2013; Missildine et al., 2013). Moreover, CTML theory was adopted in the flipped classroom environment because students were asked to watch uploaded video lessons. Students accessed the videos at their convenience time and also re-watched the lessons, which was not possible in the conventional method of teaching. In addition, we found that students were highly satisfied with and positive about the flipped classroom. This resulted in greater learning motivation. Previous studies (Davies et al., 2013; Mason et al., 2013) have also provided the similar results.

The findings also revealed that low achievers in the experimental group performed better than the control group, but the performance of the high and average achievers remains same. This outcome supports the previous study by Bidwell (2014). In the flipped classroom mode, low achievers got more attention from the teachers, and they discussed the problems to understand the mathematical concept. Therefore, the flipped classroom mode may help low achievers to improve their performance in mathematics.

A considerable number of high-school students belong to the category of low achievers. Thus, it is necessary to develop teaching methods to help low achievers better understand mathematical concepts. The present study has contributed to mathematics education by providing the empirical evidence of the potential of the flipped classroom to support teaching and learning in mathematics. The findings of this study support that a student-centered approach in the flipped classroom is better than the teacher-centered approach in the conventional method of teaching. This study has also shown that the flipped classroom benefits lower achievers more than high and average achievers. To
conclude, with the advancement of technology and the adoption of the flipped classroom, both teaching and learning mathematics can be made more enjoyable and effective.

**Limitations and future research**

The duration of this study was limited to six weeks; this is one of the limitations of this study. Future studies should cover at least one module for two to three months for more concrete scientific findings. Another limitation of this study was that students were not able to ask their questions immediately while watching the lesson videos. Future studies may provide an online discussion forum for giving opportunity and motivating the students for pondering questions to develop critical thinking and engagement. Interactive video lessons are also recommended for further studies in the flipped classroom to make learning more meaningful.

**Acknowledgements**

This project is supported in part by the Ministry of Science and Technology, Taiwan, R.O.C. (previously named National Science Foundation of Taiwan) under the Grant number NSC102-2515-S-103-001-H, “Aim for the Top University Project” sponsored by the Ministry of Education, Taiwan, R.O.C., and “International Research-Intensive Center of Excellence Program” of NTNU and Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. MOST 104-2911-I-003-301.

**References**


The Effect on Pupils’ Science Performance and Problem-Solving Ability through Lego: An Engineering Design-based Modeling Approach

Yanyan Li*, Zhinan Huang, Menglu Jiang and Ting-Wen Chang

R&D Center for Knowledge Engineering, School of Educational Technology, Beijing Normal University, Beijing, China // lyy@bnu.edu.cn // chinanwong@mail.bnu.edu.cn // jiangml@mail.bnu.edu.cn // tingwenchang@bnu.edu.cn

*Corresponding author

(Submitted July 3, 2015; Revised September 15, 2015; Accepted December 29, 2015)

ABSTRACT

Incorporating scientific fundamentals via engineering through a design-based methodology has proven to be highly effective for STEM education. Engineering design can be instantiated for learning as they involve mental and physical stimulation and develop practical skills especially in solving problems. Lego bricks, as a set of toys based on design technique, are in line with the cognitive characteristics of students and provide a good game-based learning tool for engineering education. Many studies have incorporated Lego to investigate students’ scientific attitudes, science inquiry skills, and problem-solving strategies, yet few address the effects of engineering design-based learning on pupils’ problem-solving abilities. Therefore, this paper conducted an experiment and included control group to examine how fourth-grade students’ science performance and problem-solving abilities change over the engineering design-based science learning by using Lego bricks. Results indicate that: (1) pupils’ science performance significantly improved on both the control and the experimental groups, (2) pupils’ gains of problem-solving ability in the experimental group were significantly improved, and (3) the males made a significant progress in problem-solving ability than the females in the experimental group. The key findings, possible reasons behind them, and potential benefits in the context of learning are also discussed.

Keywords

Engineering design, Problem-solving ability, Science performance, Elementary school

Introduction

Many studies have addressed the need to enhance K–12 STEM (science, technology, engineering, and mathematics) education, in which engineering has been prominently recognized as an important recommended component for children from early childhood through high school (e.g., Brophy, Klein, Portsmore & Rogers, 2008; Bybee, 2011; Carr, Bennett & Strobel, 2012). Moreover, researchers have also argued that engineering design (an inquiry-based pedagogical strategy that learning across disciplines) should be an important component of the precollege education of all youths (e.g., Diefes-Dux, 2015; Mentzer, Huffman & Thayer, 2014) because engineering design can provide students with an effective approach to tackle open-ended problems. Previous research has found engineering design to be a supportive context for science learning, and designing functional artifacts may ground children’s exploration of scientific concepts (Bethke Wendell & Rogers, 2013). However, it is evident that the “E” of STEM has been virtually ignored in most elementary schools in China. Science courses lack any engineering design principles, and the students obtain discrete engineering knowledge unsystematically (Wang, 2014).

On the other hand, proponents of game-based learning suggest that educational games induce a positive experience that may be harnessed for learning (Connolly, Boyle, MacArthur, Hainey & Boyle, 2012). Game-based learning may encourage students to acquire knowledge and offer a rich context that allows students to reinforce and consolidate their knowledge through practice. Lego bricks, as a set of toys based on design technique, are in line with the cognitive characteristics of students and provide a good game-based learning tool for engineering education. A number of studies have explored the effectiveness of engineering design to improve science learning by using Lego. Most of these studies focused on various aspects of problem solving, such as problem-solving attitude and problem-solving strategies (Hussain, Lindh & Shukur, 2006; Sullivan, 2008; Bethke Wendell & Rogers, 2013). Indeed, researchers have indicated the importance of students’ problem-solving abilities in using engineering design, which may help students better recognize the utility of knowledge learned in the classroom for solving real-world problems (Fortus, Krajcik, Dershimer, Marx & Mamlok-Naaman, 2005). However, there are few studies focusing on the problem-solving ability and few of the studies have been empirically validated or experimentally verified. In this regard, we note that more research is needed to determine the efficacy of the engineering design-based approach and to examine whether it contributes to science learning for a wider range of students.
In this study, we proposed an engineering design-based modeling approach to provide a new instructional strategy in the science learning. During the research, we used control and experimental groups to investigate whether engineering design-based learning changed elementary student problem-solving abilities and science knowledge. Our experimental results show that engineering design-based learning can effectively improve students’ problem-solving abilities as well as science performance. The remainder of this paper presents a brief discussion of the application of engineering design, the model used in this study. The research design and analytical methods are provided in the third section, followed by results, discussion, and conclusions.

Engineering design

The theoretical foundation underpinning this study adheres to the constructionism theory (Papert, 1980): Children construct knowledge and build on their background knowledge. Papert emphasizes that students learn most effectively when they actively produce tangible artifacts with manipulative materials in the real world. Engineering design provides a way for students to gain a deeper understanding through “learning by making,” which is considered an important tenet of constructivism (Williams, Ma, Prejean, Ford & Lai, 2007). Although many researchers have discussed engineering design, a consensus of its definition has not yet been reached. Dym, Agogino, Eris, Frey and Leifer (2005) articulates the definition of engineering design as: “Engineering design is a systematic, intelligent process in which designers generate, evaluate, and specify concepts for devices, systems, or processes whose form and function achieve clients' objectives or users' needs while satisfying a specified set of constraints.” Thus, engineering is the application of science to problem solving, and design is the creative expression of knowledge. In this section, we begin with a review of the engineering design application in science learning, and then follow with the model of engineering design used in this study.

Engineering design application in science learning

The existing research concerning engineering design-based science learning has mainly focused on contributions to science content learning. These studies generally describe significant results that promote the personalized comprehension of science concepts involving physics, i.e., Newton’s laws of motion, distances and angles, and kinematics (Berland, Steingut & Ko, 2014; Mitnik, Nussbaum & Soto, 2008). For example, Bethke Wendell and Rogers (2013) investigated the effect of an engineering design-based curriculum on elementary student science attitudes and science content using Lego. Their results showed significantly better performance in science content but minimal difference in science attitude.

In addition to science content learning, some studies emphasize the development of skills through engineering design by examining the effect of engineering design on various aspects of problem solving, such as problem-solving attitude and problem-solving strategies. For example, the work of Lindh and Holgersson (2007) concerned attitudes towards problem solving in pupils aged 11–12 and 15–16 years, but no statistical evidence has shown that the average pupil gains ability to solve logical problems from Lego training, and Sullivan (2008) analyzed how students used thinking skills (observation, estimation, and manipulation) and science process skills characteristic of scientifically literate individuals to solve a design challenge while Capobianco, Yu and French (2015) examined the effects of engineering design on students’ science learning to explore their engineering identity development.

However, the observed results for the studies were inconclusive because the authors presented both significant and non-significant improvement in problem-solving strategies and attitude (Hussain et al., 2006). Furthermore, few studies have experimentally verified how engineering design-based learning affects problem-solving ability. To facilitate the research process, we examined fourth-grade students to determine the efficacy of the engineering design-based approach and examine whether it contributes to science learning.

Model of engineering design

A general model of engineering design is illustrated by Sheppard, Macatangay, Colby, and Sullivan (2009). The process begins with problem identification; continues through the development of a conceptual design, prototyping, and testing; and ultimately culminates in a sustainable implementation plan resulting in the creation of a marketable
product designed to solve the initial problem. The model is especially helpful in visualizing the “iterative and intertwined nature of defining, generating, testing, and evaluating ideas.”

In moving from the workplace and the practice of engineering to the school setting, there are also multiple representations of the engineering design process that highlights key features of a process appropriate for a school or a classroom. The Massachusetts K–12 curriculum framework provides a comprehensive and accurate representation of the engineering design process (Massachusetts Department of Education, 2006), as shown in Figure 1. The model comprises eight steps, yet it is difficult for students to experience the entire cycle in the limited time afforded by a classroom lesson. Moreover, the steps of “communicate the solution” and “redesign” run through the entire design process. These two steps initiate a re-examination of the need and restart the engineering design cycle. Thus, we simplified the model as a five-step version for practical use in the classroom. The five steps are described in Table 1.

Table 1. Engineering design used in this study

<table>
<thead>
<tr>
<th>Steps</th>
<th>Engineering design process</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Find a problem</td>
<td>Students are asked to unpack the overarching engineering design challenge and the problem by considering what conditions already exist and what they still need to solve.</td>
</tr>
<tr>
<td>2</td>
<td>Develop possible solutions</td>
<td>Students are required to develop possible solutions that utilize knowledge and skills needed for the overarching design challenge.</td>
</tr>
<tr>
<td>3</td>
<td>Decide the optimal solution</td>
<td>Students need to consider how their findings inform their decision-making process about the design for all possible solutions.</td>
</tr>
<tr>
<td>4</td>
<td>Build a prototype</td>
<td>Students are assigned to work in groups to construct an artifact that meets the requirements of the overarching design challenge.</td>
</tr>
<tr>
<td>5</td>
<td>Test the prototype</td>
<td>Students are asked to test and improve their solutions to the overarching design challenge, and present an explanation of how it works.</td>
</tr>
</tbody>
</table>

Research design

This research design is written in subsections discussing participants, the hypothesis, methods, procedures, and learning scenarios.
Participants

The participants were a class of fourth-grade students (10 years old) recruited from an affiliated primary school located in Beijing, China. The participants comprised 30 students (10 female and 20 male), who were randomly divided into two groups. Both groups were taught by the same teacher but with different pedagogies, as shown in Table 2. The control group (CG) used commonly used science pedagogy with Lego bricks. The teaching steps included “creating situations,” “analyzing problems,” “building a prototype,” and “testing a prototype,” while the experimental group (EG) used an engineering design-based pedagogy with Lego bricks. Unlike the CG, the second step in the EG was divided into three more detailed sections. Pupils were required to clearly describe the problems on a design card, to analyze the constraints and draw a design diagram on a design card, and to decide the optimal solution after comparing potential solutions with a score sheet. Pupils in both groups were divided into smaller groups (namely, working groups) with identical numbers of high-, medium-, and low-performing students (to maintain relative homogeneity), according to the results of their pre-tests. Each working group of three pupils (two male and one female) was given a set of Lego components and an assistant who was trained in how to handle the Lego bricks and prepare the students to solve the highly intricate questions.

Table 2. Pedagogy used in the CG and the EG

<table>
<thead>
<tr>
<th>Commonly used pedagogy (CG)</th>
<th>Engineering design-based pedagogy (EG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creating situations</td>
<td>Creating situations</td>
</tr>
<tr>
<td>Analyzing problems</td>
<td>Identifying problems</td>
</tr>
<tr>
<td>Building a prototype</td>
<td>Analyzing potential solution</td>
</tr>
<tr>
<td>Testing a prototype</td>
<td>Deciding an optimal solution</td>
</tr>
<tr>
<td>Students start to construct an artifact.</td>
<td>Students then decide the optimal solution after comparing potential solutions with a score sheet.</td>
</tr>
<tr>
<td>Students are asked to test and modify their solutions.</td>
<td>Students are asked to test and modify their solutions.</td>
</tr>
</tbody>
</table>

Hypothesis

In this study, we investigated the relationship between learning science through Lego engineering design activities and students’ performance in science subjects as well as problem-solving abilities (dependent variable) using two distinct pedagogical methods (independent variables). The research questions were as follows: Are there significant differences in science performance between the EG and the CG? Are there significant differences in problem-solving ability between the EG and the CG? Two hypotheses were set up:

- Pupils in the engineering design-based group (EG) will develop better knowledge in science than pupils of the commonly used learning group (CG) by using Lego bricks.
- Pupils in the engineering design-based group (EG) will develop better problem-solving ability than pupils of the commonly used learning group (CG) by using Lego bricks.

Methods

In this study, we mainly used quantitative research methods, supplemented with a qualitative method. The quantitative methods consist of different tests in physics and problem solving to examine how both the CG and the EG groups’ performances changed from pre-test to post-test and whether one, both or neither improved over time. The qualitative method was observation. Every assistant involved in this study was required to document the entire problem-solving process as thoroughly as possible by providing detailed descriptions of the cognitive processes and concepts developed by the students.
Physics tests

Two sets of physics tests were used in this experiment as a pre-test and a post-test. To avoid any possible bias resulting from a similarity in the nature of the questions between the pre-test and the post-test and to ensure that the scores students received were comparable, the two tests were constructed differently with the same level of difficulty. Both tests were selected from the previous years’ fourth-grade science final examination developed by the teacher and the researchers. Furthermore, the researchers validated the questions word by word to enhance clarity and to avoid any possible misunderstanding of the questions by the students.

As shown in Table 3, the structures of the two tests were the same. Both of the tests involved three knowledge points, such as knowing how to reduce power by increasing the length of the power arm to balance the lever, finding the gear ratio by simply identifying how many teeth are on each gear, or recognizing that idler gears do not affect the gear ratio of the overall train. Each of the knowledge points contains two true/false questions and three multiple-choice questions. In total, there were fifteen items, each worth one point.

<table>
<thead>
<tr>
<th>Knowledge point</th>
<th>Question type</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lever</td>
<td>True/False</td>
<td>Multiple choice</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Gear</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Pulley</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

Problem-solving ability questionnaire

The problem-solving ability questionnaire, administered to the pupils subjectively before and after the experiment, adopted a Likert four-point scale (ranging from “1” for “never” to “4” for “always”) which consisted of five dimensions: “find a problem,” “develop possible solutions,” “decide the optimal solution,” “build a prototype,” and “test the prototype.” Each dimension contained five items. This questionnaire consisted of 25 items in total and was modified from the Problem-Solving Ability Self-Checking Questionnaire (Li, 2003); the Cronbach’s alpha coefficient was 0.895 in this experiment, suggesting that the questionnaire is reliable.

Assessment of artifacts

In order to assess the artifacts built by the pupils objectively, we used a rubric modified from the Technical Works Assessment Scale (Lin, 2011) to evaluate the final products in terms of accuracy, integrity, and practicality. Each dimension comprises detailed description ranging from “1” for “poor” to “5” for “excellent.” The teacher scores the final products from each group throughout the assessment rubric.

Observation

During this research, observation is also needed to evaluate the process of pupils’ problem solving (i.e., figuring out the key problem of the tasks and answering conceptual questions). While designing and building the prototype, the pupils’ actions were documented to see how they work, think about and accept new challenges and seek solutions, as well as the frequency at which they revise or iterate while considering the problems. The collected information was successively used to illustrate how the children developed their knowledge and help the researchers analyze the pupils’ problem solving during the process.

Procedure and learning scenarios

As shown in Table 4, the researchers developed a set of science activities for fourth-grade students by consulting the teachers according to the school syllabus concerning science knowledge, such as gear use, the lever principle, etc.
Both the control group and the experimental group were engaged in the same five activities. The first activity was a warm-up for both the CG and the EG to familiarize the working-group members with each other and the Lego bricks. The only difference between the groups was that the teachers introduced the engineering design process in the experimental group. The training time was approximately two hours per week for five weeks, extending from November to December 2014. The assistants were directed to observe the EG to follow up on how well the work continued and what eventually was changed during that time.

Table 4. Activities used in the study

<table>
<thead>
<tr>
<th>Design activity</th>
<th>Assessment item</th>
<th>Task description</th>
</tr>
</thead>
</table>
| Garden          | *               | Task 1: Design a tower, the taller the better.  
|                 |                 | Task 2: Design a garden that is your favorite. |
| Crane           | Pulley principle| Design a crane that can easily hoist weights, the heavier the better. |
| Fan             | Gear use        | Design a fan with a gear, the faster the better. |
| Platform scale  | Lever principle | Design a platform scale that can weigh goods, the heavier the better. |
| Folding table   | Shape stability | Design a folding table that can easily open, fold, and lift weights, the heavier the better. |

Note. * means the task was unstructured to allow students to become familiar with the engineering design process.

Each activity followed a similar engineering design process. The teacher proposed an ill-structured problem in the manner of creating a situation. For example, in the "crane" activity, the students were required to design a crane that could easily hoist weights to the roof of the building, although the weights could not be carried up directly with certain restricted materials. The students worked in groups to discuss and write a response to the exploration problem.
reflecting with clarity on their solution tactics. Next, the students designed possible acceptable solutions on their design cards and adequately compared or integrated them to decide the optimal solution. As shown in Figure 2, the students were asked to build an artifact in each activity, and the assistants interviewed them to investigate the underlying science principles during and after the building process, to document the problem-solving process, and to ascertain whether they gained a correct understanding of the illustrated principle. All the final products were evaluated by the teacher according to the assessment of the artifacts through the rubric.

**Results**

In order to compare the different influences between the two teaching pedagogies, initially, we analyzed the data about learning performance and problem-solving ability to find the changes the pupils had after the experiment both in the CG and the EG. To do so, a Wilcoxon signed-rank test was conducted on the pre-test and post-test scores of learning performance and problem-solving ability in both groups. We found that pupils had a significant improvement about learning performance both in the CG and the EG and no significant changes about problem-solving ability either in the CG or the EG. To mine more information about the changes, we compared the gains which were computed by subtracting the pre-test scores from post-test scores in the CG and the EG. An independent-sample test on the gains between the CG and the EG in learning performance and problem-solving ability showed pupils had significantly different changes between the CG and the EG in problem-solving ability.

To explore how the engineering-designed Lego activities worked, we conducted a Wilcoxon signed-rank test on the pre-test and post-test scores about each dimension of problem-solving ability in the EG. The students showed significant improvement in the ability to identify optimal solutions to the problem at hand after using engineering design-based pedagogy. Moreover, in order to examine the gender difference in problem-solving ability of the EG, a Mann-Whitney test on gains of girls and boys in problem-solving ability showed there was a significant gender difference in problem-solving ability.

**Learning performance analysis results**

Tests were designed to measure learning performance and were distributed (30 for the pre-test and 30 for the post-test in the experimental and control groups). A Wilcoxon signed-rank test on the pre-test and post-test scores was conducted on both the experimental group and the control group. The results are shown in Table 5.

The difference between the post-test scores and the pre-test scores for both groups was three points, which implies that students in both groups improved significantly. To uncover more meaningful information regarding the impact of commonly used pedagogical activities and engineering design-based pedagogical activities, we conducted an independent-sample test to measure the improvement in learning performance. As shown in Table 6, students in both groups improved by the same magnitude, which implies that there are no significant differences (effect size $r = 0.013$) between the two teaching pedagogies with respect to measuring the students’ performance improvement in learning science.

| Table 5. Learning performance gains by the CG and the EG |
|------------|-------------|-------------|-------------|
| Group | N | Pre-test | Post-test | Gains |
| CG | 15 | 7.733 | 1.668 | 10.733 | 2.576 | 3.000 | –3.020** |
| EG | 15 | 7.800 | 1.699 | 10.867 | 2.356 | 3.067 | –3.000** |

*Note. **p < .01.*

| Table 6. Independent-sample test on learning performance gains |
|------------|-------------|-------------|
| Group | CG | EG | t | df |
| Performance | 3 | 3.067 | -0.065 | 28 |
Problem-solving ability analysis results

To measure problem-solving ability, we reverted to a modified version of the well-known self-checking questionnaire (Li, 2003). Thirty copies of the questionnaire were utilized for the pre-test, and another thirty copies were used for the post-test. The results are shown in Table 7.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pre-test Mean</th>
<th>SD</th>
<th>Post-test Mean</th>
<th>SD</th>
<th>Gains</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG</td>
<td>15</td>
<td>81.400</td>
<td>11.587</td>
<td>78.733</td>
<td>12.792</td>
<td>-2.667</td>
<td>-1.217</td>
</tr>
<tr>
<td>EG</td>
<td>15</td>
<td>80.533</td>
<td>12.252</td>
<td>84.533</td>
<td>12.558</td>
<td>4</td>
<td>-1.857*</td>
</tr>
</tbody>
</table>

Note. *p < .10.

Examining the results above, we can observe a 2.667 point decrease in the scores of the control group, whereas the experimental group has an improvement of four points. These results indicate that there are no significant improvements for the control group. However, it is noteworthy that the p-value of the EG is nearly significant, which may imply that use of an engineering design-based pedagogical method enhances problem-solving abilities. Table 8 shows a significant difference (effect size $r = 0.400$) between the two teaching pedagogies with respect to measuring the students’ problem-solving ability.

<table>
<thead>
<tr>
<th>Group</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG</td>
<td>-2.667</td>
<td>28</td>
</tr>
<tr>
<td>EG</td>
<td>4</td>
<td>-2.305*</td>
</tr>
</tbody>
</table>

Note. *p < .05.

Table 9. Problem-solving gains in five dimensions by the EG

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Pre-test Mean</th>
<th>SD</th>
<th>Post-test Mean</th>
<th>SD</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a problem</td>
<td>16.733</td>
<td>3.327</td>
<td>17.467</td>
<td>3.021</td>
<td>-0.828</td>
</tr>
<tr>
<td>Research possible solutions</td>
<td>15.533</td>
<td>3.137</td>
<td>16.333</td>
<td>3.478</td>
<td>-1.191</td>
</tr>
<tr>
<td>Decide the optimal solution</td>
<td>15.733</td>
<td>3.807</td>
<td>17.800</td>
<td>3.005</td>
<td>-2.275*</td>
</tr>
<tr>
<td>Build a prototype</td>
<td>16.067</td>
<td>2.549</td>
<td>15.400</td>
<td>2.947</td>
<td>-0.366</td>
</tr>
<tr>
<td>Test the prototype</td>
<td>16.467</td>
<td>3.543</td>
<td>17.533</td>
<td>2.850</td>
<td>-1.253</td>
</tr>
</tbody>
</table>

Note. *p < .05.

Table 10. Gender difference in problem-solving ability gains

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>N</th>
<th>Pre-test Mean</th>
<th>SD</th>
<th>Post-test Mean</th>
<th>SD</th>
<th>Gains</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG</td>
<td>Boy</td>
<td>10</td>
<td>81.700</td>
<td>11.908</td>
<td>81.800</td>
<td>13.538</td>
<td>0.1000</td>
<td>-1.539</td>
</tr>
<tr>
<td>EG</td>
<td>Boy</td>
<td>10</td>
<td>79.000</td>
<td>12.605</td>
<td>85.800</td>
<td>12.717</td>
<td>6.8000</td>
<td>-2.517*</td>
</tr>
<tr>
<td></td>
<td>Girl</td>
<td>5</td>
<td>83.600</td>
<td>12.260</td>
<td>82.000</td>
<td>13.267</td>
<td>-1.6000</td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05.

To understand how the engineering design-based pedagogical method affected problem-solving ability, five Wilcoxon signed-rank tests were conducted on each ability dimension. The results are displayed in Table 9; only the dimension “decide the optimal solution” showed a significant improvement.

We examined improvements in problem-solving ability by gender using a Mann-Whitney test on the experimental group. The results are shown in Table 10. According to these results, male participants increased their scores by 6.8 points, whereas female participants decreased their scores by 1.6 points. Table 10 indicates that there is a significant difference in problem-solving ability between genders; male problem solving progressed more than that of females.
Assessment of artifacts

To make the assessment of the artifacts during the four activities for both the experimental group and control group, we evaluated the final products for each activity through a rubric involving three dimensions such as accuracy, integrity and practicality. Each dimension ranges from “1” for “poor” to “5” for “excellent,” and the total score of the rubric is fifteen. The results are shown in Figure 3. We can find that there are minor variations of the scores of the final products over the four activities in each group in the CG, whereas there is an upward trend of the scores of the final products over the four activities in each group in the EG. The results imply that the engineering design-based learning helps to improve the pupils’ problem-solving ability and helps them to develop better designed products.

![Figure 3. Scores of the artifacts in the CG and the EG](image)

Discussion

Performance of engineering design-based pedagogy

The experimental group and the control group both improved in their science tests, a result which is consistent with previous research (Hussain et al., 2006; Williams et al., 2007). Additionally, there were no significant differences between the academic performances of the two groups. We believe that one possible reason for this result is that teaching with Lego improves student enthusiasm versus traditional learning experiences in school. Students perform better through learning by doing using exercises such as activities with Lego. We also interviewed the students and determined that most preferred the Lego learning activities to traditional classroom activities. The students responded that Lego learning incorporates fun, and by playing, they learn through their own experiences and deepen their understanding of the content.

Problem-solving ability in engineering design-based pedagogy

A number of studies have examined the effect of engineering design on various aspects of problem solving, such as attitude towards problem solving and strategies for problem solving (Lindh & Holgersson, 2007; Williams et al., 2007). Although few studies focused on problem-solving abilities, Hussain et al. (2006) found that students did not show a positive attitude towards problem solving in a Lego activity. In contrast, Lindh and Holgersson (2007) investigated the logical problem-solving abilities of students but did not reach a conclusion regarding student can gain a lot by becoming an engaged and active member from design-based activities. To further explore this approach to improve students’ problem-solving abilities, we examined student performance in learning subjects while using an engineering design-based pedagogy. The experimental results showed significant difference in problem-solving abilities between the experimental group and the control group.

Our results show that engineering design-based pedagogy tends to enhance student problem-solving abilities through Lego activities. We believe that there are two possible reasons:

First, engineering design-based pedagogy provides a good scaffold for students to learn science. The teacher proposed ill-structured problems through the creation of situations and guided the students to construct their own
problems step by step. The students then designed all possible solutions on a design card while discussing and revising. They scored each solution using a score sheet, as shown in Figure 4, to measure originality, stability and artistic effect. Students can select the best solution from their designs or combine them into an integrated design solution. This process allowed the students to discover the final optimal solution.

![Figure 4. Solution score sheet](image)

During our research, we also found that some working groups in the CG only briefly discussed or jumped directly into implementation without analyzing the problem, which led to many revisions and iterations in the later part of the exercise. In contrast, with adequate discussion for identifying different strategies built into the process, the students in the EG were provided a systematic blueprint for analyzing and choosing physical structures, a method that helped them in the final part of the exercise. The students in the EG spent fewer times building or disassembling the Lego bricks than their counterparts. Obviously, the engineering design-based pedagogy helped improve the students’ ability to analyze and solve problems.

Second, the drawing process utilized by the engineering design-based pedagogy allowed the students to identify key factors for solving the problems. For example, as shown in the top three design cards in Figure 5, every solution in the pulley activity was a little abstract, and the students were almost unable to discover the key points of the problem. In the next gear activity, as shown in the bottom three design cards, the students grasped the key point of the problem quickly and focused on analyzing the gears. Everyone redrew their solutions because they realized that the gear was the key factor in building a faster fan. This phenomenon further verifies that an engineering design-based pedagogy using Lego-based activities plays a significant role in promoting problem solving in the decision-making strategy.

We also note an interesting discovery from the design cards and the observation material, that the pupils with well and detailed design diagrams got higher scores of the final products. In the first activity, the students in the EG did not consider carefully the problems that may appear in the construction phase, which resulted in a large discrepancy between the design implementation and the preliminary sketches. As the activities progressed, most students gradually perceived the importance of a systematic design and started to think more about the problems that may exist in the construction stage when they initially sketched the model. The students’ drawings became more thorough and their final design work more similar to their sketches. The students’ final products were positively associated with the degree of completion of their design diagram that helped the students to get a deeper understanding of the science content and better performance in solving problems.

During the drawing and scoring process, students learned how to compare the advantages and disadvantages of different solutions from multiple dimensions and to integrate different solutions into an optimal one. This enabled students to make rapid decisions from the complex and diverse solutions, which helped improve the dimension of their ability to decide the optimal solution. However, students were required to develop all the possible solutions and to build artifacts by themselves without any guidance during the process of generating solutions, building a
prototype, and testing the prototype. The lack of guidance on how to think and build led to no improvement in the problem-solving ability involving the other four dimensions (namely finding a problem, developing possible solutions, building a prototype, and testing the prototype).

<table>
<thead>
<tr>
<th>Activity</th>
<th>Solution 1</th>
<th>Solution 2</th>
<th>Solution 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulley Activity</td>
<td>![Image](153x373 to 261x661)</td>
<td>![Image](273x519 to 384x661)</td>
<td>![Image](400x373 to 508x661)</td>
</tr>
<tr>
<td>Gear Activity</td>
<td>![Image](273x373 to 386x519)</td>
<td>![Image](400x373 to 508x661)</td>
<td><img src="528x59" alt="Image" /></td>
</tr>
</tbody>
</table>

Figure 5. Design cards

During the five activities, there was no obvious difference between students in the CG in their seeking and decision-making strategies. In the cooperative activities, two working groups could not reach an agreement in the construction phase, because they did not discuss the activity properly and neglected the challenges. These working groups could not finish the work in time and teachers had to intervene and assist them. In contrast, students in the EG expended significant time seeking solutions based on the previously described systematic thinking process. This difference resulted in a relatively short construction period, which raised their confidence to solve questions and problems. Generally speaking, an engineering design-based pedagogy can help improve student problem solving.

The problem-solving abilities of different genders

It can be observed from the research results that the problem-solving ability score, using an engineering design-based pedagogical method, increased by 0.1 point on average for male participants in the CG, whereas the score decreased 8.2 points for female participants in the CG. We also observed that the scores for male participants in the EG improved by an average of 6.8 points, while the female participants’ score was reduced by 1.6 points. Therefore, the impact of an engineering design-based pedagogy was more visible on male participants than on female.

Some existing research conducted using Lego-type activities demonstrated that female participants are not particularly actively engaged in Lego building (Cavallo, 2003; Norton, 2006). Throughout classroom observation, we found that in both the control and the experimental groups, female participants at first did not appear very engaged when faced with the Lego artifact-building challenges, and the entire process in every working group was led by male students. The male students were more supportive of the female students, helped them get acquainted with the Lego bricks, and worked together with them to exchange ideas and/or build the artifacts. Especially in the discussion and analysis phase, male participants were more willing to express and share their opinions, whereas female
participants took the role of an assistant during this process. This imbalance minimized opportunities for the female participation in building and revising phases, which resulted in the lack of a sense of achievement in solving problems.

**Conclusion**

In this paper, we explored the feasibility and effectiveness of engineering design-based modeling approach within the science activities. The result of this study showed that performances of all students improved in the science learning regardless of whether they used the commonly used pedagogy or the engineering design-based pedagogy. However, students demonstrated an increase in the problem-solving ability after using engineering design-based pedagogy, whereas there was a decrease in this ability after the use of commonly used pedagogy. Students also showed significant improvement in the ability to identify optimal solutions to the problem at hand after using engineering design-based pedagogy. Furthermore, the study demonstrated that the problem-solving ability improved more in male participants than in female participants after using the engineering design-based pedagogy.

The engineering design-based pedagogy provides an effective instructional strategy in the science subject. By introducing engineering design in learning activities, teachers can follow the simplified five steps to better guide students in solving problems systematically. Students can take design cards as the scaffold materials, which can help them in identifying the problems and writing down possible solutions. Students can exchange and discuss ideas in groups to find the best solution by using appropriate evaluation criteria rubric. This approach encourages students to be more actively engaged in the scientific inquiry. This approach can also help students in solving ill-structured problems, because teachers can refer to the card records of the students’ written or drawn solutions in order to understand their individual science performance, problem-solving ability, and the ability to employ effective strategies.

The engineering design requires students to analyze problems in detail and thoroughly compare various possible solutions. Teachers can help students in this process by guiding them through specific steps to clearly understand the key points of the problems. This can lead to the improvement in students’ problem-solving ability. When students use engineering design-based pedagogy to solve problems in groups, they exchange their ideas with other students, prompting them to also express themselves. Teachers can observe this exchange and intervene in order to improve students’ thinking.

A number of further research directions have emerged out of this study. When looking at open-ended problems, it is important for students to have appropriate context for proper understanding of the breadth and the depth of the problem. If the context is too narrow, the learning conditions will be incomplete, leading to insufficient understanding. If the context is too wide, it can distract students’ attention even though it can provide more possible approaches to inquiry, influencing the accuracy of the grasp and the understanding of the problem. Thus, future research should focus on how to construct an appropriate problem context to get meaningful inquiry learning.

In the engineering design-based pedagogy, teachers have to monitor the students to be able to help them. This approach can work in small class sizes but is not scalable because it is not practical for teachers to be able to monitor large number of students at the same time. Learning analytics approaches can be explored to find patterns based on students’ actions, individual characteristics, problem-solving abilities, performances and other parameters, which can help teachers in identifying students who need immediate help, who should be given more challenging problems, and so on.

Future research directions also need to consider some of the limitations of the current study. The current study was limited to 30 students. A larger sample size would be able to better verify the reliability of that approach. The shorter duration of the current study (five weeks) did not allow for assessing the retention of skills beyond their current course. Experiment with a longer duration, preferably a longitudinal study, will help ascertain the retention of skills as well as students ability to use those skills in the different contexts. Therefore, generalizability of those skills can also be explored in future research. Finally, the student activities could also be video-recorded so that a more in-depth analysis about students’ behaviors could be conducted.
Acknowledgments

This research work is supported by National Natural Science Foundation of China (NSFC: 61075048), the Special Found for Beijing Common Construction Project, and the Fundamental Research Funds for the Central Universities (SKZZY2014094).

References


Improving Pupils’ Mathematical Communication Abilities through Computer-Supported Reciprocal Peer Tutoring

Euphony F. Y. Yang¹, Ben Chang²*, Hercy N. H. Cheng³ and Tak-Wai Chan¹

¹Department of Network Learning Technology, National Central University, Jhongli City Taoyuan County, Taiwan // ²Graduate Institute of Learning and Instruction, National Central University, Jhongli City, Taoyuan County, Taiwan // ³National Engineering Research Center for E-Learning, Central China Normal University, Wuhan, China // tnyfuyds@gmail.com // bchang.tw@gmail.com // hercycheng.tw@gmail.com // chan@cl.ncu.edu.tw

*Corresponding author

(Submitted June 10, 2015; Revised October 19, 2015; Accepted December 22, 2015)

ABSTRACT

This study examined how to foster pupils’ mathematical communication abilities by using tablet PCs. Students were encouraged to generate math creations (including mathematical representation, solution, and solution explanation of word problems) as their teaching materials and reciprocally tutor classmates to increase opportunities for mathematical communication during a semester. A reciprocal peer-tutoring-enhanced mathematical communication system was designed for supporting students’ math creations and reciprocal peer-tutoring activities. An experiment involving 51 second-graders was conducted to evaluate their improvement in mathematical communication abilities. While the control group received one-to-one self-learning mathematical materials and teacher-led instruction, the experimental group was engaged in computer-supported reciprocal peer tutoring in the same environment with the same materials. Both groups were evaluated by using a mathematical communication ability assessment. The result indicated that the experimental group outperformed the control group in the assessment. Additionally, math creations were analyzed for assessing students’ formative development. The results showed that students’ math creations became clearer and more efficient. In other words, their mathematical representations and solution explanations became more accurate after the learning activity.

Keywords

Mathematical communication ability, Reciprocal peer tutoring, Math creation

Introduction

Mathematical communication, a fundamental mathematics educational objective that involves cognitive and social activities (Baroody & Ginsburg, 1990), is used to engage students in communicative situations for increasing learning interaction with others to obtain mutual mathematical ideas (Silver & Smith, 1996), share mathematical thoughts, develop mathematical concepts and strategies, and reflect on their current mathematical understanding (Whitin & Whitin, 2000; Cooke & Buchholz, 2005). Mathematical communication abilities also include expressing mathematical thought by using mathematical language clearly, precisely, and succinctly (National Council of Teachers of Mathematics, 2000); understanding others’ mathematical equations and concepts (Lin & Lee, 2004; Lin, Shann, & Lin, 2008); and evaluating others’ mathematical concepts by, for example, asking meaningful questions and explaining the reasons for others’ incorrect mathematical thought (Lin & Lee, 2004).

Various means of fostering mathematical communication abilities have been proposed. For example, Baroody and Ginsburg (1990) suggested that students should communicate mathematical ideas through representing, listening, discussing, reading, and writing. Cobb, Boufi, McClain, and Whitenack (1997) also claimed that students’ mathematical discourse in classrooms can support their conceptual development, while Shimizu and Lambdin (1997) revealed that students who write about the thinking process of their solutions can organize complex thoughts and evaluate their own opinions. Similarly, Steele and Arth (1998) argued that reflecting on how to solve a problem by writing their solutions can facilitate students to explain their thinking process more clearly, thereby benefitting themselves in learning mathematics and learning to communicate mathematically by constructing mathematical artifacts as well as developing and evaluating mathematical arguments (National Council of Teachers of Mathematics, 2000). Similarly, students can adopt written mathematical communication by using text, figures, tables, pictures, diagrams, or mathematical symbols to provide critical evidence of their mathematics ideas and concepts (Mooney, Hansen, Ferrie, Fox, & Wrathmell, 2012; Whitin & Whitin, 2000). Additionally, students can learn mathematics by observing, interacting with, and manipulating physical objects and the representations of objects and
concepts (Sedig, 2008). Besides, students may support their claims by describing and recognizing patterns, generalizing rules, and using various types of representations (Moschkovich, 2012). These studies have demonstrated various methods for expressing or explaining mathematical ideas by writing or creating mathematical materials. The content of mathematical communication can be treated as an artifact of reflection, refinement, discussion, and modification. Therefore, mathematical communication ability should be fostered in students by simultaneously training their oral expression and various mathematical representations for explaining their understanding of mathematical ideas and strategies concretely as well as sharing their work with one another (Dacey & Eston, 2002).

To improve students’ mathematical communication ability, several studies have shown the potential of using computers. For instance, Koedinger, Anderson, Hadley, and Mark (1997) built Practical Algebra Tutor (PAT) system to engage students in investigating real-world problems and using algebraic tools to generate multiple representations (tables, graphs, and symbols) for solving algebra problems and communicating results. Stahl (2009) designed a Virtual Math Team (VMT) system for students’ exploring and discussing mathematical topics with peers for improving the quality of mathematical conception through online mathematical text-based chatting. Furthermore, Tsuei (2012) adopted G-Math, a synchronous peer-tutoring system, for pupils to discuss mathematical word problem solving via a sharing mechanism, which takes advantage of the availability of students’ works to communication and obtain mutual perspectives. Additionally, Bruce, McPherson, Sabeti, and Flynn (2011) allowed students to express their mathematical thought and learn their strengths from peers’ feedback during discussions.

Most importantly, the aforementioned studies, which allow students to interact with computers in pairs or small groups, have demonstrated positive learning effects (Johnson & Johnson, 1999). Previous studies have also indicated that peer tutoring may be a useful approach to facilitate students’ abilities of helping each other (Webb & Mastergeorge, 2003; Walker, Rummel, & Koedinger, 2011), potentially increasing their mathematical communication. Furthermore, Sorsana (2005) argues that peers’ mutual dialogues may benefit students’ performance more than a teacher’s direct instruction can because students as tutors can instruct their own concepts to peers as tutees. Whereas the tutor learns by doing and teaching, the tutee learns by observing, analyzing, and offering performance-related feedbacks (Topping, 2005; Berghmans, Neckebroeck, Dochy, & Struyve, 2013).

Because playing the role of either tutors or tutees has different learning effects, reciprocal peer tutoring, in which a pair of students interchangeably play both roles (Pigott, Fantuzzo, & Clement, 1986), may be more effective as students mutually explain and convey their opinions, thoughts, and strategies (Brendefur & Frykholm, 2000). For doing so, students have to prepare themselves for instruction, evaluation, and reinforcement, thereby creating mutual assistance and social support (Fantuzzo, King, & Heller, 1992). In this vein, reciprocal roles are intended to promote mutuality in the tutoring process and provide equivalent opportunities. Furthermore, both roles can be engaged in various cognitive and metacognitive activities by using mathematical language (De Backer, Van Keer, & Valcke, 2012). For example, because students need to provide immediate learning information for a partner, they have to disclose their thoughts to each other. Moreover, they have to understand classmates’ thought processes as well as the strengths and weaknesses of the proposed explanation, thereby providing additional feedback (Mosston & Ashworth, 2002). Due to questioning, explaining, and self-monitoring of learning, both students may benefit from providing help through teaching each other (King, Staffieri, & Adelgais, 1998).

During reciprocal peer tutoring, tutors usually start from expressing their ideas to peers, so that expressing coherent and organized concepts is a determinant of students’ mathematical communication and learning. The reasons are as follows. First, expressing one’s concepts is a direct result of students engaging in mathematical communication activities, such as explaining, manipulating different representations, questioning, answering, and correcting others’ errors (King, 1998). Second, students can rehearse their knowledge, integrate prior knowledge into new knowledge, and generate new ideas. Third, to produce relevant explanations, students have to prioritize the information and decide which concepts are most related to core topics then rearrange conceptual connections (Chi, Roy, & Hausmann, 2008). Fourth, students may monitor and reflect knowledge building of expressing ideas that help them evaluate the breadth and depth of their own knowledge and improve incorrect or insufficient ideas (Roscoe & Chi, 2007; 2008). Fifth, students should carefully break down examples to many steps and linking them to underlying principles thereby gaining a deeper understanding (Atkinson, Renkl, & Merrill, 2003). Therefore, to train students’ mathematical expression, facilitating students’ mathematical communication abilities and learning performance is necessary.
Among many strategies for enhancing students’ mathematical expression, asking students to generate explanations for others may be a better one (Fiorella & Mayer, 2014). Previous studies have indicated that students who can explain their solution steps of word problems are more successful in transferring knowledge (Aleven & Koedinger, 2002). Cox (1999) also notes that multiple representations can benefit students’ expression because higher level of cognitive representation is essential for advanced problem solving skills and communication. Thereby, enhancing students’ mathematical communication abilities can be achieved by asking students to integrate different representational forms to solve word problems and to communicate with others. In a sense, the multiple representations are also considered as a communication tool. Based on these reasons, the aim of this study is to explore the students’ mathematical teaching material construction effects. The mathematical teaching material used for reciprocal peer tutoring, named math creation, includes multiple representations. The math creation activities emphasize students’ mathematical problem solving and communication abilities, which require students to clarify how and why they know, and what they already know about the problems (Banger-Drowns, Hurley, & Wilkinson, 2004) via drawing representations and writing explanations. In addition, previous studies have indicated that computer-supported reciprocal peer tutoring may be an adequate and efficient learning approach to improve students’ mathematical communication ability. Therefore, the research questions of this study are: (1) Does a reciprocal peer-tutoring-enhanced mathematical communication (RPTMC) activity improve students’ mathematical communication ability? (2) Does students’ math creation in RPTMC enhance their performance in expressing their respective mathematical concepts?

RPTMC activity flow

To improve students’ mathematical communication ability, a reciprocal peer-tutoring-enhanced mathematical communication activity flow was designed and implemented. Figure 1 illustrates the details of the activity flow. Before this activity, the researchers described the purpose and procedure of RPTMC activity and system functions to both teacher and students. Later on, every two students were paired as a mathematical communication group. The learning activity involved four sub-activities: creating, reciprocal peer tutoring, revising, and staging. The four sub-activities are described as follows.

![Figure 1. Paired mathematical communication activity flow](image)

Creating

The sub-activity of creating, which required students to prepare tutoring materials, math creations, involved four steps: understanding the problem, drawing a representation, writing a solution, and explaining the solution. These
steps were designed according to Polya’s findings (1973) about problem solving, i.e., understanding the problem, devising a plan, carrying out the plan, and looking back. The following steps were implemented in the Sketch Board Zone (Figure 3) to assist students in developing a math creation:

- **Understanding the problem:** Students read the word problem on their own tablet PCs and discussed the solution with their peers to understand the conditions given and the problem asked.
- **Drawing a representation:** Students used words, symbols, models, and manipulative materials as their mathematical representations to devise a plan as well as to convey their ideas and communicate information.
- **Writing a solution:** Students wrote their mathematical equations for solving the problem.
- **Explaining the solution:** Students reflected on how and why they had solved the problem and explained their solution in writing. Because students may need guidance in learning how to express their mathematical concepts before they could write a complete sentence explaining their solutions, a text-based scaffold was provided (see Figure 2).

**Reciprocal peer tutoring**

In the second sub-activity, paired students sat together to reciprocally teach their math creations. One student, who played the role of a tutor, taught his/her peer why and how to solve the word problem by displaying his/her math creation in the Sharing Zone (Figure 3), while the other student, who played a tutee, received instruction with the tutor’s math creation on his/her own tablet PC. Subsequently, the tutee had to ask the tutor questions about the solution strategy. The paired students then switched roles and continued the sub-activity. Figure 2 shows examples of the paired students’ math creations.

**Revising**

In this sub-activity, the students had to revise their math creations based on peer feedback in the previous sub-activity for improving the correctness and clarity of their own math creations. Revising also served as a time for self-reflection and preparation for the next sub-activity of staging. Meanwhile, the teacher monitored the students’ math creations and helped their revision.

**Staging**

Finally, the teacher encouraged the students in each group to display their math creations to the whole class. As their practice in the second sub-activity, they had to explain their solutions with their representations to the audiences. Then they had to answer questions asked by the audiences. In the end, the teacher used students’ works to demonstrate how to explain the mathematical concepts and to clarify some mistakes made by students for preventing similar ones next time. Moreover, the teacher may ask some relevant questions to promote students’ thinking for communicating their own mathematical concepts and thinking with others.
System design

To explore students’ mathematical communication abilities improvement in a one-to-one learning environment (Chan et al., 2006), i.e., one student to one computer, a mathematical communication system is developed. This system provides representational tools for students to construct math creations effectively in a computer-based mathematics learning environment (Hwang, Su, Huang, & Dong, 2009; Sedig, 2008) and express personal mathematical concepts as the math creation of teaching materials to tutor each other. The RPTMC includes two major functions: Sketch Board (Figure 3) designed for constructing mathematics creations, and Sharing Zone for displaying and sharing students’ mathematics creations.

Sketch board

The Sketch Board is designed for students to express and explain mathematical concepts by drawing and writing. In the zone, students may complete a math creation individually in the sub-activity of creating. The system also provides a mathematical component library, which includes coins, building blocks, number lines, etc., for students to construct their math creations in order to concretely illustrate the problem-solving procedures and communicate mathematical concepts with others.

Additionally, the Sketch Board has a text-based scaffolding to help students get familiar with solution explanation. In the initial activities, students were scaffolded to think relevant mathematical concepts by completing the keywords of word problems. Students may thus learn how to explain their mathematical ideas to others by imitating similar explanatory patterns. Students started from filling in one-blank sentence, later on filling in more blanks, and after several activities they should write a complete sentence by using conjunction scaffolding. The format of the conjunction scaffolding likes: “[First, I used] one calculating method, [because] (I) wanted to... [Then, I used] another calculating method, [because]...”

Sharing zone

The Sharing Zone is designed to facilitate students to display their math creations easily and instantly. Teachers and students can select a classmate’s ID on their own, and then the students’ math creations are displayed on the screen for them to observe. This function assists students in taking advantage of their own tablet PCs to view each other’s work, enabling them to tutor each other reciprocally for training mathematical oral communication and to get more perspectives and suggestions. This zone also helps students learn various mathematical expressions from others’ math creations and return to modify theirs. In addition, through the Sharing Zone, teachers are able to monitor students’ progress, analyze students’ problems, and examine their knowledge status for subsequently providing real-time feedback and recommend revisions.
Besides, the Sharing Zone provides the enquiring and testing questions sheet revised from Mason (2010) to prompt students’ mathematical thought and ask questions to help peers mutually examine the correctness of their mathematical representation, solution, and solution explanation. Enquiring questions may guide learners to understand and ask questions about every meaning from several divided problem-solving steps of representation and solution. If students do not have their own questions, they can adopt the question sheet to participate in the reciprocal peer tutoring. As well, the enquiring questions can also be regarded as hints to remind students what and how to express their individual concepts to a peer or to others. Furthermore, the testing questions sheet is used for students to ask advanced solution explanation in case that the explanations of peers are unclear or doubtable (see Table 1).

The following example demonstrates how the students participate in the RPTMC activity. At the Creating Stage listed in Figure 1, each student receives a mathematical word problem on the Sketch Board delivered by the RPTMC system. At this stage, the students start to create their own teaching materials, try to understand the problem, then construct their math creations (representation, solution, and solution explanation). The student also has a text-based scaffolding to help him/her develop solution explanation. Once the students have their own math creations, the students practise reciprocal peer tutoring via the Sharing Zone. At the stage of reciprocal peer tutoring, the student who plays the tutor introduces his/her solution processes to the student who plays the tutee. The tutee then ask the tutor questions either based on the enquiring questions listed in Table 1 or the questions generated by himself/herself. After the tutor answers the questions proposed by the tutee, the tutee is assigned a set of testing questions to check the correction of the math creations constructed by the tutor. The paired students exchange their roles once this sub-activity is completed. The following is the revision stage, in which the students revise their math creations according the feedbacks generated in above sub-activities. At the final stage, staging, the teacher selects one or two students from each group to illustrate their math creations for the whole class. And then, the teacher gives some comments to those students’ math creations and makes a conclusion.

<table>
<thead>
<tr>
<th>Enquiring questions</th>
<th>Testing questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) What does this problem ask for?</td>
<td>(1) If there is something unclear or incorrect in the Instructor’s illustration, please question him/her.</td>
</tr>
<tr>
<td>(2) How do you illustrate six boxes of milk, with each box having twelve bottles of milk?</td>
<td>(2) Is his/her solution correct? If there is something wrong, please tell him/her the correct solutions.</td>
</tr>
<tr>
<td>(3) How do you illustrate the milk drunk?</td>
<td>(3) Is his/her solution explanation correct? If there is something wrong, please tell him/her how to modify it.</td>
</tr>
<tr>
<td>(4) How do you illustrate the milk left?</td>
<td></td>
</tr>
<tr>
<td>(5) What does each calculation mean?</td>
<td></td>
</tr>
</tbody>
</table>

**Evaluation**

**Setting**

For answering the research questions, an experiment was conducted. Fifty-one second-graders (aged eight to nine years) from two classes at a primary school in Taiwan were involved. The school had established a one-to-one learning environment in their math program. These students learned math by using one-to-one technology in every formal math course since they were in the first grade; thus, they had sufficient computer operation abilities to participate in the RPTMC activity. Specifically, the activities were conducted 13 times over an entire semester. Each time took 80 minutes each week. In this study, one class was assigned to be the experimental group (25 students, 13 boys and 12 girls), and the other one served as the control group (26 students, 14 boys and 12 girls).

The control variable was their daily learning approach. In other words, both groups had the same mathematics learning time in the same one-to-one self-learning mathematics environment. However, the control group practised mathematics by teacher-led instruction for solving various word problems, while the experimental group participated in the RPTMC activity to solve related word problems chosen by the teacher and researchers. For evaluating students’ mathematical communication ability, the pretest and posttest were conducted before and after the experiment, respectively. Each test took 40 minutes.
Pretest and posttest: Mathematical communication ability assessments

For the first research question, the mathematical communication assessment was used to assess the students’ mathematical communication abilities in terms of three sub-abilities: expressing their respective mathematical concepts, understanding others’ mathematical equations, and comprehending others’ mathematical thought (Lin & Lee, 2004). The assessment was consisted of multi-step word problems, including continuous addition, mixed addition and subtraction, mixed addition and multiplication, and mixed subtraction and multiplication. The pretest and posttest used the parallel problems all within the 2nd-grade mathematical curriculum (see Figure 4). These problems were collaboratively designed and developed by two educational technology experts and an elementary school teacher who has ten years teaching experience. Each question represented one mathematical communication sub-ability, and each sub-ability included two to three criteria to test different evaluative approaches.

<table>
<thead>
<tr>
<th>Original assessment material</th>
<th>Translated assessment material</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. Understand the meaning of a mathematics problem: What does the problem ask for?</td>
</tr>
<tr>
<td></td>
<td>2. Use mathematical symbols to solve the problem: Please list the equation to solve the problem.</td>
</tr>
<tr>
<td></td>
<td>3. Apply mathematics language and representations to explain mathematical concepts: Please explain why you calculated in that way.</td>
</tr>
</tbody>
</table>

Figure 4. Example of mathematical communication assessment

Evaluation criteria

Table 2 lists the detailed evaluation criteria for mathematical communication abilities (Lin & Lee, 2004). The criteria and scores were devised through discussions with three educational assessment experts and seven schoolteachers.

Table 2. Evaluation criteria for mathematical communication abilities

<table>
<thead>
<tr>
<th>Sub-ability</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Express individual</td>
<td>(1) Understand the meaning of a mathematics problem (multiple choice, 1 point)</td>
</tr>
<tr>
<td>mathematical concepts</td>
<td>(2) Use mathematical symbols to solve the problem (word problem, 2 points)</td>
</tr>
<tr>
<td>(8 points)</td>
<td>(3) Apply mathematics language and representations to explain mathematical concepts (explanation problem, 5 points)</td>
</tr>
<tr>
<td>Comprehend others’</td>
<td>(1) Understand others’ mathematical equations and evaluate the correctness (true/false question, 1 point)</td>
</tr>
<tr>
<td>mathematical equations</td>
<td>(2) Provide meaningful explanations for correct equations or explain reasons for incorrect equations to others (explanation problem, 5 points)</td>
</tr>
<tr>
<td>(6 points)</td>
<td></td>
</tr>
<tr>
<td>Comprehend others’</td>
<td>(1) Realize others’ mathematical thought and evaluate the correctness (true/false question, 1 point)</td>
</tr>
<tr>
<td>mathematical thought</td>
<td>(2) Use mathematical language to transform others’ mathematical thought into mathematical equations (if true) or ask meaningful questions and explain reasons for the incorrect thought of others (if false) (explanation problem, 5 points)</td>
</tr>
<tr>
<td>(6 points)</td>
<td></td>
</tr>
</tbody>
</table>
To ensure the reliability of the assessment, two raters evaluated the assessment independently. The inter-rater reliability of the pretest was 0.912, \( p < .05 \), and that of the posttest was 0.905, \( p < .05 \).

**Results**

Table 3 presents the averages and standard deviations of the pretests and posttests of both groups. A \( t \) test showed no significant differences between the pretest scores of the two groups \( [t(49) = .48, SE = 1.23, p = .80] \), indicating both groups had equivalent mathematical communication ability. Besides, the pretest scores of both groups were only around half of the total score, suggesting that students needed more trainings to enhance their mathematical communication abilities, especially those for comprehending others’ mathematical thought.

<table>
<thead>
<tr>
<th>Mathematical communication ability</th>
<th>Experimental group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((n = 25))</td>
<td>((n = 26))</td>
</tr>
<tr>
<td>Express students’ respective concepts (8 points)</td>
<td>Mean: 4.28, SD: 1.34</td>
<td>Mean: 3.96, SD: 1.25</td>
</tr>
<tr>
<td>Understand others’ mathematical equations (6 points)</td>
<td>Mean: 3.00, SD: 2.55</td>
<td>Mean: 3.65, SD: 2.86</td>
</tr>
<tr>
<td>Comprehend others’ mathematical thought (6 points)</td>
<td>Mean: 2.44, SD: 1.83</td>
<td>Mean: 2.69, SD: 1.49</td>
</tr>
<tr>
<td>Total</td>
<td>Mean: 9.72, SD: 4.62</td>
<td>Mean: 10.30, SD: 4.18</td>
</tr>
</tbody>
</table>

A two-way ANOVA (\( \alpha = 0.05 \) significance level) was performed to analyze the effect of RPTMC activity on the total scores of the mathematical communication assessments. The variance homogeneity tests between group linearity were not significant (\( \rho = .40 > .05 \)). The results revealed a significant interaction between the groups and the pre/posttest \( [F(1, 49) = 23.76, MSE = 271.80, p = .00; \eta^2 = .15] \). Further analysis indicated that the experimental group showed significant improvement in total score \( [t(24) = -7.64, SE = .89, p = .00] \), whereas the control group showed no significant improvement \( [t(25) = -2.27, SE = 1.00, p = .79] \). In other words, students engaged in RPTMC activities may develop mathematical communication abilities superior to those who received one-to-one self-paced learning and teacher-led instruction. The reason may be that the RPTMC activity could promote student use of mathematical language for expressing mathematical concepts and sharing math creations in the Sharing Zone; thus, the students may learn various forms of expression for communicating their solutions with tutored peers. In addition, when monitoring the students’ performance of math creation, the teacher asked some students to provide more detailed explanations of their solutions and drawings, and thereby develop their mathematical communication abilities after several activities. Contrarily, the control group had fewer opportunities to express their mathematical ideas about solving their problems, which resulted in lower performance on the posttest.

In addition, the experimental group showed a significant improvement on the sub-abilities of expressing their respective mathematical concepts \( [t(24) = -7.96, SE = .36, p < .01] \), understanding others’ mathematical equations \( [t(24) = -2.21, SE = .51, p < .05] \), and also comprehending others’ mathematical thought \( [t(24) = -6.16, SE = .45, p < .001] \). The results suggested that all three sub-abilities were fostered by training in the RPTMC activities. In other words, sufficient practices on finding solutions and explaining them through writing/drawing and verbal forms may assist students in expressing their own mathematical concepts and understanding others’ mathematical thought.

**Math creations: Expressing individual mathematical concepts**

For the second question, this study further analyzed the performance of students’ math creations in the experimental group. Besides, the researchers also interview the class teacher and students about students’ learning differences and difficulties before and after the RPTMC activities.
Evaluation criteria

The scores of expressing their respective mathematical ideas were calculated according to the criteria in Table 4, in which the correctness and completeness of their creations were considered. Each criterion was scored as five points. The criteria were designed to evaluate how well the students (1) applied various mathematical representations for explaining their respective mathematical concepts, (2) used mathematical symbols to solve problems, and (3) explained the solutions critically.

To observe the development of the students’ math creations, the 13 weeks of math creations were divided into three stages according to the mathematical learning content: The initial stage (weeks 1 to 2), in which two-digit subtraction and two-digit by one-digit multiplication were learned; the middle stage (weeks 3 to 7), in which continuous addition, continuous subtraction, and mix of addition and subtraction; and the final stage (weeks 8 to 13), in which mix of addition, subtraction, and subtraction were learned.

Table 4. Evaluation criteria for the math creations

<table>
<thead>
<tr>
<th>Score</th>
<th>Mathematical representation</th>
<th>Solution</th>
<th>Solution explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect or not related</td>
<td>None or drawing something hard to understood</td>
<td>Incorrect equation</td>
<td>Explaining incorrectly</td>
</tr>
<tr>
<td>(1 point)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct but without a critical concept</td>
<td>Drawing objects without relevant marks</td>
<td>Only part of a correct equation</td>
<td>Explaining irrelevantly</td>
</tr>
<tr>
<td>(2 points)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct but focused on the calculation</td>
<td>Drawing objects as an equation without their amounts or meanings</td>
<td>Correct equation with an incorrect calculation</td>
<td>Explaining only the calculation</td>
</tr>
<tr>
<td>(3 points)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct but incomplete</td>
<td>Drawing objects as an equation with their amounts or meanings</td>
<td>Correct equation without an answer or with an incorrect answer</td>
<td>Explaining correctly without essential concepts</td>
</tr>
<tr>
<td>(4 points)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct and complete</td>
<td>Drawing spatial relations between the objects of the problem and the solution strategy</td>
<td>Correct equation with a correct answer</td>
<td>Fully elaborating solution strategy with spatial relations among the various objects of the problem</td>
</tr>
<tr>
<td>(5 points)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results

The average scores for the quality of the students’ creations at the three stages are listed in Table 5. Besides, to illustrate the math creations, a typical example of one’s math creations selected based on the average scores of the three stages is shown in Figure 5.

Table 5. Three stages of expressing students’ respective mathematical ideas

<table>
<thead>
<tr>
<th>Stage</th>
<th>Mathematical representation</th>
<th>Solution</th>
<th>Solution explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>1.86 (1.54)</td>
<td>3.66 (2.14)</td>
<td>—</td>
</tr>
<tr>
<td>Middle</td>
<td>2.04 (1.01)</td>
<td>4.76 (0.90)</td>
<td>2.39 (0.91)</td>
</tr>
<tr>
<td>Final</td>
<td>3.29 (1.60)</td>
<td>4.94 (0.44)</td>
<td>3.06 (1.66)</td>
</tr>
</tbody>
</table>

At the initial stage, the students were arranged to get familiar with the system and the rules for math creation, so they did not need to write the solution explanation. Half of the math creations were basically correct but without essential concepts. Specifically, they used various math components as mathematical representation for expressing the amounts of the variables in the problems. Some students’ mathematical representations were not comprehensible, whereas the other students’ presented only partial equations for the solutions. These creations showed that most students lacked sufficient understanding to transform the details of the problem into their representations; therefore they could not apply appropriate mathematical representations to express their mathematical concepts at this stage. Besides, in the interview some students regarded the mathematical representation as the solution, and thereby they did not write the solutions.
At the middle stage, most students started to use coins or blocks provided by the system to represent the equations for their solutions and calculative processes. Specifically, their mathematical representations showed an amount and order similar to their equations instead of solution strategies at this stage. In addition, when explaining solutions with scaffolds, many students did not know the answers and left the blanks, thereby resulting in a relatively low average score. By contrast, the average solution score increased, because the students were more familiar with the differences between mathematical representation and the solution, and recognized the mathematics problem meaning. Hence, they were more capable of using correct mathematical symbols to complete the solutions.

![Initial stage](image1)

![Middle stage](image2)

![Final stage](image3)

*Figure 5. Example of students’ math creations at the three stages*

At the final stage, the teacher reflected that “students began to use various explanatory methods and also examined the process and reasoned one step after another,” indicating that the students attempted to apply various mathematical representations and reasons for expressing their solution ideas and could think how to solve a word problem logically. The students started to draw squares or circles with numbers as their mathematical representations, and then their scores increased. As a result, from tutors’ mathematical representations, tutees could better observe the solution strategies and the relationship of every condition to the assigned problem. Regarding the solutions, although they were assigned more difficult problems, the average score increased. In addition, the average of the solutions explanations also slightly increased because they were requested to write a complete explanation. However, most students merely wrote arithmetic operations instead of sentence. There were seldom students expressed their reasons for representing the meaning of each number. The results indicated that the students became more capable of using mathematical representations and make equations, but they still needed more practice in using mathematical language for explaining their solutions.

In sum, the students exhibited considerable difficulty in drawing their mathematical representations and solving problems at the initial stage. Meanwhile, the students became increasingly capable of constructing their math creations as mathematical communication artifacts at the middle stage, and at the final stage, their math creations demonstrated more sophisticated representations and explanations.

**Conclusions**

Mathematical communication emphasizes people’s interactions and exchanges of mathematical ideas, which is a crucial ability for students in expressing their respective mathematical concepts, comprehending and evaluating other students’ mathematical equations and thoughts. Therefore, this study aims at enhancing pupils’ mathematical communication abilities. To assist students in constructing math creations for reciprocally tutoring their classmates (e.g., to convey mathematical concepts and understand peers’ mathematical ideas), an RPTMC system was designed and implemented. For evaluating mathematical communication ability, both math creation and mathematical communication assessments were applied.

The evaluation showed that the experimental group students’ mathematical communication abilities had improved. More specifically, the RPTMC activity could facilitate pupils’ three sub-abilities: (1) to express their respective mathematical concepts, (2) to understand others’ mathematical equations, and (3) to comprehend others’ mathematical thought. The findings provided insights into the manner of their progress. Regarding comprehending others’ mathematical thought, the students improved significantly in the assessments after sufficient RPTMC
activities. At the pretest stage, because most students were unable to communicate math with others, they wrote only equations or simply repeated the statements of a problem. However, when playing the role of tutees in the reciprocal peer tutoring, they had to use appropriate mathematical language to transform their partners’ mathematical thought into mathematical equations. Such a collaborative and shared activity may allow them to realize peers’ mathematical thought, evaluate the correctness, and further provide meaningful questions and explanations for the peer’s incorrect thought. The improvement on the assessment indicated those students’ experiences of the RPTMC activity may enable them to transfer other’s knowledge to self-expression for more complete and understandable creations. Therefore, in the posttest, most students understood the meaning of mathematical problems and knew which mathematical symbols should be used in solving their math problems. Most importantly, their thoughts were sharpened by connecting to various presentation techniques to achieve the effect of mathematical communication.

However, although the students could express their mathematical solutions and explain others’ mathematical equations and thoughts with complete sentence in their posttest, most students did not write complete sentences to explain their solutions in their math creations. Such phenomena reflected that students might have learned how to express their mathematical concepts, but some factors still limited their performance in the math creations. For instance, students were used to typing rather than handwriting words on tablet PCs. Even if they knew how to explain their solutions, they did not do so. In addition, as some math creations without explanation were shown in the Sharing Zone, more and more students imitated such creations and eventually formed an improper classroom atmosphere. This fact revealed that the Sharing Zone can facilitate the students’ mathematical communication, but some students might misuse the function. Thereby when an instructor finds learning problems spread by the learning system, the instructor’s immediate correction is necessary.

Regarding the sub-ability effects of understanding other students’ mathematical equations, the experiment demonstrated significant learning gains, possibly because the control group was unfamiliar with understanding other students’ mathematical equations. In other words, teacher-led instruction with one-to-one self-learning lacked opportunities for mathematical communication with their peers possibly prevents them from understanding other’s equations.

Overall, the RPTMC activity not only improved students’ mathematical communication ability but also provided instructors and system designers with the developmental process of mathematical communication ability. By doing so, instructors are more capable of analyzing the quality and improvement of students’ mathematical communication ability (i.e., depicting the quantitative relationships of mathematical problems, completing mathematical representation, and explaining solutions). By using the Sharing Zone function of the RPTMC system, instructors could have the students’ learning statuses instantly. However, system designers still need to improve the system by developing categorization or grouping functions for instructors to assess and monitor the students’ learning performance of each group. In addition, once students complete the RPTMC activity, the system should provide students with the opportunity to exchange suggestions for mutual math creations, thereby enhancing mathematical communication further. The system and learning approach demonstrated in this study is only one of the methods for stimulating mathematical communication abilities. Further class computer-supported reciprocal peer tutoring methods investigations are needed for realizing computer-supported reciprocal peer tutoring learning approaches in class.

Acknowledgements

This project was supported in part by the Ministry of Science and Technology of Taiwan, Republic of China under the grant MOST 104-2511-S-008-009-MY3, and Research Center for Science and Technology for Learning, National Central University, Taiwan. The authors particularly thank the teacher and fifty-one second graders for their participation.

References


ABSTRACT

Seamless language learning promises to be an effective learning approach that addresses the limitations of classroom-only language learning. It leverages mobile technologies to facilitate holistic and perpetual learning experiences that bridge different locations, times, technologies or social settings. Despite the emergence of studies on seamless language learning, there is a lack of instruments specifically designed to measure students’ motivation and learning strategies in such technology-enhanced learning environments. This study aimed to develop and validate an instrument, namely, the Mobile-Assisted Seamless Chinese Learning Questionnaire (MSCLQ), to measure students’ motivation and the learning strategies they use in a seamless learning environment. Confirmatory factor analyses (CFA) were conducted to validate the psychometric properties of the instrument. Subsequently, structural equation modeling (SEM) was conducted to examine how students’ intrinsic value and self-efficacy for learning Chinese language predicts their perception of the various dimensions of meaningful seamless learning. The findings indicated that the MSCLQ is a valid and reliable questionnaire. Moreover, it was also found that students’ motivation predicted the learning strategies they used in a seamless learning environment. Implications are discussed.

Keywords

Mobile-assisted, Seamless learning, Mobile-assisted language learning, Motivation, Learning perceptions and strategies

Introduction

In recent years, language learning scholars (e.g., Little, 2007; Tedick & Walker, 1995) have become cognizant of the limitations of classroom-only language learning. Salient criticisms of classroom-only language learning include de-contextualization of the learning material and processes, and the lack of autonomous learning and authentic social interactions. These issues typically undermine learners’ holistic language development, especially for real-life interactions. In turn, there is an emerging consensus that the language learning process could be extended beyond the classroom, and that learners could be provided with opportunities to use the target language meaningfully and extensively in their daily life (Benson, 2013; Canagarajah & Wurr, 2011).

The advancement of mobile technologies could potentially address the aforementioned problems that bedevil classroom-only language learning. Mobile technologies offer pedagogical affordances that educators can leverage to promote meaningful learning among learners, both inside and out of the classroom. The most salient pedagogical affordance of mobile devices is that they allow learning to happen in the real world, which contributes to the authenticity of the learning and situated meaning making (Pachler, 2010).

Armed with their mobile devices, learners can actively construct digital artifacts whenever and wherever they have the intention to learn. Subsequently, they can upload the constructed artifacts for sharing, peer critiquing and co-construction, thereby making learning more collaborative. Researchers have characterized this form of learning wherein there is 24/7 access to at least one mobile device (1:1) as seamless learning (Chan et al., 2006). Since 2006, emerging designs of seamless learning that aim to create holistic and perpetual learning experiences have been reported. Formal and informal learning, individual and social, and physical and digital spaces are thus woven together with the mediation of mobile technologies (Wong & Looi, 2011; Wong, Milrad, & Specht, 2015). Among the 40 seamless learning projects identified in a recent review (Wong, Chai, & Aw, 2015), ten projects were dedicated to the design of language learning tasks in multiple settings.
Despite the emergence of studies on seamless language learning (SLL), few researchers have explicitly examined the role of motivation and learning strategies in SLL. This is an important gap given that the likelihood of success in any learning activity is largely determined by students’ motivation and learning strategies (Weinstein, Husman, & Dierking, 2000). This lack of research is partly due to the lack of appropriate measures of motivation in technology-enhanced learning environments, as well as the lack of instruments to measure the relevant learning strategies that students use when they are engaged in seamless learning. Furthermore, most SLL studies reported to date have been limited by small sample sizes (e.g., Ogata et al., 2008; Wei, 2012).

Students’ perceptions of and strategies for SLL offer valuable data that can allow researchers to assess the efficacy of their designed activities, and that can be used for further refinement of the learning designs. In addition, how students’ motivation contributes to SLL activities is also a crucial issue to consider, especially in second language acquisition (SLA) settings. Student-centered learning design is associated with learners’ motivation and autonomy (self-directedness) to learn. This research therefore aims to address the stated gap by developing and validating an instrument, namely, the Mobile-Assisted Seamless Chinese Language learning Questionnaire (MSCLQ), to measure students’ motivation and perceptions of strategies for seamless Chinese language learning through confirmatory factor analyses. Although the current questionnaire was focused on Chinese language learning, we believe that the instrument can easily be adapted for use with other languages.

Subsequently, structural equation modeling was conducted to test the hypotheses that students’ intrinsic motivation and self-efficacy in learning Chinese language contributes to the various dimensions of meaningful seamless learning.

In the following section, we first reviewed the current issues related to second language education. This was followed by a review of seamless learning and meaningful learning. Based on the review, the four dimensions of seamless Chinese learning were formulated. Finally, a review of motivation to learn, specifically intrinsic goal orientation and self-efficacy, was presented. Hypotheses regarding the relationships among these dimensions were then formulated based on the literature review.

**Literature review**

**Current issues in second language education**

A number of empirical studies (e.g., Liao & Yang, 2012; Liu, Goh, & Zhang, 2006; Plank & Condliffe, 2011) have found that many K-12 second language classroom practices typically fall short in the following aspects: (1) Incorporating an excessive amount of decontextualized materials; (2) Unbalanced instructional emphases, i.e., predominantly teacher-centric, and emphasizing language input over language output activities; (3) Disintegrated instructions of language knowledge/skills, i.e., listening, speaking, reading and writing skills are taught and practiced separately; (4) Exercising the presentation, practice, production (PPP) procedure in a linear fashion; and (5) Lack of autonomous learning and authentic social interactions. Such classroom practices are not conducive to developing learners’ communicative skills and elevating/sustaining their learning motivation.

The growing dissatisfaction with such pedagogical approaches undergirded by behaviorist and cognitivist models of language instruction has prompted scholars to investigate alternative approaches to language learning. The current theorizing which adopts sociocultural approaches to understand SLA views language production and thinking as tightly interwoven, i.e., language production mediates thinking (Lightbown & Spada, 2013). Social interaction, rather than isolated input or output, is foregrounded as the context in which language use and language learning co-occur (Min, 2006). Participation in social activities is emphasized as the primary goal, and the learning of language occurs naturally to fulfill this goal. Effective language learning is characterized by active and constructive production of thoughtful linguistic artifacts in authentic social settings (Ellis, 2000).

In addition, motivation to learn a second language has long been a perennial concern that SLA educators face (Masgoret & Gardner, 2003). In Singapore, the practice of bilingualism has promoted learning English as the first language while the ethnic mother tongues are treated as second languages. Consequently, the mother tongue language teachers, including Chinese language teachers, are facing challenges in motivating students whose proficiency levels in their mother tongues are declining (Wong, Chai, Chen, & Chin, 2013; Wong, Gao, Chai, &
Finding pedagogically engaging ways to foster students’ motivation is thus an important issue. Technology-enhanced learning that helps students connect their language learning endeavours to their daily life has been explored as one possible avenue for improving students’ learning motivation.

**Synthesis of meaningful learning, seamless learning and second language learning**

Technology-enhanced learning emphasizes students’ agency in meaning making. It is student-centric and constructivist/social constructivist oriented (Voogt, 2010). It encourages students to take charge of their own learning, especially after formal instruction has ended. However, most SLL studies have been conducted from learning technologists’ perspectives that typically foreground mobile affordances. Some SLL studies have facilitated learning through repetitive language learning activities in both formal and informal settings, across time and locations (e.g., Redd, 2011; Wei, 2012). In addition, the learning efforts in different learning spaces are disconnected from each other. Explicit efforts to bridge the activities will help to foster intentional learning across the spaces rather than the current compartmentalized practices. Another salient observation is that seven out of the ten SLL projects identified were in essence one-off studies. A more pervasive and immersive process of learning is crucial as the appropriation of language knowledge and skills is cumulative, and it has to be genuinely integrated into learners’ daily life. Furthermore, some studies (e.g., Koh, Loh, & Hong, 2013; Ogata et al., 2011; Wei, 2012) lack language learning theoretical support. In sum, it seems clear that SLL needs more theorizing and model building supported by current language learning theories.

To improve the design of SLL, with consideration of the evolution of SLA and the review of current SLL studies, the following key design principles are proposed:

- Create opportunities for authentic learning activities;
- Facilitate learners’ collaborative construction of linguistic knowledge through social interaction;
- Integrate language input with output activities for integrated skills development;
- Encourage autonomous learning among learners.

These principles of language learning can be easily synthesized with the dimensions of meaningful learning (Howland et al., 2012) to be supported by mobile technologies. These dimensions are supported by previous research (Wong, 2013; Wong, Chin, Tan, & Liu, 2010) that was only applied to a specialized domain of seamless idiom learning.

First, mobile technologies can enhance authentic learning as they enable students to go beyond the classroom into the real world to learn (Kukulska-Hulme & Traxler, 2013). Vocabulary, which includes mainly verbs, nouns, and adjectives, can be more readily found in the real world than within the classroom confines (Pavlenko, 2009). Taking photos or video clips of what is happening in the real world and using these digital resources to help students to deepen their understanding of vocabulary learned in class helps to bridge formal learning with everyday living experiences (e.g., Koh et al., 2013; Ogata et al., 2011).

Second, active and constructive learning can be actualized through students’ active construction of linguistic artifacts. Technologies afford efficient means for students to construct and mix different forms of media to represent their understanding, thus encouraging active learning mediated through the creation of artifacts (Sadik, 2008). In particular, Wong (2013) reported a study that illustrated how language inputs (i.e., vocabulary learned in the classroom) were integrated with the output activities of creating artifacts. Students in Wong’s study actively enacted scenarios for vocabulary they wanted to learn, took photos or video clips, and constructed sentences or paragraphs about the recorded enactment.

Third, the notion of intentional learning is akin to self-directed or self-regulated learning. Intentional learning is supported through students’ goal-directed and self-regulated use of the technologies to seek information related to the learning goals, organize and keep track of the information identified, and work on the information gathered (Howland, Jonassen, & Marra, 2012). Technologies provide students with ample choices of learning resources that can help them use different ways to represent what they have learned. Deliberating on various choices is likely to foster self-directedness that is essential for lifelong learning.
Fourth, online platforms are widely recognized as facilitating social interactions for collaborative learning (Kreijns, Kirschner, & Vermeulen, 2013). This is especially apt for language learning since the ultimate aim of language learning is to communicate. When students are exposed to a range of possible usage of vocabulary accompanied by illustrative photos shared by their peers, they are likely to acquire many more examples and thus become more fluent in using the vocabulary. Current evaluations conducted by Aw, Wong, Zhang, Li and Quek (2016) and Wong, Chai, King, and Liu (2015) indicate that students are able to use the vocabulary learned through SLL much better than those learned through the traditional method.

Taken together, successful seamless learners are those who engage in authentic learning, create different linguistic artifacts, are self-regulating, and are collaborative. Students who use these strategies are able to fully harness the power of the seamless learning environment.

**Research on motivation**

Motivation has been found to support student engagement and enhance the attainment of learning outcomes (Pintrich, 2003). Motivated students devote efforts to performing learning tasks, persist when they encounter problems, and regulate their learning (Masgoret & Gardner, 2003). Most importantly, studies from the past two decades have established a strong positive association between student motivation and performance (Pintrich, 2004). While there is general consensus that motivation is manifested in “the process whereby goal-directed activity is instigated and sustained” (Pintrich & Schunk, 2002, p. 5), motivation in learning is a multidimensional phenomenon (King & McInerney, 2014).

Pintrich, Smith, Garcia, and McKeachie (1991) measured six aspects of motivation from the sociocognitive framework that includes intrinsic goal orientation, extrinsic goal orientation, task value, control of learning belief, self-efficacy and test anxiety. They created a self-report instrument entitled the Motivated Strategies for Learning Questionnaire (MSLQ). This instrument consists of two parts, the first measuring the six aspects of motivation mentioned above, and the second measuring nine types of learning strategies (e.g., rehearsal, organization, elaboration, peer learning). It is one of the most widely used instruments for measuring students’ academic motivation and learning strategies (Duncan & McKeachie, 2005).

While the MSLQ was originally developed for college students, it has also been adapted for elementary students (Ocak & Yamac, 2013). In particular, Ocak and Yamac’s (2013) research was conducted in the context of a Turkish primary school for fifth graders learning mathematics. They reported that intrinsic goal orientation, task value and self-efficacy predicted students’ cognitive and metacognitive learning strategies. In Singapore, the MSLQ was adapted to study Singapore secondary school students’ motivation in learning Chinese language, known as MALLI (Motivation and Attitudes for Language Learning Inventory) (Wong et al., 2013).

The MSLQ was created in the early 1990s, and dimensions related to technology-enhanced learning were not included as part of the learning strategies subscale. In developing the MSCLQ, we replaced the learning strategies portion of the MSLQ with key seamless language learning strategies drawn from the literature review articulated above, that is, authentic learning, construction of language learning artifacts, self-regulated learning, and collaborative learning. The reason we did not use the original learning strategies question in the MSLQ was that the original subscales are more relevant for learning in traditional classroom settings that are not mediated by technology. In technology-enhanced seamless learning environments, a different set of learning strategies becomes crucial. Thus, we measure these key constructs in this paper.

**Method**

**Background and participants**

The participants in this study were 259 Primary 3 students (127 girls and 132 boys) whose ages ranged from 9-10 years from one primary school in Singapore. In general, the students’ first language is English and they possess mixed abilities in Chinese as a second language. In terms of their experience of the use of Information and Communication Technology (ICT) in learning, the students were quite well-experienced in using different software
(e.g., Microsoft Word and PowerPoint) and several e-learning portals since Primary 1, as they are part of the school structured ICT training and ICT integrated lessons. However, they had no prior experience of using smartphones for learning purposes. The school has equipped all Primary 3 students with smartphones and data connection plans. In addition, a cloud-based platform named MyCLOUD was specifically designed to facilitate language learning with the mobile devices. The intervention took place during February-November 2013. The students constitute the entire Primary 3 level of the school and were taught by six different teachers. To implement SLL, the six teachers co-designed the lesson plans with the researchers and implemented the lessons in their respective classes.

The lesson usually began with classroom teaching of a prescribed text to be read and vocabulary to be learned. Subsequently, the students used their smartphones to participate in a range of learning activities including (a) intentional selections of unfamiliar vocabulary and the checking of their meanings and examples of usage from the web as a means to promote self-directed learning, (b) taking pictures and making sentences associated with the vocabulary items learned for authentic, active and constructive learning, and (c) posting the artifacts online and writing comments for their peers’ artifacts on the platform for collaborative learning. These activities were initially modeled in the classroom. Subsequently, students were assigned vocabulary that they should use to construct sentences after class. The students were encouraged to go beyond the assigned tasks by creating more artifacts in a self-directed manner. The teachers periodically reviewed the students’ online posts and discussed their digital artifacts during class time. Around eight thousand digital artifacts (sentences, comments, or sentences with photos) were constructed during four months of engaging in the seamless Chinese learning, averaging about two digital artifacts created per student per week. These outcomes indicate that the students were familiar with the use of smartphones for seamless Chinese language learning.

**Instrument**

To explore the students’ motivation and perceptions of the seamless Chinese language learning practices, a questionnaire entitled Motivation for Seamless Chinese Learning Questionnaire (MSCLQ) was constructed. This questionnaire contains six subscales: intrinsic value (IV), self-efficacy (SE), artifact creation (AC), authentic learning (AL), self-directed learning with technology (SDT) and collaborative learning with technology (CLT).

Items for intrinsic value and self-efficacy for Chinese learning were adapted from a previous study which aimed to explore Singaporean secondary students’ motivation to learn Chinese language (Wong et al., 2013). Items for the self-directed and collaborative learning with technology were adapted from another study among Singaporean secondary students in the context of promoting self-directed and collaborative learning with technology, which is the current focus of the local Masterplan for ICT (Tan et al., 2011). Both studies validated the instruments through factor analysis. The items adopted were reviewed, contextualized and simplified with input from the teachers.

Authentic learning and artifact creation were created by the authors for this study based on the literature review. Minor changes involved adding qualifiers such as “in learning Chinese” and specifying “smart phones and computers” as the devices. Authentic learning is understood to involve connecting what is learnt to one’s daily life or to the real world (Howland et al., 2012). Artifact creation represents students’ active and constructive learning (Sadik, 2008; Wong, 2013) whereby students create digital artifacts to illustrate their understanding of the linguistic knowledge acquired. The authors brainstormed a list of possible behavior indicators that represent the two constructs with reference to the lesson activities that they co-designed with the teachers. After drafting the initial set of items, the 30-item MSCLQ was reviewed by two professors in education. The professors were asked to check if the items corresponded to the constructs we wanted to measure. Both professors, who are experienced in quantitative research, provided their feedback. The items were revised before we gave the questionnaire to the teachers to review again to check if the language and content were appropriate for Primary 3 students.

**Data collection and analysis**

To validate the instrument, the participants should have prolonged experience of SLL. Since the students did have such experience, the sampling strategy for this study is purposive sampling. All Primary 3 classes in the researched school were invited to participate. The data were collected four months after the intervention commenced. After
cleaning the raw data, 259 valid responses (97.4% of the P3 students) were keyed into the SPSS. Some students were absent, while some responses were incomplete. These cases were excluded from further analysis.

An analysis of skewness and kurtosis was first conducted to check if the data could be considered as normally distributed. The results indicate that all items were within the acceptable range of |2|, indicating that the data could be treated as normally distributed. The data were then subjected to CFA to test the construct validity of the responses to the questionnaire. After removing items with insufficient factor loadings or items that exhibited multicollinearity, the means, average variance extracted (AVE) and the composite reliability for each factor were computed. The correlations between the factors were then computed. Finally, structural equation modelling was used to test the hypotheses. The data analysis procedures were in accordance with the recommendations of Hair, Black, Babin, and Anderson (2010).

Results
Validation of the MSCLQ with CFA

The first aim of this study was to test the factor structure of the MSCLQ, which was developed to study students’ motivation and perceptions of seamless learning. The MSCLQ is comprised of six factors. The final version of the questionnaire with the factor loadings of each item is presented in Appendix 1. Four items were removed due to low factor loadings. When CFA was conducted on the remaining 26 items, a satisfactory model fit was obtained ($\chi^2 = 432.01$, $\chi^2/df = 1.54$, $p < 0.001$, RMSEA = 0.046, SRMR = 0.045, CFI = 0.96, GFI = 0.88). Based on the recommended value for instruments with 12 to 30 items for a sample of more than 250 participants ($\chi^2/df < 3$, CFI > 0.92, RMSEA < 0.07, SRMR < 0.08) (Hair et al., 2010), the model fit obtained supported the construct validity of the MSCLQ. Table 1 reports the means, SDs and the AVE and composite reliability (CR) values. The AVE values are all above the recommended value of 0.5, and the CR values are all above 0.7. These indicators attest that the MSCLQ is a reliable instrument.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>SD</th>
<th>AVE</th>
<th>CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic Value (IV)</td>
<td>3.79</td>
<td>0.97</td>
<td>0.51</td>
<td>0.75</td>
</tr>
<tr>
<td>Self-efficacy (SE)</td>
<td>3.66</td>
<td>0.98</td>
<td>0.53</td>
<td>0.85</td>
</tr>
<tr>
<td>Authentic learning (AL)</td>
<td>3.75</td>
<td>1.03</td>
<td>0.56</td>
<td>0.84</td>
</tr>
<tr>
<td>Self-directed learning with technology (SDT)</td>
<td>3.50</td>
<td>1.03</td>
<td>0.56</td>
<td>0.83</td>
</tr>
<tr>
<td>Artifact creation (AC)</td>
<td>3.65</td>
<td>1.12</td>
<td>0.56</td>
<td>0.82</td>
</tr>
<tr>
<td>Collaborative learning with ICT (CLT)</td>
<td>3.44</td>
<td>1.09</td>
<td>0.51</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Note. SD = standard deviation; AVE = average variance extracted; CR = composite reliability.

Correlations among perceptions of learning practices and motivation

Table 2 shows the correlations among the six factors in the students’ perceptions of seamless Chinese learning practices with the motivation factors. The six factors were significantly correlated with each other ($r = 0.47$ to $r = 0.76$).

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic Value (IV)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-efficacy (SE)</td>
<td>0.64*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authentic learning (AL)</td>
<td>0.66*</td>
<td>0.59*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-directed learning with technology (SDT)</td>
<td>0.59*</td>
<td>0.55**</td>
<td>0.70**</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artifact creation (AC)</td>
<td>0.51*</td>
<td>0.47**</td>
<td>0.64**</td>
<td>0.76***</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Collaborative learning with ICT (CLT)</td>
<td>0.48*</td>
<td>0.48**</td>
<td>0.63**</td>
<td>0.66**</td>
<td>0.57**</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. *p < 0.01.
Structural equation modeling

To test the hypothesis that student motivation would positively predict students' perceptions of seamless Chinese learning practices, SEM analysis with AMOS was conducted (see Table 3). The fit indices indicated that the model had a good fit to the data ($\chi^2 = 486.89$, $\chi^2/df = 1.70$, $p < 0.001$, RMSEA = 0.052, SRMR = 0.048, CFI = 0.94, GFI = 0.87). Table 3 reports the hypotheses and the estimates. Out of the eight hypotheses, four regarding intrinsic value were supported, while the four related to self-efficacy were not supported. Interestingly, self-efficacy was a significant negative predictor of the seamless learning perceptions.

Table 3. Hypothesis testing with SEM

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Unstandardized estimates</th>
<th>S.E.</th>
<th>C.R.</th>
<th>$p$</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1 IV positively predicts AC</td>
<td>4.21</td>
<td>1.32</td>
<td>3.21**</td>
<td>.001</td>
<td>Yes</td>
</tr>
<tr>
<td>H2 IV positively predicts AL</td>
<td>2.49</td>
<td>.77</td>
<td>3.25**</td>
<td>.001</td>
<td>Yes</td>
</tr>
<tr>
<td>H3 IV positively predicts SDT</td>
<td>4.28</td>
<td>1.31</td>
<td>3.28**</td>
<td>.001</td>
<td>Yes</td>
</tr>
<tr>
<td>H4 IV positively predicts CLT</td>
<td>2.74</td>
<td>.88</td>
<td>3.12**</td>
<td>.002</td>
<td>Yes</td>
</tr>
<tr>
<td>H5 SE positively predicts AC</td>
<td>-3.14</td>
<td>1.20</td>
<td>-2.61**</td>
<td>.009</td>
<td>No</td>
</tr>
<tr>
<td>H6 SE positively predicts AL</td>
<td>-1.47</td>
<td>.69</td>
<td>-2.13*</td>
<td>.034</td>
<td>No</td>
</tr>
<tr>
<td>H7 SE positively predicts SDT</td>
<td>-3.04</td>
<td>1.19</td>
<td>-2.55*</td>
<td>.011</td>
<td>No</td>
</tr>
<tr>
<td>H8 SE positively predicts CLT</td>
<td>-1.87</td>
<td>.80</td>
<td>-2.34*</td>
<td>.019</td>
<td>No</td>
</tr>
</tbody>
</table>

Note. *$p < 0.05$; **$p < 0.01$; S.E. = standard error; C.R. = critical ratio.

It was surprising to find that SE was a negative predictor in the model, while its bivariate correlation with the other outcomes was positive. This is a case of the negative suppression effect which occurs when two independent variables have a positive zero-order correlation with the dependent variable and correlate positively with each other. In such cases one of the independent variables may become a negative predictor in a regression equation or path model (Conger, 1974; Darlington, 1968). Psychometricians have identified a high correlation between the two independent variables as a possible cause of negative suppression (Pandey & Elliott, 2010).

One possible way to deal with the suppression effect is to delete one of the predictors if it is completely redundant. Another possibility is to combine the constructs into one omnibus construct (Maassen & Bakker, 2001). We found this suggestion to be more valid because the bivariate correlation between IV and SE was quite high ($r = .64$, $p < .001$), but they were not completely redundant.

We re-ran the SEM model with a higher-order construct underpinned by the first-order latent constructs of intrinsic value and self-efficacy (see Table 4). We termed this higher order construct general motivation. The results of the analysis were in line with the theoretical expectations. General motivation positively predicted AC, AL, SDT, and CLT. The model had a good fit to the data ($\chi^2 = 497.07$, $\chi^2/df = 1.71$, $p < 0.001$, RMSEA = 0.052, SRMR = 0.047, CFI = 0.94, GFI = 0.87).

Table 4. Hypothesis testing with SEM

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Unstandardized estimates</th>
<th>S.E.</th>
<th>C.R.</th>
<th>$p$</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1 Motivation positively predicts AC</td>
<td>1.40</td>
<td>.194</td>
<td>7.21***</td>
<td>.001</td>
<td>Yes</td>
</tr>
<tr>
<td>H2 Motivation positively predicts AL</td>
<td>1.28</td>
<td>.194</td>
<td>7.45***</td>
<td>.001</td>
<td>Yes</td>
</tr>
<tr>
<td>H3 Motivation positively predicts SDT</td>
<td>1.58</td>
<td>.20</td>
<td>7.90***</td>
<td>.001</td>
<td>Yes</td>
</tr>
<tr>
<td>H4 Motivation positively predicts CLT</td>
<td>1.13</td>
<td>.17</td>
<td>6.80**</td>
<td>.001</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note. **$p < 0.01$; ***$p < 0.001$. S.E. = standard error; C.R. = critical ratio.

Discussion and conclusion

Recent technological advancements in cloud computing and nanotechnology have enabled computing devices to shrink in size and to increase their functionality. Wireless technology is facilitating better connectivity between the learners and the learning resources, and among learners in learning communities (Specht et al., 2012; Wong, Milrad, & Specht, 2015). Such technological advancements are unlikely to be reversible, and they are driving changes in the
social and pedagogical milieu. Given such technological conditions, language teaching practices should involve active construction of digital artifacts that can be refined progressively supported by a face-to-face and/or online community (Wong, 2013).

This paper aimed to develop and validate the MSCLQ as a reliable and valid instrument to measure student motivation in and strategies for seamless Chinese language learning. To fulfil this aim, the relevant literature on meaningful learning (Howland et al., 2012), seamless learning (Wong, 2012; Wong, Milrad, & Specht, 2015) and SLA (Lightbown & Spada, 2013) were reviewed and synthesized. As a result, four subscales which measure students' perceptions of meaningful seamless learning were formed (AL, AC, SDT, CLT). The creation of the MSCLQ was inspired by the MSLQ (Duncan & McKeachie, 2005; Pintrich et al., 1991) and by replacing the learning strategies portion with strategies associated with meaningful seamless learning supported by technologies. This study contributes to the research on motivation and technology-enhanced learning.

Motivation remains an important factor to consider in any form of learning. Students who are more motivated achieve better grades and have higher quality learning (Pintrich, 2003). This has been attested by thousands of psychological studies. Researchers in technology-enhanced learning will need to draw upon motivation research to provide an additional angle to examine the influences of the technology on learning as well as to examine how motivation influences the way students learn in technology-rich settings. Too often, learning technologists have neglected the important role of motivation by exclusively focusing on the technological affordances and technical features of new learning technologies (e.g., Nehme, 2010). While these studies have certainly advanced our understanding, the examination of motivation provides an additional layer of nuance that would be left unexplored if the focus was exclusively on the technical and pedagogical features of various technologies. For example, given the same technology-enhanced learning environment, how do we know who among the students will learn better? How do we account for these individual differences in motivation to make learning maximally effective? These types of questions can only be answered if one takes the role of motivation into account.

Currently, most SLL studies reported were limited by small sample sizes (e.g., Ogata et al., 2008; Wei, 2012). However, as one-to-one computing is becoming more prevalent, the importance of SLL is likely to be raised with more schools tapping into the pedagogical affordances of mobile technology. Given the access to a sizable sample, this study was able to validate the MSCLQ. The findings of the CFA, coupled with the AVE and CR, provide sufficient evidence for the MSCLQ to be accepted for future use in assessing and comparing different designs of mobile-assisted seamless learning interventions. This study therefore has important implications for educators who might be interested in taking up the socio-technological affordances of mobile computing devices to scale up seamless learning.

In addition, we found that self-efficacy and intrinsic value were highly correlated in our current sample, which necessitated their combination into an omnibus variable. It is possible that the Singaporean Chinese students in our study did not make fine-grained distinctions between self-efficacy and task value. Cross-cultural research on motivational constructs has shown that East Asian students from collectivist societies make fewer distinctions between overlapping motivational constructs (e.g., Ho, Hau, & Salili, 2007; King, 2015; King & McInerney, 2014). Psychologists have attributed this to the possibility that academic achievement has become highly internalized among East Asian students. For these students, valuing the task (intrinsic value) and doing it well (self-efficacy) are highly overlapping as they both propel students towards greater task engagement.

This study has several important academic and practical implications. In terms of its academic implications, it advances the literature on mobile seamless learning by developing a questionnaire that could be used to study perceptions of meaningful Chinese seamless learning. The lack of a standardized questionnaire that could be used to measure seamless learning has hindered meaningful integration across different studies. While initially developed for the Chinese language context, it is possible that the questionnaire could be modified by other researchers who are interested in promoting seamless learning in other subject areas. The most obvious ones would be those related to language learning such as learning English or learning a foreign language. However, it is also possible that the questionnaire could be modified to study seamless learning in non-language domains such as science or mathematics.

In terms of practical implications, the results of the current study could be used to aid educators who are interested in enhancing the seamless learning experience of their students. For mobile seamless learning to be meaningful,
engaging students in self-directed and collaborative learning mediated through linguistic artifacts created around students’ life experiences is a pedagogically viable way forward. As reflected by the mean scores of the four dimensions of learning which are all above the mid-point of three (see Table 1), the students involved in this study are reportedly engaged in various forms of learning. In particular, this study suggests the need to focus on intrinsic value. Teachers could foster a greater degree of intrinsic valuing of a subject by designing meaningful learning activities beyond the classroom walls. With their mobile computing devices in hand, students can easily engage in cross-boundary seamless learning and connect what they have learned with what they encounter daily. This will help to alleviate common problems associated with language instruction such as decontextualization and isolated language learning.

One limitation of the current study is that the validation of the instrument was confined to one level of primary three students from a single Singapore school. Future studies involving more schools and perhaps higher levels of students (e.g., secondary school) are desirable. In addition, collecting data with a larger sample size (> 500) will allow researchers to perform CFA using split-half techniques. This will subject the MSCLQ to more rigorous statistical examination to further enhance its validity, reliability and generalizability across different schools and year levels. It is also advisable to adopt a mixed methods approach. Another limitation of the current study is that it has only tested two dimensions of motivation. Motivation is a multifaceted construct, and different theoretical perspectives on motivation focus on different motivational constructs. In future research, more dimensions could be included to further unpack the relationships between mobilized seamless learning and students’ motivation. Lastly, we drew exclusively on a Singapore sample in our study. It is possible that relationships among the variables might be different across cultures. Thus, future pancultural studies are needed that would shed light on the cultural similarities and differences in terms of the variables examined (King & McInerney, 2014). We would also like to encourage future researchers to conduct focus group interviews with students in order to strengthen understanding of their perceptions directly from the students’ first-hand accounts. There could be other dimensions of motivation or seamless learning practices that emerge naturally from students’ experience and initiatives in using technology.

Acknowledgments

The study was funded by the Office of Education Research, National Institute of Education, Nanyang Technological University, Singapore (project ID: OER 16/12 WLH).

References


Modelling Mathematics Teachers’ Intention to Use the Dynamic Geometry Environments in Macau: An SEM Approach

Mingming Zhou, Kan Kan Chan* and Timothy Teo
Faculty of Education, University of Macau // mmzhou@umac.mo // kankchan@umac.mo // timothyteo@umac.mo

*Corresponding author

(Submitted September 21, 2014; Revised November 27, 2015; Accepted December 2, 2015)

ABSTRACT
Dynamic geometry environments (DGEs) provide computer-based environments to construct and manipulate geometric figures with great ease. Research has shown that DGEs have positive impact on student motivation, engagement, and achievement in mathematics learning. However, the adoption of DGEs by mathematics teachers varies substantially worldwide. This paper described Macau secondary schools teachers’ intention to use DGEs within the theory of planned behaviour (TPB). Using the structural equation modelling approach, results revealed that subjective norm and perceived behavioural control were significant positive predictors of intention, while attitude was not. Further, perceived behavioural control had a significant and positive influence on teachers’ attitudes, whereas subjective norm did not. Implications for theory and practice were given.

Keywords
Theory of planned behaviour, Teachers’ intention, Dynamic Geometry Environments, Macau

Introduction
Technology is important to the process of teaching and learning of Mathematics (National Council of Teachers of Mathematics [NCTM], 2000). It increases teachers’ productivity in teaching and engages students in their learning of mathematics (Pierce & Ball, 2009). Thus, calls for the integration of technology in mathematics education have been made in various educational systems as seen in their official curriculum documents (Curriculum and Development Council & Hong Kong Examination and Assessment Authority, 2007; Curriculum and Planning Division, 2012; NCTM, 2000). Thus far, educational systems in the Western nations have responded to the call more rapidly than those in the East such as Hong Kong and Mainland China (Wong, 2003). In recent years, Macau, a special administration region of China, has become active in promoting technology integration in mathematics education among schools and encouraging teachers to use technology in the classroom. Dynamic geometry environments (DGEs) emerge as a technological tool to support mathematical teaching and learning and it has been integrated in the mathematics classroom in various countries (NCTM, 2000). Given its affordances in shaping mathematics education, the sole teacher education provider in Macau, University of Macau, has recently revised its curriculum to incorporate the use of DGEs into the pre-service and in-service courses.

Dynamic Geometry Environments (DGEs)
Dynamic Geometry Environments (DGEs) are tools originally designed for learning geometry in the 1990s. Nowadays, DGE has evolved to include features that can be used in teaching various subjects such as algebra, calculus and statistics. It provides a computer-based environment in which geometric figures can be constructed and manipulated with great ease (Strässer, 2002). In DGEs, users manipulate geometric figures using a visual interface and receive constructive feedback. In addition, multiple representations of the mathematical objects could be linked dynamically (Laborde, 2007). Using DGEs, teachers could represent and manipulate mathematical objects much more easily than before. Research on DGEs provided evidence in support of its ability to transform mathematics teaching and learning. These included the ability of DGEs to facilitate discovery or inquiry-based instruction (Laborde, 2007), to scaffold geometric problem solving (Strässer, 2002), such as using an inductive approach to explore and investigate mathematical proof, which are not usually achievable within a limited time span (Laborde, 2007). In doing so, students develop their higher-order thinking.

However, the use of DGEs in teaching mathematics is not without limitations (Anthony & Clark, 2011). An example is the low adoption of DGEs among teachers. In order to integrate DGEs into the classroom effectively, teachers have
to change the existing mindsets and practices. Essentially, teachers have to learn how to use a new tool, spend more time in designing technology-based tasks to engage students in learning, and revamp their assessment protocols to be compatible with existing standards (Chan, 2015). These reasons could account for the low adoption rate among teachers who would merely use the DGEs as a supplementary tool instead of exploiting its affordances to facilitate students’ learning (Ruthven, Hennessy, & Deaney, 2008).

**Theory of Planned Behaviour (TPB)**

Many models exist that describe and predict the use of technology, such as the Technology Acceptance Model (TAM) (Davis, Bagozzi, & Warshaw, 1989) and Theory of Planned Behaviour (TPB) (Ajzen, 1991). Among these models, the TPB has been used extensively, which assumes that the decision-making process of performing behaviours in a new situation is determined by one’s behavioural intentions, which in turn is affected by three major variables: attitude towards the behaviour, subjective norms, and perceived behavioural control. In addition to the direct influence of the three determinants on intention, research has also shown that subjective norm and perceived behavioural control have a direct influence on individuals’ attitude towards technology (Teo, Lee, & Chai, 2008; Bellone & Czerniak, 2001).

In this study, a total of four variables were examined. Behavioural intention (BI) and attitudes towards the behaviour (ATT) are the two endogenous variables. ATT is a personal variable which is a person’s overall judgment of the behaviour. The judgment comprised beliefs about the consequence of the behaviour which might be positive or negative, and the evaluation of such belief regarding the appropriateness of the behaviour. Another variable in the TPB, subjective norm (SN), refers to the extent one believes that the significant personnel around him or her would support or disapprove a particular behaviour; in this case, the use of DGEs in class. If the person significant to the teacher is supportive of using DGEs in class, the motivation for the teacher to adopt that behaviour would be greater. The third variable, perceived behavioural control (PBC) refers to the extent to which a person feels capable of enacting a behaviour. This belief includes both internal and external constraints or resources that influence one's behaviour. For example, when teachers consider themselves fluent in the use of DGEs, they would tend to perceive greater control of using this tool in class.

Researchers have made the effort to investigate the intention to use of DGEs in classrooms using the TPB model. For example, Stols and Kriek (2011) conducted an exploratory study on 22 mathematics teachers to examine their beliefs and intention to use DGEs. They found that teachers’ perceived usefulness and their level of technological proficiency were the most important predictors of their intention to use such technology.

**Moderating effects**

From the literature, several demographic variables have been identified to have acted as moderators between the intention to use technology and its key determinants. In this study, we focused on age, gender, years of teaching, and technology experience as potential moderators by examining possible differences in the key determinants of technology use intention by teacher profiles.

**Age**

Prensky (2001) argued that teachers’ intention to use of technology varied by their age. For example, those born after 1980, labeled as “digital natives,” had grown up surrounded by technology and therefore were more comfortable in using it than those who were born earlier, labeled as “digital immigrants.” While the digital immigrant teachers struggle to learn and use technologies and are thus reluctant to adopt them, digital native teachers would face with fewer problems with technology integration (O’Bannon & Thomas, 2014). It is possible that ATT and PBC in technology use would be moderated by age. Furthermore, age was also found to moderate the relationship between subjective norms and behavioral intention, with the relationship stronger for older users (Wang, Wu, & Wang, 2009). On this ground, we expected a stronger effect of ATT and PBC on technology use intention for younger teachers than for older teachers, whereas a stronger effect of SN on technology intention for older teachers.
Gender

In several studies, gender has been found to have a moderating effect on the relationship between behavioral intention and its determinants. For example, females were found to be more significantly influenced than males by attitude when making decisions toward technology-based services (Zhu & Chang, 2014). Gender also affects the relationship between social influences and behavioral intention with a stronger effect for women (Huang, Hood, & Yoo, 2013). As such, we expected a stronger effect of PBC on intention to continue using DGEs for male teachers than for female teachers, whereas a stronger effect of ATT and SN on technology intention for female teachers.

Teaching experience

Teaching experience has been found to be strongly associated with teachers’ attitudes and perceptions toward technology usage (Jimoyiannis & Komis, 2007), with more experienced teachers showing more negative attitudes than less experienced teachers (Anderson & Williams, 2012). Further, teachers with less teaching experience reported higher levels of self-efficacy in regards to technology (Kemp, 2002). As these are significant variables in determining their intentions to use technology, teaching experience would also play a role in teachers’ technology intention. Hence, we expected a stronger effect of ATT, PBC and SN on intention to continue using DGEs for teachers with less experience than for those with more experience.

Technology experience

Empirical evidence has demonstrated that technology experience moderates the relationship between key determinants of behaviour intention and technology use intention, such that attitude toward behaviour was more salient with increasing technology experience (Taylor & Todd, 1995) and the subjective norm becomes less important with increasing levels of experience (Venkatesh, Morris, Davis, & Davis, 2003), when predicting technology use intention. Hence, we expected a stronger effect of ATT and PBC on intention to continue using DGEs for teachers with more technology experience and a stronger effect of SN on intention to continue using DGEs for teachers with less technology experience.

The present study

Given that DGEs have the potential to foster more effective teaching and learning of mathematics, it is important to understand the impetus of mathematics teachers’ intentions to continue the use of DGEs (Ertmer, 2005). This study aimed to examine Macau secondary school teachers’ intention to use DGEs in the mathematics classroom and it contributes to the current literature in four ways. First, although the TPB has been applied to examine teachers’ intention to use technology in many studies (e.g., Sadaf, Newby, & Ertmer, 2012), these were typically related to general computing such as computer, PowerPoint or Web 2.0. According to Ajzen (2006), further examination of the predictive powers of the TPB model should be conducted with specific behaviours. Second, a limited number of empirical studies has been conducted to examine the use of TPB to understand teachers’ intention to use DGEs (e.g., Chan, 2015; Stols & Kriek, 2011). Third, the current study could shed light on the cross-cultural validity of the TPB. Macau is characterized by an educational system highly influenced by Confucian values where examinations drive both teaching and learning (Morrison & Tang, 2002). Results from this study would have strong implications for exam-oriented classrooms. Finally, we adopted the structural equation modelling (SEM) approach in the current study, in contrast to the mainly descriptive, correlation statistics, and regression analysis in existing studies. SEM is useful for analysing the relationships between latent and observed variables, with random errors in the observed variables estimated directly. This is not possible with traditional techniques (e.g., multiple regression). Consequently, SEM produces more precise measurements of the constructs in research (Teo & Zhou, 2014).

The following questions guide this study:

- To what extent is the TPB a valid model to predict mathematics teachers’ intention to continue using DGEs?
- What direct and indirect effects do the TPB constructs have on mathematics teachers’ intention to continue using DGEs?
• In what way is the relationship between the intention to continue using DGEs and its key determinants moderated by teachers’ age, gender, teaching experience, and technology experience?

From the above discussion of related literature, the following hypotheses were formulated:

H1: ATT has a direct and significant influence on teachers’ intention to continue using DGEs in teaching mathematics
H2: SN has a direct and significant influence on teachers’ intention to continue using DGEs in teaching mathematics
H3: PBC has a direct and significant influence on teachers’ intention to continue using DGEs in teaching mathematics

Research has also shown that SN and PBC have a direct influence on ATT of the individual (Teo et al., 2008). Hence, the following hypotheses were also tested.

H4: SN has a direct and significant influence on teachers’ ATT.
H5: PBC has a direct and significant influence on teachers’ ATT.

Figure 1 summarizes the hypotheses proposed in this study.

![Proposed research model](image)

ATT=Attitude; SN=Subjective Norm; PBC=Perceived Behavioural Control; BI=Behavioural Intention

In addition, the moderating effects of gender, age, years of teaching and technology experience on the relationships depicted in the model are well supported in the literature as reviewed above. Hence, the following hypotheses were tested:

H1a: ATT influences teachers’ intention to use DGEs more strongly for younger teachers than for older teachers.
H2a: SN influences teachers’ intention to use DGEs more strongly for older teachers than for young teachers.
H3a: PBC influences teachers’ intention to use DGEs more strongly for younger teachers than for older teachers.
H4a: SN influences teachers’ ATT more strongly for older teachers than for younger teachers.
H5a: PBC influences teachers’ ATT more strongly for younger teachers than for older teachers.

H1b: ATT influences teachers’ intention to use DGEs more strongly for female teachers than for male teachers.
H2b: SN influences teachers’ intention to use DGEs more strongly for female teachers than for male teachers.
H3b: PBC influences teachers’ intention to use DGEs more strongly for male teachers than for female teachers.
H4b: SN influences teachers’ ATT more strongly for female teachers than for male teachers.
H5b: PBC influences teachers’ ATT more strongly for male teachers than for female teachers.
H1c: ATT influences teachers’ intention to use DGEs more strongly for less experienced teachers than for more experienced teachers.
H2c: SN influences teachers’ intention to use DGEs more strongly for less experienced teachers than for more experienced teachers.
H3c: PBC influences teachers’ intention to use DGEs more strongly for less experienced teachers than for more experienced teachers.
H4c: SN influences teachers’ ATT more strongly for less experienced teachers than for more experienced teachers.
H5c: PBC influences teachers’ ATT more strongly for less experienced teachers than for more experienced teachers.

H1d: ATT influences teachers’ intention to use DGEs more strongly for teachers with more technology experience than teachers with less technology experience.
H2d: SN influences teachers’ intention to use DGEs more strongly for teachers with less technology experience than teachers with more technology experience.
H3d: PBC influences teachers’ intention to use DGEs more strongly for teachers with more technology experience than teachers with less technology experience.
H4d: SN influences teachers’ ATT more strongly for teachers with less technology experience than for teachers with more technology experience.
H5d: PBC influences teachers’ ATT more strongly for teachers with more technology experience than for teachers with less technology experience.

Method

Participants and procedure

A survey design was used in this study. An invitation letter was sent to the principals of all 35 secondary schools in Macau where Chinese was used as the medium of instruction. The letter introduced the purpose of the questionnaire and invited mathematics teachers to take part in this study. The link of the online survey was emailed to the mathematics department head in each school who was responsible for sharing the link with all other mathematics teachers in the department. A total of 171 mathematics teachers voluntarily completed an online survey, while 19 completed a paper version of the survey due to a lack of access to the Internet at the point of completing the questionnaire. This led to a total of 190 participants in this study, representing 40% of the whole population of mathematics teachers (475) in Macau at the time of data collection. The survey consisted of two sections (1) demographic information of the participants and (2) items designed to measure teachers’ intention to use DGEs in class (in this study GeoGebra was promoted to Macau teachers, see Figure 2).

Figure 2. A screenshot of GeoGebra
Table 1 shows the profile of the participants in this study. Approximately 36% of the participants were new teachers (with less than four years of teaching experience) and one third of them have taught Mathematics for more than 10 years. Over 70% of the teachers rated their computer proficiency at a medium level. The majority of teachers (75%) had experienced using DGEs in the classroom.

**Measures**

For the direct measures of TPB, the original scales by Ajzen (2006) and Francis et al. (2004) were adapted by specifying the type of technology being examined. A sample item was: “In general, I believe the use of DGEs in geometry lesson is convenient.” A standard back-translation procedure was followed, wherein the second author translated the English items into Chinese and back-translated into English by two other researchers who had no knowledge about the scale. The final translated items were included in the reconciled version only after all three researchers had reached a consensus. The final questionnaire contains 11 items that measured behavioural intention, attitudes, subjective norms and perceived behavioural control, on a 7-point Likert scale, with 1 = “strongly disagree” to 7 = “strongly agree.” The original English scales have been shown satisfactory validity and reliability in previous studies (Hammer & Vogel, 2013).

For the indirect measures of TPB, the items were elicited from 30 mathematics teachers in Macau who were knowledgeable about and have employed DGEs in their teaching. Their responses to an open-ended questionnaire were coded to generate the indirect measures for each TPB variable. The detailed coding process and results could be referred in Chan (2015). We invited three graduate students to respond to the indirect measures to examine the validity of the measure.
Data analysis

Structural equation modelling (SEM) was employed in this study its ability to analyse the relationships between the latent and observed variables and estimate random errors in the observed variables directly, giving rise to more precise measurements of the items and constructs in the survey. As the focus of the present study is on a modified TPB model, SEM has an added advantage over traditional data analysis techniques by modelling the relationships among latent variables thus aligning with the practice of hypotheses testing (Hoyle, 2011). Using the two-step approach to SEM, we first estimated the measurement model (CFA model) to describe how well the observed indicators (survey items) measure the unobserved constructs. In the second step, we estimated the structural part of the SEM to specify the relationships among the exogenous and endogenous latent variables. The measurement model and structural model were tested with AMOS 7.0 using the maximum likelihood estimation (MLE) (Schumacker & Lomax, 2010). Several goodness-of-fit indices were used to evaluate the structural model: the chi-square test, root mean square error of approximation (RMSEA) and Standardized Root Mean Residual (SRMR), Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) (Kline, 2010). Values of .90 or more for the CFI and TLI, and values of .08 or less for RMSEA and SRMR represent an acceptable fit (Hair, Black, Babin, & Anderson, 2010).

To examine the moderation effects of age, gender, teaching experience, and technology experience on the key relationships in the proposed model, a series of multi-group analyses were performed. We used the median value to divide each moderator into two groups to represent the upper and lower limits, except for DGEs experience, which was divided by “with experience” and “no experience.” The individual parameter estimate of each path in the model was compared for high and low levels of each variable (Zweig & Webster, 2003). We used the Z-score as a critical ratio (Hunter & Schmidt, 1990) when comparing the mean correlations of moderator subgroups.

Results

Descriptive statistics

All 11 items measuring the four constructs in the TPB were examined for their mean, standard deviation, skewness, and kurtosis. The mean scores ranged from 3.01 and 5.16, indicating an overall positive response to the items that were used to measure the constructs in this study. All standard deviations were above 1.00, ranging from 1.04 to 1.42, indicating a narrow spread of scores around the mean. Skewness and kurtosis indices were small and well within the accepted the accepted level of |3| and |10| respectively (Kline, 2010). On the recommendations from Kline (2010), the data in this study were considered to be univariate normal.

Evaluation of the measurement model

On what is considered an adequate sample size for SEM, researchers recommend a sample size of 100 to 150 cases (Kline, 2010). In addition, researchers also refer to the Hoelter’s critical $N$, which refers to the sample size for which one would accept the hypothesis that the proposed research model is correct at the .05 level of significance. The Hoelter’s critical $N$ for the model in this study is 136 and, given that the sample size of this study is 190, structural equation modelling is regarded as an appropriate technique for data analysis. As multivariate normality of the observed variables is assumed in the MLE procedure, the data in this study were examined using the Mardia’s normalized multivariate kurtosis value. The Mardia’s coefficient for the data in this study was 44.186, which is lower than the value of 143 computed based on the formula $p(p+2)$ where $p$ equals the number of observed variables in the model (Raykov & Marcoulides, 2008). On this basis, multivariate normality of the data in this study was assumed.

The results of the CFA are shown in Table 2. All parameter estimates were significant at the $p < .01$ level, as indicated by the $t$-value (greater than 1.96). The standardized estimates ranged from .448 to .963, and these were regarded as acceptable (Hair et al., 2010). The internal consistency of all constructs ranges from .64 to .76. Finally, the measurement model has a good model fit ($\chi^2 = 71.321, \chi^2/df = 1.981, TLI = .921, CFI = .948$, RMSEA = .072, SRMR = .053). The adequacy of the measurement model indicated that all items were reliable indicators of the hypothesized constructs they were purported to measure.
Table 2. Results for the measurement model

<table>
<thead>
<tr>
<th>Item</th>
<th>UE</th>
<th>t-value</th>
<th>SE</th>
<th>Cronbach Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT1</td>
<td>.649</td>
<td>5.614</td>
<td>.448</td>
<td>.70</td>
</tr>
<tr>
<td>ATT2</td>
<td>.922</td>
<td>8.672</td>
<td>.719</td>
<td></td>
</tr>
<tr>
<td>ATT3</td>
<td>1.000</td>
<td>---</td>
<td>.832</td>
<td></td>
</tr>
<tr>
<td>SN1</td>
<td>1.603</td>
<td>5.255</td>
<td>.865</td>
<td>.64</td>
</tr>
<tr>
<td>SN2</td>
<td>1.000</td>
<td>---</td>
<td>.543</td>
<td></td>
</tr>
<tr>
<td>PBC1</td>
<td>.762</td>
<td>6.379</td>
<td>.539</td>
<td>.76</td>
</tr>
<tr>
<td>PBC2</td>
<td>1.000</td>
<td>---</td>
<td>.866</td>
<td></td>
</tr>
<tr>
<td>PBC3</td>
<td>.789</td>
<td>6.853</td>
<td>.583</td>
<td></td>
</tr>
<tr>
<td>INT1</td>
<td>.419</td>
<td>4.835</td>
<td>.421</td>
<td>.75</td>
</tr>
<tr>
<td>INT2</td>
<td>.606</td>
<td>5.917</td>
<td>.557</td>
<td></td>
</tr>
<tr>
<td>INT3</td>
<td>1.000</td>
<td>---</td>
<td>.943</td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .01. UE: Unstandardized Estimate; SE: Standardised Estimate; --- This value was fixed at 1.00 for model identification purposes.

Evaluation of the structural model and hypothesis testing

There was a good fit for the structural model: $\chi^2 = 71.207; \chi^2/df = 2.034; TLI = .916; CFI = .947; RMSEA = .072; SRMR= .053$. Overall, three out of five hypotheses in this study were supported by the data. Except for $H_1$ and $H_4$, all hypotheses were supported in this study. Two endogenous variables (behavioural intention to use, attitude towards use) were tested in the research model (Figure 1) and the variance explained in ATT was higher than BI with an $R^2 = 0.523$ and $R^2 = 0.195$, respectively. This means that together, SN and PBC accounted for 52.3% of the variance in ATT. On the other hand, ATT, SN, and PBC accounted for only 19.5% of the variance in BI. The summary of the hypotheses testing results is shown in Table 3.

Table 3. Hypothesis testing results

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Path</th>
<th>Path coefficient</th>
<th>t-value</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>ATT $\rightarrow$ BI</td>
<td>.130</td>
<td>1.510</td>
<td>Not supported</td>
</tr>
<tr>
<td>$H_2$</td>
<td>SN $\rightarrow$ BI</td>
<td>.186</td>
<td>2.313</td>
<td>Supported</td>
</tr>
<tr>
<td>$H_3$</td>
<td>PBC $\rightarrow$ BI</td>
<td>.197</td>
<td>1.994</td>
<td>Supported</td>
</tr>
<tr>
<td>$H_4$</td>
<td>SN $\rightarrow$ ATT</td>
<td>.213</td>
<td>1.850</td>
<td>Not supported</td>
</tr>
<tr>
<td>$H_5$</td>
<td>PBC $\rightarrow$ ATT</td>
<td>.584</td>
<td>4.442</td>
<td>Supported</td>
</tr>
</tbody>
</table>

Tests of direct and indirect effects

Table 4 shows the standardized total effects, direct and indirect effects associated with each of the four variables. A coefficient linking one construct to another in the path model (Figure 2) represents the direct effect of a determinant on an endogenous variable. An indirect effect reflects the impact a determinant has on a target variable through one or more other intervening variables in the model. A total effect on a given variable is the sum of the respective direct and indirect effects. The effect sizes with values less than 0.1 were considered small, those with less than 0.3 are medium, and values with 0.5 or more considered large (Cohen, 1988). The most dominant determinant of behavioural intention is perceived behavioural control, with a total effect of 0.427. This is followed by subjective norm and attitude with a total effect of 0.334 and 0.200, respectively.

Table 4. Tests of direct and indirect effects

<table>
<thead>
<tr>
<th>Endogenous variable</th>
<th>Determinant</th>
<th>Standardized estimate</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Direct</td>
<td>Indirect</td>
<td>Total</td>
</tr>
<tr>
<td>Behavioural intention</td>
<td>Attitude</td>
<td>.200</td>
<td>.000</td>
<td>.200</td>
</tr>
<tr>
<td></td>
<td>Subjective norm</td>
<td>.212</td>
<td>.042</td>
<td>.334</td>
</tr>
<tr>
<td></td>
<td>Perceived behavioural control</td>
<td>.310</td>
<td>.117</td>
<td>.427</td>
</tr>
<tr>
<td>Attitude</td>
<td>Subjective norm</td>
<td>.212</td>
<td>.000</td>
<td>.212</td>
</tr>
<tr>
<td></td>
<td>Perceived behavioural control</td>
<td>.585</td>
<td>.000</td>
<td>.585</td>
</tr>
</tbody>
</table>
Moderation analyses

An examination of the differences by gender revealed that male teachers expressed more negative relationships in their attitudes towards DGEs and the intention to use technology \( (\beta = -0.34, \text{Table 5}) \). An opposite pattern was observed in female teachers whose attitudes towards DGEs had a positive impact on their intention to use technology \( (\beta = 0.34) \). Teaching experience also moderated this path with a similar pattern: teachers who have taught over 10 years tended to report a negative relationship between attitudes and intention to use technology \( (\beta = -0.97) \) whereas teachers who have taught for less than 10 years showed a positive attitude towards using technology \( (\beta = 0.20) \). Further, the path between PBC and attitudes was moderated by teaching and technology experience. Specifically, teachers with more experience in teaching or with using DGEs reported a stronger path \( (\beta = 0.82; \beta = 0.57, \text{respectively}) \) than those with less experience in teaching or with using DGEs \( (\beta = 0.22, \beta = 0.28, \text{respectively}) \). No moderating effects were observed for age.

<table>
<thead>
<tr>
<th>Path</th>
<th>Moderator: Gender</th>
<th>Critical ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT → BI</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>-0.34</td>
<td>-0.34</td>
</tr>
<tr>
<td>Moderator: Teaching Experience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBC → ATT</td>
<td>More Experience</td>
<td>Less Experience</td>
</tr>
<tr>
<td></td>
<td>0.82**</td>
<td>0.22</td>
</tr>
<tr>
<td>ATT → BI</td>
<td>Less Experience</td>
<td>More Experience</td>
</tr>
<tr>
<td></td>
<td>-0.97</td>
<td>0.20</td>
</tr>
<tr>
<td>Moderator: Technology Experience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBC → ATT</td>
<td>With Experience</td>
<td>No Experience</td>
</tr>
<tr>
<td></td>
<td>0.57***</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Note. *** \( p < .001 \); ** \( p < 0.01 \); * \( p < 0.05 \).

Discussion

Straesser (2002) noted that DGEs may be one of the highly-featured software within mathematics education research due to its wide use in schools around the globe. However, past research had mainly focused on the effectiveness of DGEs in improving student's performance (e.g., Dogan & İçel, 2011). The current study explored the determining factors that affect teachers' decision to adopt DGEs in classroom. Consistent with past findings (Cheon, Lee, Crooks, & Song, 2012; Stols & Kriek, 2011), perceived behavioural control was found to be the most powerful factor in explaining Macau teachers' intention for using DGEs.

From development, DGEs were developed for users to explore mathematics concepts with its interactive diagrammatic interface (Laborde, 2007) and this feature poses a heavy demand on teachers' classroom activity design (Jones, 2002). To use DGEs effectively, teachers need to possess a sufficient level of domain knowledge, conceptions of mathematics learning as well as pedagogical strategies (Koehler & Mishra, 2009). Thus, what teachers' believe about their knowledge and skills to use this software are important factors in their decision to use DGEs.

Next, subjective norm was a significant variable in explaining Macau teachers' intention to use the DGEs. Consistent with current research, it appeared that the teachers in this study had regarded their subject leaders, school principals, and colleagues as important influences to their use of DGEs in the Macau classroom (Teo & Zhou, 2014; Lee, Ceereto, & Lee, 2010). In the Macau context, decisions on technology procurement and adoption are largely made by the school leaders. There is limited autonomy on the part of the teachers on technology integration hence the big impact of their normative beliefs on their behavioural intention to use the technology.

The missing support for the positive effect of attitudes on teachers' intention to continue using DGEs in teaching mathematics \( (H_1) \) was unexpected as attitude towards technology has been documented to be a strongest predictor of intention to use technology (e.g., Cheon et al., 2012; Stols & Kriek, 2011). Although the majority of the teachers in this study had used DGEs, they had not received any formal training in using it. Consequently, they may not have optimised the use of DGEs to guide students' learning and were unaware of various innovative and pedagogical uses of DGS in the mathematic classroom-experiences that may have shaped teachers' attitude. Coffland and Strickland
(2004) noted that teachers’ attitudes and technology awareness were positively related and those who were trained in the integration of technology were more likely to develop positive attitudes towards using technology for teaching. To an extent, the interaction between knowledge and experience in DGEs of the participating teachers in this study might have weakened the predictive power of attitudes on technology use intention.

Also noteworthy was that perceived behavioural control had a significant influence on teachers’ attitudes (H3), despite the missing direct effect of attitudes on intention. This was in conflict with studies where attitudes had mediated the relationship between perceived behavioural control and intention (e.g., Yang & Zhou, 2011). However, considering that perceived behavioural control is jointly dependent on motivation (intention) and ability (behaviour control) (Ajzen, 1991), the motivation dimension seemed to be more closely relevant, when linked with other variables such as attitudes, explaining the significant influence from PBC to attitudes.

As an external variable, subjective norm did not have a significant effect on attitudes. Although some researchers have argued that the expectations from the important “others” had an impact on their attitudes toward technology use (Teo, 2009), which in turn influenced the intended use of that technology, our data showed that this was not the case. In places where key decisions are made collectively, as in the case of Macau, individual preferences may not have much influence on the outcomes, such as the type of technology to use for instruction. This may be the sentiments of the Macau teachers in our study who felt it more important to comply with policies or decisions made by their school leaders, rather than acting from their personal views, explaining the significant influence subjective norm had on intention, but not on attitude towards the use of DGEs.

Although there was evidence in support of a positive relationship between each demographic variable and the intention to use technology, the findings in this study revealed only a few to be statistically significant. One possible reason is that this study had focused on teachers’ individual background in the selection of moderators and, while these could have had an impact on teacher’s attitudes and their technology intention, or between PBC and attitudes, it may be possible that other variables related to a teacher’s perceptions had greater influences in other paths in the model, such as self-efficacy (Islam, Khan, Ramayah & Hossain, 2011) or voluntariness (Venkatesh et al., 2003). Future studies are warranted for further investigation.

Implication for theory and practice

There are few studies which reveal teachers’ intention of using special mathematical software. This study provides empirical evidence to support the argument that the role of three direct variables of TPB model as predictors of intention of using DGEs is different with that of general software such as PowerPoint (Lee et al., 2010). This study verifies the importance of specifying precise behaviour when applying TPB, and contributes to a greater understanding of the validity of the TPB as a model to explain the intention to use technology for teaching mathematics with a sample of Asian teachers, in particular, among Macau teachers. Both teachers’ beliefs that they have sufficient knowledge and skills to use the technology tool as well as their subject leaders, school principals, and colleagues’ recommendations emerge as important factors to their use of technology.

Findings of this study suggest that decision makers in schools can continue playing a role in promoting the use of DGEs. Technology integration in the Macau education system is not as clear as those in other countries such as Australia and the United Kingdom. Therefore, creating a vision of using technology to facilitate teaching and learning of mathematics could increase teachers’ intention of using such technology. Nirode (2012) noted that teachers’ use of DGEs was restricted to convergent tasks of basic construction and measurements. School leaders should create conducive environments to support mathematics teachers’ adoption of DGEs in the classroom by providing time and access to resources for teachers to explore various technologies and ways to use them, such as properly installing software in school computers, organizing workshops and seminars, promoting peer sharing of lesson plans to allow for the production of dynamic lessons and discussions of related pedagogical issues and so forth. These measures could enhance teachers’ perceived behaviour control which in turn strengthens their attitudes towards DGEs.
Limitations

Several limitations exist in this study. First, the participants in this study were recruited from the secondary schools. Hence, replication studies are needed in order to generalize the results of this study to all in-service teachers. In addition, future research should endeavour to understand students’ perceptions since they are key stakeholders in education. Finally, since a web-based questionnaire is used in this study, a bias towards low computer anxiety and high computer self-efficacy could be present in the data thus skewing the respondents’ inclinations of adopting technology (Svendsen, Johnsen, Almås-Sørensen, & Vittersø, 2013). Finally, future research could include a qualitative study to cross-validate the proposed model in this study.

Conclusion

This study aimed to examine Macau teachers’ intention of using DGEs in mathematics classroom using the theory of planned behaviour. Results showed that Macau teachers were affected by their important others and the perceived behavioural control. Evidence of this study confirmed that the predictive power of direct determinants of intention would be different in different technology.

References


Web-based Learning System for Developing and Assessing Clinical Diagnostic Skills for Dermatology Residency Program

Fan-Ray Kuo1, Yi-Ying Chin2 3, Chao-Hsien Lee4, Yu-Hsien Chiu5, Chien-Hui Hong6 7, Kuang-Lieh Lee8, Wen-Hsien Ho5 and Chih-Hung Lee9 10*

1National Taiwan University, Taipei, Taiwan // 2Kaohsiung Municipal Hsiao-Kang Hospital, Kaohsiung, Taiwan // 3Kaohsiung Medical University Hospital, Kaohsiung, Taiwan // 4National Taiwan University of Technology, Taipei, Taiwan // 5Kaohsiung Medical University Hospital, Kaohsiung, Taiwan // 6National Yang Ming University, Taipei, Taiwan // 7Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan // 8National Kaohsiung Normal University, Kaohsiung, Taiwan // 9Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan // 10Chang Gung University College of Medicine, Taoyuan, Taiwan // revonkuo@gmail.com // ninakim515@yahoo.com.tw // chlee@ntut.edu.tw // chiuyh@kmu.edu.tw // zieben@gmail.com // science.lee@livemail.tw // whho@kmu.edu.tw // dermlee@gmail.com *Corresponding author

(Submitted October 26, 2014; Revised July 11, 2015; Accepted October 27, 2015)

ABSTRACT

Few studies have explored the learning difficulties and misconceptions that students encounter when using information and communication technology for e-learning. To address this issue, this research developed a system for evaluating the learning efficiency of medical students by applying two-tier diagnosis assessment. The effectiveness of the proposed system was evaluated in 11 resident doctors of dermatology enrolled according to a one-group pre- and post-test experimental design. After using the proposed e-learning system, the resident doctors had significantly improved learning efficiency and high satisfaction with the system functions. Although the number of subjects in this study was low, the significant improvements in clinical skills observed at the end of the educational training program indicate that the proposed integrated e-learning system is user-friendly and effective for self-education in the diagnosis skills required by medical students and dermatology residents.

Keywords

Meaningful learning, Two-tier diagnosis assessment, Web-based learning, Clinical diagnosis, Clinical dermatology

Introduction

As information and communication technology rapidly evolves and matures, computer-based learning and web-based learning are being replaced with mobile learning approaches (Chen et al., 2005). Holzinger et al. (2005) created mobile interactive learning objects that run on mobile phones to enable medical staff and students to access information quickly and conveniently. The main role of the student has evolved from a passive receiver of knowledge to an active learner, and the role of the educator has shifted from a teacher of students to a facilitator in acquiring knowledge and in improving learning skills (Albarrak, 2010). That is, the learning context has shifted from a teacher-centered model to a learner-centered model in which learning motivation and interaction are higher than ever (Hu et al., 2008).

One form of e-learning is web-based learning, in which learning materials for a given purpose are delivered to students via the Internet (Hsu, 2004). One example is APoMT, a technology-based modeling tool developed by Wu et al. (2010). By applying scaffolding theory and recommended guidelines, APoMT assists students in performing scientific modeling. Wu et al. (2010) showed that APoMT effectively supports the construction of conceptual understanding and learning performance in students who use the tool. Hsu et al. (2008) developed an e-learning platform to teach 75 high school students how the varying distance between the sun and the earth affects the weather during each season. Their findings showed that the platform effectively helped students to model science concepts. Various e-learning platforms have also been used in medical education courses. For example, Bogoch et al. (2012) used a physical examination blog as a tool for enhancing and supplementing evidence-based classes in bedside physical examination. Notably, 77% of the students agreed that, because the blog included an evidence-based component, it was a helpful or very helpful supplement to traditional classes in bedside physical examination. Additionally, 73% agreed that the blog was helpful or very helpful for learning how to perform a physical examination. Fraser et al. (2011) investigated the e-learning attitudes of 117 otolaryngology trainees. Their survey
indicated that most of the trainees had been exposed to e-learning throughout their training, and, for most trainees, current self-rated computer literacy was high. Finn (2010) investigated the use of e-learning for teaching cardiopulmonary resuscitation (CPR) to medical students. A survey of the students indicated that most agreed that the e-learning platform enhanced their understanding of CPR concepts and substantially increased their learning motivation. In Harrison et al. (2013), a website was used to deliver feedback about the objective structural clinical examination (OSCE) to clarify how students engage with various forms of feedback in this context. The study showed that students who performed well on the OSCE viewed significantly more web pages compared to students who performed poorly. Marshall et al. (2011) established and evaluated an e-learning platform for teaching medical students when a radiologic examination should be ordered. The results showed that the platform significantly improved the quality of radiologic examination orders, especially in terms of logistics and radiation safety. Another study of medical students by O’Leary and Janson (2010) investigated whether simulations delivered by an e-learning system improve performance in pediatric resuscitation. The research findings showed that the simulations significantly improved both the knowledge and competence of medical students in performing pediatric cardiopulmonary resuscitation. A study by O’Neill et al. (2011) developed a web-based infection prevention and control program for medical students. The 517 medical students who participated in the study showed statistically significant improvements in their knowledge after completing two modules. Most students who participated in the assessment of the online modules gave positive feedback about their learning experience. For medical students, Schmeling et al. (2011) developed a web-based INMEDEA simulator that included 15 virtual cases of death. The participants had strong positive overall perceptions of the simulation system. Mohamadirizi et al. (2014) developed an e-learning system for enhancing interaction between prenatal care patients and medical personnel by providing the patients with relevant medical information. After using the system for 4 weeks, the patients showed significantly improved understanding of prenatal care concepts and significantly improved satisfaction with the care received. In Weiner et al. (2014), an online mentoring database was established to improve research support by medical students and to improve the effectiveness of shadowing. The students reported that the proposed mentoring database was definitely helpful or potentially helpful for matching them with a mentor. The above studies of the effectiveness of e-learning in medical education indicate that e-learning not only improves the learning performance of medical students in various disciplines, but also increases their learning motivation and satisfaction.

However, student learning performance can be affected by many factors, including learner competence, learning materials, family socioeconomic background, learning method, learning strategy, etc. (Oloruntegbe & Ikpe, 2010). Of these, learning method and strategy are considered the most important determinants (Tsai & Shen, 2009; Hwang & Kuo, 2011; Kuo et al., 2012). Studies indicate that integrating technology in learning methods or strategies enhances the learning performance of medical students (Kurowski et al., 2013). For example, in Lin (2013), a cooperative learning approach that integrated technology improved the critical thinking skills of nursing students when performing catheterization. In Garrett and Jackson (2006), a wireless personal digital assistant was integrated with the use of a reflection strategy to support professional reflection by healthcare professionals in a clinical learning environment. Experiments showed that application of the personal digital assistant and the reflection strategy by nursing students enhanced their learning performance.

Based on recent medical advances and learning theory, Farrimond et al. (2006) developed an e-learning package for teaching skin examination. A survey of the learners indicated that they considered the e-learning system an effective adjunct system that helped them build confidence in performing skin examinations. Nast et al. (2009) proposed the use of online lectures in a learning strategy for dermatology students. The participants responded positively to the online lectures in dermatology and indicated that the learning system may be a good means of improving their learning performance. Holzinger et al. (2009) applied both pedagogical and psychological expertise in the development of user-centered simulation software for medical students. Experiments showed that the simulation software enhanced the use of cognitive skills by medical students when learning complex concepts. Silva et al. (2011) designed an e-learning program for dermatology students and evaluated its impact on learning efficiency. Their evaluation showed that students who participated in face-to-face activities and associated online discussions (blended learning) in support of their classroom learning had significantly higher posttest scores compared to students who only participated in classroom learning activities. Some studies have also investigated the use of tablet PCs to deliver learning materials, e.g., textbooks and actual health records, to support the development of critical decision-making skills in medical students (Bloice et al., 2014).
The literature agrees that various learning technologies can be used to deliver media that improve the learning performance and attitude of medical students. However, few studies have investigated the learning difficulties and misconceptions of students during the learning process. The two-tier diagnostic assessment developed by Treagust (1986) is a research instrument that is widely used to quantify understanding in students (Treagust, 1995). The questions used in the two-tier diagnostic assessment are mainly designed to identify alternative conceptions or misconceptions in a limited context. For each item, the first tier contains a multiple-choice (three or four choices) content question that usually includes essential concepts of the learning module. The second tier of each question contains a set of four or five possible explanations for the answer given in the first tier. Two-tier diagnostic assessment has been used in general science (Tsai & Chou, 2002) and in specific disciplines such as chemistry (Tan & Treagust, 1999), biology (Goh et al., 1993), and mathematics (Wang, 2014). Thus, educators can use two-tier diagnostic assessment technique to measure conceptual understanding when evaluating the learning efficiency of individual students. To address the above issues, this study developed a medical diagnosis learning system based on two-tier diagnosis assessment technique. The purpose of the proposed system is not only to facilitate the development of clinical skills such as diagnosis of skin diseases and prescription of medications in resident doctors, but also to save health care resources.

Methods

This is a pilot study. Considering the small sample size, a single group with a pre-experimental design compares the subject performance before and after the completion of the e-learning program. The lack of a control group may drastically affect the interpretation of the study result. Hence, a complete experimental design including control groups with mock training modules and/or classic learning program could consolidate the result.

Participants

This study included 11 physicians (five males, six females) in the residency programs of dermatology departments of medical centers in Kaohsiung, Taiwan. Their age ranged from 26 to 32 years. All participants had medical licenses issued by the Taiwan MOH, and all were enrolled in 4-year residency programs in dermatology, which included comprehensive training in skin physiology, skin histology, clinical diagnostic and treatment skills, dermatoscopy, dermatological surgery, clinical decision making, etc. For all participants, the dermatology training also included interpretation of microscopic images, which is essential for accurate and prompt diagnosis. After completing their residency training, the participants were eligible to take the examination for board certification in dermatology. The study was approved by the Institutional Review Board of Kaohsiung Medical University (KMUH-IRB-20120174). Each participant was assigned a personal account name and password to access the e-learning website used in this study.

Experimental flow

Figure 1 shows the three stages of the experiment: pre-test, learning activity with learning system, and post-test. Before participating in the learning activity, the eleven residents were required to complete a pre-test to determine their preliminary competence in dermatology. Each resident then used the medical diagnosis learning system to perform 120 hours of learning activities. The subject matter included five categories of dermatological diseases: infection, neoplasm, autoimmunity, inflammation, and genetic diseases. Each category had five difficulty levels. For each difficulty level, the participant was required to answer six questions. Each participant generally required at least 20 minutes of learning to answer a question correctly. Therefore, the learning procedure required 60 hours. Assuming the learning procedure was repeated two times, the total learning time was 120 hours. This duration was consistent with Phillips and Phillips (2007), who suggested that a minimum of 120 hours is required could decide the learning efficiency of an organization. Finally, a post-test and questionnaires were administered to measure learning performance and satisfaction with the system, respectively.
Medical diagnosis learning system

To identify the misconceptions of the dermatology residents, gross and microscopic images of skin biopsies in patients with infection, neoplasm, autoimmunity, inflammation, and genetic abnormality were collected and stored in the medical diagnosis learning system. To facilitate learners in achieving the learning objectives and to minimize misconceptions about the skin diseases, meaningful learning theory (Ausubel, 2000) and two-tier diagnosis technique were applied when developing the proposed system. Figure 2 shows that the role of the user was either instructor or learner. Instructors were the domain experts; all instructors were experienced dermatologists who were certified by the Taiwan Board of Dermatology or who had senior positions in their departments. Thus, the instructors taught the inexperienced resident doctors, i.e., learners, by applying and sharing their clinical experience in performing dermatological treatments. As Figure 2 shows, the instructors first collected and analyzed clinical and pathological biopsy images. The system automatically processed the images to maintain a consistent image quality and format and automatically scaled the images to the same size. Each image was also watermarked with copyright information. The instructors then used two-tier diagnosis assessment technique to design questions regarding the clinical and pathological characteristics of the biopsy samples depicted in the images. To aid learners in developing skills in distinguishing skin diseases according to the clinical and pathological characteristics of each specimen, the range of answer choices for each question included skin diseases that are easily mistaken for the disease actually associated with the specimen. Finally, the instructors used Delphi method to rate the difficulty of each question before storing the question in the dermatology diagnosis database.

The right side of Figure 2 further shows that learners who required further professional training in dermatology were allowed to log into the proposed system and practice distinguishing skin diseases by browsing and answering the questions designed by the instructors. The proposed medical diagnosis learning system randomly selected one question from the dermatology diagnosis database. Each question included a brief text description and several clinical and pathological biopsy images. If the learner answered the question correctly, the system randomly selected another question. If the learner answered the question incorrectly, the system determined whether the learner had exceeded the default limit of three incorrect answers for that question (Wang et al., 2004; Wang, 2010). If the limit had not been reached, the system prompted the learner to attempt another answer. If the limit had been reached, the system showed the correct answer with an explanation and randomly selected another question. Therefore, the system enabled repeated practice in skin diagnosis.

The medical diagnosis learning system was also designed to enter practical data in the learning portfolios of individual learners, such as the number of incorrect answers to a question about a specific disease, login frequency, login time, logout time, exercise frequency, time spent on each diagnosis, etc. After the instructor used the two-tier diagnostic assessment tool to determine the learning status of each learner, an appropriate remedial teaching strategy was selected. For example, if the learning system identified a weakness in a specific clinical skill (e.g., if the learner had a low score in the neoplasm category), a remedial teaching strategy was used to improve that skill. The strategy
included intensive weekly educational courses led by a senior dermatologist with expertise in dermatopathology. For example, the course required students to distinguish between slides of normal and diseased tissues. For the neoplasm category, the remedial teaching strategy was designed to improve skills in describing pathological patterns, in making pattern-based differential diagnosis, and in making clinical decisions.

![Figure 2. The proposed clinical diagnosis learning system](image)

**Two-tier diagnostic assessment**

For each participant, learning effectiveness was measured using the two-tier diagnosis assessment technique proposed by Huang et al., (2009). The score for each question was calculated according to the difficulty of the question. The equation used to calculate the score was as follows:

\[
S = \frac{\sum_{j=1}^{n} L_j}{\sum_{i=1}^{k} T_i} \times 100, \tag{1}
\]

- **S**: test score
- **T_i**: difficulty of a given question \( i \) scored on a scale from 1 (least difficult) to 5 (most difficult)
- **i**: range, 1 to \( k \)
- **k**: number of test questions
- **L_j**: difficulty of question \( j \) in two-tier diagnosis assessment scored on a scale from 1 (least difficult) to 5 (most difficult)
- **j**: range, 1 to \( n \)
- **n**: number of test questions answered correctly in two-tier diagnosis assessment

The difficulty of each test question was pre-determined by two domain experts. For each participant, learning efficiency was measured by comparing pre- and post-test scores. Figure 3 shows an example of a question used in the two-tier diagnosis assessment by computerized adaptive testing (CAT) (Kuo & Wu, 2013). The first step of the
assessment required the learner to identify the disease based on a given image. The second step required the learner to justify the diagnosis made in the first step. The following question was selected according to whether or not the previous question was answered correctly. For example, if the learner incorrectly answered a test question with a difficulty level of 3, a test question with a difficulty level of 2 was selected; if the learner correctly answered the question, a test question with a difficulty level of 4 was selected.

Based on their scores, the learning system rated the proficiency of the participants from level 1 to level 5 as follows: level 1 (score 0 to 20), level 2 (score 21 to 40), level 3 (score 41 to 60), level 4 (score 61 to 80), and level 5 (score 81 to 100).

Figure 3. Screenshot of two-tier diagnostic assessment used for one test question

Figure 4 shows the post-test results entered in the learning portfolios, which were used to assess the development of diagnosis skills. Each learner took the post-test and post-questionnaire after using the clinical diagnosis learning system for 120 hours. The learning system first asked level 5 questions related to inflammation, which had been difficult for most students in the past. The learners were required to answer twenty test questions about five major categories of diseases at an appropriate difficulty level based on their previous answers. For each two-choice question, the system gave the difficulty level of the question and the disease category. Students could review their learning portfolios to determine what dermatology concepts needed further study. For each learner, a post-test score was calculated using Eq. (1) as shown below:

$$ S = \frac{[(2+3+4+5+3+4+5+4+5+4)]}{(5+3+2+3+4+5+5+4+3+4+5+4+5+5+4+5+4+5+4+5+4+3)} \times 100 $$

$$ = \left[ \frac{44}{83} \right] \times 100 = 53. $$

The numerator is the sum of the difficulty levels of the test questions answered correctly in the two-tier diagnosis assessment ($n = 11$); the denominator is the sum of the difficulty levels of the questions ($k = 20$). In this example, the learner has a post-test score of 53, which is equal to level 3.
Measuring tools

The questionnaire survey was administered in two parts. The first part collected personal and academic data, and the second part assessed perceptions of the proposed learning system in four dimensions: perceived ease of use, perceived usefulness, use intention, and learning satisfaction. The 20 questionnaire items included 5 perceived ease of use items developed by the authors in this study, 5 perceived usefulness items developed by Goodhue and Thompson (1995), 5 use intention items developed by Davis (1989), and 5 learning satisfaction items developed by Lewis (1995).

The four dimensions were measured on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Moreover, all items were modified and adapted through an iterative personal interview process with three dermatology experts, including two senior medical doctors and one professor of dermatology, to verify the completeness, wording, and appropriateness of the instrument and to confirm its content validity. All four dimensions had high Cronbach α values (0.94, 0.91, 0.92 and 0.93 for perceived ease of use, perceived usefulness, use intention, and learning satisfaction, respectively), which indicated that the questionnaire had high reliability (Cohen, 1988).

Research results and discussion

Table 1 shows the basic statistical data contained in the learning portfolios for the 11 resident doctors who had completed the learning activity within the required time. For each participant, all data were recorded in the database, including “online learning attempts,” “learning hours,” “questions viewed,” and “mean attempts per question.” “Online learning attempts” was the number of times the participant had engaged in an online learning activity during the period of the study. Cumulative online learning time was recorded in the “learning hours” field. The number of questions viewed during the learning time, including questions viewed repeatedly, was also recorded in the “questions viewed” field while the mean number of times the participant had viewed the question was indicated in
the “mean attempts per question” field. The average values for online learning attempts, learning hours, and questions viewed, and mean attempts per question were 12.73, 31.81, 463.3, and 1.97, respectively. Not all participants completed the minimum 120 hours of learning activities before the deadline due to their other professional responsibilities (average, 31.81 hours). The participants averaged at least ten online learning attempts and 30 learning hours for the five categories of dermatological diseases (infection, neoplasm, autoimmunity, inflammation, and genetic diseases). The statistical data further showed that, as the online learning attempts and the learning hours increased, the mean attempts per question decreased. Therefore, the participants accurately identified the questions that were problematic. The system repeatedly showed the question until the participant answered correctly. Therefore, as the learning hours increased, the misconceptions about skin diseases decreased.

### Table 1. Descriptive statistical data for online learning activity of the participants

<table>
<thead>
<tr>
<th>Participant</th>
<th>Online learning attempts</th>
<th>Learning hours</th>
<th>Questions viewed</th>
<th>Mean attempts per question</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>12</td>
<td>31.17</td>
<td>433</td>
<td>1.88</td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>40.25</td>
<td>510</td>
<td>1.36</td>
</tr>
<tr>
<td>C</td>
<td>11</td>
<td>38.23</td>
<td>556</td>
<td>2.15</td>
</tr>
<tr>
<td>D</td>
<td>13</td>
<td>35.19</td>
<td>542</td>
<td>1.87</td>
</tr>
<tr>
<td>E</td>
<td>12</td>
<td>33.34</td>
<td>477</td>
<td>1.96</td>
</tr>
<tr>
<td>F</td>
<td>11</td>
<td>28.13</td>
<td>501</td>
<td>2.53</td>
</tr>
<tr>
<td>G</td>
<td>11</td>
<td>25.56</td>
<td>434</td>
<td>2.45</td>
</tr>
<tr>
<td>H</td>
<td>14</td>
<td>33.27</td>
<td>465</td>
<td>1.77</td>
</tr>
<tr>
<td>I</td>
<td>13</td>
<td>29.46</td>
<td>389</td>
<td>2.19</td>
</tr>
<tr>
<td>J</td>
<td>12</td>
<td>29.88</td>
<td>414</td>
<td>1.95</td>
</tr>
<tr>
<td>K</td>
<td>16</td>
<td>25.45</td>
<td>375</td>
<td>1.56</td>
</tr>
</tbody>
</table>

Average 12.73 31.81 463.3 1.97

Table 2 shows the results of a paired-sample t test used to compare learning performance between pre- and post-test. The T test results showed that the proposed medical diagnosis learning system significantly improved learning performance \((T = -3.42, p < .01)\); that is, the use of meaningful photographs and two-tier assessment technique significantly improved the diagnosis skills of the participants. Notably, the system used photographs to enhance learning efficiency. According to information processing theory, photographs contribute to meaningful learning by providing graphic representations of abstract or implicit concepts and by showing relationships among concepts. Photographs also help learners to organize ideas, to link new information with prior knowledge, and to store and retrieve data. That is, photographs help learners construct meaning (Epçaçan et al., 2010) by visually representing the meaning they construct from reading. Thus, the visual representations used in the study enhanced learning in the participants by facilitating them in recognizing associations between skin disease and pathology. Studies also indicate that the two-tier diagnosis assessment is useful for formative assessments because it can be used to assess understanding by students and because it encourages them to think about concepts rather than memorize facts (Chandrasegaran et al., 2007; Tsui & Treagust, 2010; Mutlu & Sesen, 2015). The two-tier diagnosis assessment stimulates discussion and thinking and provides meaningful information about the level of understanding in students.

### Table 2. Results of paired-sample t-test of learning performance in resident dermatologists \((N = 11)\)

<table>
<thead>
<tr>
<th>Test</th>
<th>SD</th>
<th>Mean</th>
<th>SE</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>9.95</td>
<td>47.73</td>
<td>3.00</td>
<td>-3.42**</td>
</tr>
<tr>
<td>Posttest</td>
<td>16.03</td>
<td>60.45</td>
<td>4.83</td>
<td></td>
</tr>
</tbody>
</table>

Note. **p < .01.

The questionnaire survey results for the four dimensions assessed by the questionnaire survey (Table 3) indicated that over 90% of the participants agreed that the learning system had high usefulness and high ease of use. Additionally, over 90% of the participants were highly satisfied with the effectiveness of the learning system in correcting their misconceptions about dermatology. The participants strongly agreed that the system maintained their interest in learning dermatology and that they would be interested in using the system in the future. These results indicate that the proposed diagnosis learning system can facilitate resident doctors in improving their knowledge and diagnosis skills in dermatology.
Table 3. Descriptive statistics for four dimensions of user acceptance

<table>
<thead>
<tr>
<th>Dimension/Items</th>
<th>1(N/%)</th>
<th>2(N/%)</th>
<th>3(N/%)</th>
<th>4(N/%)</th>
<th>5(N/%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived usefulness (Cronbach α = 0.94)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. I can improve my learning efficacy by using the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>2. I can improve my learning performance by using the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>3. I find the system useful.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>6/55%</td>
<td>5/45%</td>
</tr>
<tr>
<td>4. I have improved my diagnosis skills by using the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>6/55%</td>
<td>7/45%</td>
</tr>
<tr>
<td>5. I have improved my overall diagnosis skills by using the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>Perceived ease of use (Cronbach α = 0.91)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. The system is easy to use.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>2. The system is easy to learn.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>3. The functions of the system can be applied quickly and easily in my work.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>1/9%</td>
<td>7/64%</td>
<td>3/27%</td>
</tr>
<tr>
<td>4. The instructions for operating the system are easy to understand.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>1/9%</td>
<td>7/64%</td>
<td>3/27%</td>
</tr>
<tr>
<td>5. Learning how to use the system is not time-consuming.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>1/9%</td>
<td>7/64%</td>
<td>3/27%</td>
</tr>
<tr>
<td>Learning satisfaction (Cronbach α = 0.92)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. I clearly understand the relationship between skin symptoms and pathology when using the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>2. I could control my learning path and progress when using the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>1/9%</td>
<td>7/64%</td>
<td>3/27%</td>
</tr>
<tr>
<td>3. I could identify my misconceptions by using the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>4. Using the system was a good decision.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>5. Overall, I am very satisfied with the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
<tr>
<td>Intention to use (Cronbach α = 0.93)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. I will continue using the system to improve my diagnostic skills.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>1/9%</td>
<td>7/64%</td>
<td>3/27%</td>
</tr>
<tr>
<td>2. I intend to increase my use of the system.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>1/9%</td>
<td>7/64%</td>
<td>3/27%</td>
</tr>
<tr>
<td>3. I intend to use the system to improve my learning performance in other disciplines.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>1/9%</td>
<td>7/64%</td>
<td>3/27%</td>
</tr>
<tr>
<td>4. I will use the system frequently to improve my diagnostic skills.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>1/9%</td>
<td>7/64%</td>
<td>3/27%</td>
</tr>
<tr>
<td>5. I hope the system is applied in other disciplines.</td>
<td>0/0%</td>
<td>0/0%</td>
<td>0/0%</td>
<td>7/64%</td>
<td>4/36%</td>
</tr>
</tbody>
</table>

The quantitative analysis results showed high scores for all four dimensions of user acceptance. For further evaluation of the effectiveness of the learning system for developing diagnostic skills, six participants were randomly invited to interviews after completion of the program. Similarly, the interpretations of the six participants were arranged according to the four dimensions. The interview data were coded according to the convention for qualitative research. For example, the code “A-2-20140210” indicated the second sentence stated by resident A, who was interviewed on February, 10, 2014. The qualitative analysis results are summarized below.

**Perceived usefulness**

The participants agreed that the system was helpful for identifying their misconceptions about dermatology before they began their residency. Examples of statements by the participants included the following: *I found that the system helped me to recognize my misconceptions about dermatology (A-2-20130110, B-1-20130220); The learning system showed me actual photos of skin diseases that I had difficulty understanding (C-2-20130308, F-2-20130601); The system identified mistakes that I made repeatedly (D-5-20130406).*
Perceived ease of use

The participants agreed that viewing their diagnostic results was very easy. Statements by the participants included the following: Since the learning system has a user-friendly interface, I can begin practicing immediately (A-3-20130110, E-3-20130503, F-4-20130601); The system provides clear instructions for users (C-4-20130308, B-2-20130220); The learning system clearly shows the difficulty level of the materials I am studying (D-3-20130406).

Learning satisfaction

The interviewed participants agreed that the system is helpful for clarifying misconceptions and for enhancing learning performance. Statements by participants included the following: By using the system, I quickly understood the relationships between skin symptoms and pathology (C-3-20130308, D-4-20130406); By using the system, I clearly identified my misconceptions about skin symptoms. For me, the system was an effective tool for learning diagnosis in the dermatology domain (A-1-20130110, B-3-20130220, E-2-20130503, F-3-20130601).

Intention to use

Finally, the six interviewed participants indicated that they would be interested in using the system to learn dermatology in the future. Statements by the participants included the following: The diagnosis mechanism is very helpful, and I plan to use it in the future (A-4-20130110, B-4-20130220, F-5-20130601); I would definitely encourage other resident doctors to use the diagnosis system (C-5-20130308, D-6-20130406, E-4-20130503).

Conclusions and future works

Previous studies indicate that advanced electronic learning technologies can improve learning performance and learning attitude in various domains. However, few studies have focused on the misconceptions of medical students in the clinical diagnosis domain. This research addressed this issue by developing a medical diagnosis learning system for resident doctors in dermatology. The system was designed to increase accuracy in diagnosing skin diseases during residency training. The effectiveness of the proposed system was investigated in 11 resident doctors in dermatology by using a one-group pretest and posttest experimental design. The research findings showed that the proposed diagnosis learning system significantly improved learning performance in the resident doctors between pre- and post-test. In terms of the perceived usefulness and the perceived ease of use of the proposed system, the participants strongly agreed that the system facilitated their recognition of various symptoms of skin diseases. Most participants also agreed that the system was easy to learn and easy to operate. Thus, they had high satisfaction with the system and high intention to use it in the future, not only in dermatology, but also in other disciplines. Although the study included only 11 dermatology residents, the authors will eventually implement the system for use by PGY doctors, interns, and clerks that rotate to dermatology training programs. The difficulty of the questions will be adjusted according to the training and experience of the participants.

The proposed medical diagnosis learning system provides users with individual portfolios, which were designed by applying two-tier diagnosis assessments. Users can review their portfolios to identify their misconceptions and their expertise levels in various skin diseases. The system provides an adaptive learning environment in which the difficulty of each question is adjusted according to the answer for the previous question (Brusilovsky, 1999; Specht et al., 2002) and learning materials are tailored to the current knowledge of the user. That is, by reviewing their learning portfolios, resident doctors can evaluate their expertise in different categories of skin diseases and can identify concepts that need further study. Data collected in dermatology courses and stored in student learning portfolios can also be used by educators to identify common misconceptions of students so that appropriate
complementary materials can be prepared for remedial teaching. The research results are consistent with previous studies (Meyer et al., 2012; Yasin et al., 2012).

Some limitations of this study are noted. The number of dermatology residents recruited for this study was low because the Ministry of Health and Welfare in Taiwan limits the number of medical students in dermatology and because recruitment of the participants was limited to the Kaohsiung area of south Taiwan, where the number of dermatology residents is relatively low. In this context, this research result should not be excessive inferences to residents of other areas. Additionally, the significant difference in clinical skills observed in each participant before and after the education intervention cannot be considered definitive evidence of the effectiveness of the learning system. Future studies should consider using a delayed post-test design to investigate the learning retention of students.

In future studies, the researchers will further validate the effectiveness and reliability of the proposed system in a larger population of dermatology residents recruited from central and northern areas of Taiwan. The quasi-experimental design will be employed to examine the effectiveness of the proposed learning system in two groups, i.e., an experimental group using the proposed learning system and a control group using a conventional e-learning system. After the participants complete the learning activities, objective data, including qualitative, quantitative, and focus group data, will be collected for comprehensive analyses.

Acknowledgements

This study was supported in part by the Ministry of Science and Technology of the Republic of China under contract number NSC 101-2320-B-037-022. The dermatology residents analyzed in this study are also appreciated for their active participation and commitment to professional improvement. Fan-Ray Kuo and Yi-Ying Chin had contributed equally on this paper.

References


Finn, J. (2010). E-learning in resuscitation training—students say they like it, but is there evidence that it works? *Resuscitation, 81*, 790-791.


Fostering Creativity in Tablet-Based Interactive Classrooms

Hye Jeong Kim1, Ji Hyeon Park2, Sungae Yoo3 and Hyeoncheol Kim4*

1Center for Teaching and Learning, Chung-Ang University, Korea // 2Hansol Middle School, Korea // 3Icheon Songjeong Middle School, Korea // 4Department of Computer Science and Engineering and Department of Computer Science Education, Korea University, Korea // hyejeongkim@cau.ac.kr // gina307@hanmail.net // sungaey@gmail.com // harrykim@korea.ac.kr

*Corresponding author

(Submitted November 11, 2014; Revised July 26, 2015; Accepted October 20, 2015)

ABSTRACT

This article aims to examine the effects of an instructional model that leverages innovative technologies in the classroom to cultivate collaboration that improves students’ comprehension, fosters their creativity, and enables them to better express and communicate their ideas through drawing. This discussion focuses on classroom interaction systems and technologies that can foster creativity, including tablets, electronic blackboards, interaction management solutions, and high-speed wireless internet. To address the study aims, we conceptualized an instructional model entitled “Visual Thinking through Tablet-based Classroom Interaction” (VTTCI) and applied it for one semester in English classes in a high-tech middle school in Korea. This model aimed to enable students to better express and communicate their ideas through drawing; to understand the lesson material’s text and topic; to conceive of core ideas related to the topic; to visually express concepts through drawings, presentations, and discussions; and to ultimately gain a better understanding of the material based on the drawings. After applying the instructional model for one semester, we assessed the creativity of the students in the six treatment classes using the Torrance Tests of Creative Thinking-Figural form (TTCT), and we compared their scores with those of the students in the six control group classes, who did not receive the treatment. The results revealed that the students in the treatment group scored significantly higher on the Originality, Abstractness of Titles, and Elaboration subscales and on overall creativity. Thus, the VTTCI model can be seen to affect creativity. The results show that the students’ creative proficiency improved with the adoption of the instructional model for tablet-based interactive classrooms.

Keywords

Learner-generated drawing, tablet computer, tablet-based interactive classroom, creativity, Torrance Tests of Creative Thinking, instructional model

Introduction

With the advent of information and communication technology, educators have become increasingly interested in adapting interactive and mobile technologies for classroom curricula and learning activities (Enriquez, 2010; Ifenthaler & Schweinbenz, 2013; Koile & Singer, 2008; Loveless, 2002). Interactive technologies, including smartphones, tablets, interactive whiteboards, and classroom response systems, have generated active discussions that focus on students’ learning performance and abilities (Gikas & Grant, 2013). Recently, tablet computers have caught educators’ attention because their versatility and mobility may afford students active and collaborative experiences. A tablet’s features and capabilities allow students to access course content and to interact with their teachers and classmates in new ways (Gikas & Grant, 2013). For example, students in a mathematics class can transmit their solutions through their individual tablets to an electronic board to review the answers with their classmates and teachers. Additionally, teachers can interact more with individual students during a limited class period, adapt activities quickly in response to pupils’ responses, and use students’ errors and misconceptions as a teaching moment for the entire class (Muijs & Reynolds, 2011). Recently, the education field has begun actively researching ways of leveraging technology-supported interactive environments to foster students’ creative experiences during classroom activities (Koile & Singer, 2006; Koile & Singer, 2008). Students can brainstorm ideas together, develop various alternatives to solve problems, share ideas and work, and develop solutions using various technological tools, such as 3D printers, painting tools, cloud-based notebooks, composition tools, and interactive management solutions for tablets and electronic boards. In particular, the relationships between the distinctive features of interactive technologies (tablet computers) and the characteristics of creativity open up new approaches and strategies for cultivating creativity in education.
Tablet-based interactive classroom technologies have the potential to generate new ways of fostering creative experiences through drawing, seeing, expressing, learning, and engaging with instructional situations. These innovative technologies help teachers offer new ways for students to practice drawing skills as a creative activity to increase their creative understanding (Ainsworth, Prain, & Tyler, 2011). Thus, in the present study, the researchers utilized an instructional model designed to foster creativity. The model allows students to develop their own drawings and interact with others in a high-tech, interactive classroom. This model seeks to help students more creatively comprehend and express what they have learned.

The current study aims to examine the effects and benefits of a model entitled “Visual Thinking through Tablet-based Classroom Interaction” (VTTCI) in terms of students’ creativity after one semester of use. We further seek to better understand the effects of the model on students’ creativity by examining the creativity subscales of the Torrance Tests of Creative Thinking-Figural (TTCT-Figural, Torrance, 1990; Torrance, 1998) and by evaluating the differences between the treatment and control groups. This article describes the process that we follow in our evaluation of the model and the outcomes. We begin by discussing the general benefits and shortcomings of drawing as a creative activity in a tablet-based classroom interaction environment. Then, we identify the pedagogical attributes of drawing activities in classrooms that are supported by a high-tech, interactive learning environment. Next, we introduce the instructional model utilized in the study. Finally, we reveal the findings.

**Literature review**

**Value of learner-generated drawings**

Drawing activities are known to promote students’ creative imagination, i.e., the ability to transform available and remembered data into new and original mental images (Lingvist, 2003). Vygotsky (1978) emphasized creative imagination as a precursor to human learning. He believed that acting in an imaginary situation teaches children how to behave in different situations. As a visual illustration tool, drawing can be used to make an experience more meaningful for learners, to increase their understanding, and to support their communication with peers (Ainsworth et al., 2011; Madsen, 2013; Van Meter, 2001). In schools, drawing is frequently used as a learning tool. However, a shift to reading and writing activities occurs when students enter secondary schools, and drawing seems to be absent, except within visual arts programs (Madsen, 2013). In traditional classrooms, learners focus on interpreting the materials presented in a textbook or lesson rather than on creating their own visual representations (Cox, 1999; Van Meter, 2001). Thus, in the progressive stages of educational development, learners have fewer opportunities to develop their own visual interpretations to promote their understanding (Ainsworth et al., 2011).

Children, who are naturally visual learners, can internalize and understand the relevance behind what they draw, and they can link their drawings to other ideas or objectives through image making (Davis, 2005). Children can draw and create visual images that communicate something to others. According to Armstrong (1994), children’s visual learning may include visualization, color cues, picture metaphors, concept maps, sketches, diagrams, and graphic symbols. The emphasis on learners’ visualization skills helps explore new ways of enhancing the visual literacy of learners (Ainsworth et al., 2011; Madsen, 2013). Thus, the classroom teaching objective of increasing visual literacy provides interesting opportunities for children to actively participate in the learning process (Madsen, 2013).

In a structured classroom setting, learner-generated drawing is a strategic process in which drawings are created to achieve a learning goal. Drawings and illustrations (e.g., street signs or instructional handbooks) are often used to help people understand complex problems or to explain concepts to others (Leenaars, Van Joolingen, & Bollen, 2013; Van Meter, 2001; Van Meter, Aleksic, Schwartz, & Garner, 2006). Researchers in various disciplines, including science education (Ainsworth et al., 2011; Leenaars et al., 2013), mathematics education (Uesaka, Manalo, & Ichikawa, 2007), music education (Bennett, 2013), and language arts/communication (Madsen, 2013; Van Meter & Garner, 2005), have addressed the benefits of visualizing ideas and thoughts (Davis, 2005; Roam, 2009). Drawing is classified as a strategic process because it is goal directed, enhances the organization of knowledge, and potentially improves learning outcomes (Paris, Lipson, & Wilson, 1983). Learners can turn abstract or complex relationships into more concrete and explanatory relationships (Leenaars et al., 2013).

Research findings have shown that drawing can lead to the construction of a mental model and to improved engagement and achievement (Ainsworth et al., 2011; Van Meter & Garner, 2005). Van Meter (2001) asserted that
drawings can be supported or unsupported by external aids to help students create accurate drawings or develop more ideas during the drawing process, such as pre-supplied backgrounds and cutout figures to render more accurate visual representations. Drawings, as external illustrations, can potentially inform learners’ internal biases. According to Ainsworth et al. (2011), drawing activities in science education can enhance students’ engagement and deepen their understanding by allowing them to generate their own representations to create and record information. Drawing can provide further opportunities for active learning in science classes. Students can use drawings to explore, coordinate, and justify their understandings, which can increase their motivation. Van Meter and Garner (2005) reviewed the scientific literature on learner-generated drawing in the classroom. They discovered that learner-generated drawing can improve observational processes and imagination in terms observing details and subtle properties. In particular, drawings can help students engage in higher-order thinking (e.g., creativity and problem solving) and promote a deeper understanding of the instructional materials. Furthermore, teachers can assess students’ grasp of the content and determine their misconceptions based on their drawings. Chan and Zhao (2010) reported a statistically positive correlation between drawing skills and creativity among children, adolescents, and young adults; in addition, they predicted how drawing skills correspond with creativity throughout different stages of development. Creativity is a complex and multidimensional phenomenon that is meaningful for a learner’s education (Dewulf, Baillie, & Britain, 1999; Sternberg, 1999; Treffinger, Young, Selby, & Shepardson, 2002). In addition, creativity and learning are clearly connected and involve imagination and imitation (Spencer, Lucas, & Claxton, 2012). Moreover, Clark and Zimmerman (2004) found that drawing skills and creativity were moderately correlated.

Educators and researchers have indicated that incorporating technology into the classroom can enhance different types of drawing projects within different academic areas by improving the communicability and representational purposes of the final illustration. According to Ainsworth et al. (2011), drawing—particularly when linked to new technologies in science education—has considerable potential and several purposes in the science classroom, such as understanding complex visualization environments, creating models and interacting with scientific systems, and broadening the existing concepts of drawing via the use of smart mobile devices or multimedia tools. Tablet computers represent an innovative technology that can be used to develop visual depictions that improve communication and representations in classroom activities. For example, popular tablet computers (e.g., the Android-based Samsung Galaxy Note or the iOS-based Apple iPad) support photo and movie applications, drawing applications, painting applications, and editing applications.

**Creative activities supported by interactive technologies**

Creative activities are often related to people’s sharing behaviors. People often share and exchange information and ideas in an iterative way to accomplish certain tasks. When students share ideas, knowledge, or opinions, some degree of interaction between individual students is needed. This interaction can occur through a variety of means, including face-to-face communication, written correspondence, and other methods. Interaction via computers is considered to have positive effects on instruction and learners’ satisfaction and performance (Zirkin & Sumler, 1995). Likewise, technology-supported interaction in the classroom has been emphasized in instructional design as an essential method and tool for education (Ohl, 2001). According to Zirkin and Sumler (1995), highly interactive classroom instruction is positively correlated with better student performance. The technology used to generate classroom interaction continues to evolve, resulting in a constant need for new studies and an understanding of interactive classrooms, such as the recent drive to examine ways of fostering creativity in high-tech, interactive learning environments.

Interaction is known to be a fundamental process that promotes creative performance and outcomes in social contexts, such as schools (Fischer, 2004; Fischer, Giaccardi, Eden, Sugimoto, & Ye, 2005; Sosa, 2011). According to Amabile’s (1983) componential theory of creativity, creativity is the generation of novel and valuable ideas that involve better products, processes, practices and problem solving. Amabile and Mueller (2002) asserted that creativity consists of four components: intrinsic motivation, domain-relevant knowledge, creativity-relevant skills, and social interactions. Additionally, exposure to heterogeneous knowledge through interactions with others can improve individuals’ creative potential and a team’s innovativeness because people’s behavioral tendencies in terms of the sharing of knowledge or ideas can promote or inhibit creative performance (Tiwana & McLean, 2005).
Typically, people are more likely to generate creative ideas when they can access a diverse range of knowledge and information by interacting with other people with dissimilar expertise (Sosa, 2011). Thus, people’s knowledge- and experience-sharing behaviors are regarded as important components that allow diverse expertise to be transferred to others (Tiwana & McLean, 2005). In terms of the sharing process among classroom peers, students can share their know-how, the technical elements of their crafts and skills, their viewpoints, their understanding, and/or their creative ideas with their classmates. Recently, these collaborative activities have been more emphasized in relation to social creativity, which is the process that leads to the emergence and sharing of creative activities and meanings in a socio-technical environment (Fischer et al., 2005). Thus, the sharing of knowledge or ideas requires close interactions among individuals to foster a shared understanding that helps them utilize others’ knowledge or their creative imaginations (Hansen, 1999; Lam, 2000). However, the typical school education system in Korea, where a teacher and many students interact in a classroom for a short period (e.g., 40 or 45 minutes), has its limits. To promote student creativity in a social context, instructional activities must enable students to create and combine diverse solutions, share ideas, search for additional information, and change their way of thinking.

Recently, after numerous discussions, some classrooms in Korea were transformed by interactive systems consisting of electronic boards, tablets, and integration solutions. These new classrooms were adapted at specially designated experimental schools, where students can experience diversified learning approaches based on innovative research models and high-tech environments (see Methods). The classroom interaction system supports teacher-to-student and student-to-student communication as well as the electronic sharing of information or materials. This innovative system can support various instructional strategies that cannot be realized in traditional classrooms. Researchers (Enriquez, 2010; Koile et al., 2007) have stated that tablet computers will potentially change the dynamics of classroom interaction by combining wireless communication and pen-based computing technology. One benefit that is often espoused in discussions is the use of tablet-based systems to facilitate students’ abilities to analyze and solve problems. Recent tablet computers, also called tablets, have additional functionality. In particular, a tablet’s touchscreen can simulate paper; thus, students can create documents or use their finger or a stylus as a “pencil” and write directly on the screen as though it were a notepad. The enhanced functionality of recent tablets is more suitable and convenient than the features of personal computers, particularly for solving, analyzing, and displaying problems that require individual sketches, diagrams, and multimedia files. A tablet-based classroom interaction system with wireless connectivity has the potential to provide an ideal environment for classroom activities, such as collaborative assignments and individual presentations. In traditional classrooms, such projects cannot be performed with a similar level of ease in terms of displaying teachers’ or students’ work to the class.

One possible beneficial function of the interactive system is the enhancement of visual representations through their immediate presentation to the wider class, which in turn can foster learners’ creativity in classroom activities through shared interpretations and collaboration to refine an idea. As a classroom activity, drawing can enhance student creativity, particularly through playful interactions with computers, which may lead to students’ preferred writing projects, such as stories (Fleith, 2000). Although many attempts have been made to examine the effects of high-tech interactive environments in the classroom, few studies have examined the extent to which creativity has been enhanced or inhibited in such interactive contexts in an educational environment, particularly a tablet-based interactive environment.

Based on the reviewed advantages of integrating technologies in classroom activities to promote creativity, an instructional model and technological framework for the model have been developed. Although the literature addresses the importance of technologies in supporting creative activities, discussion is limited regarding the influence of mobile and interactive technologies on the sharing of creative outputs in the classroom. This study examines middle school students’ experiences in a high-tech classroom environment that is conducive to collaborative drawing activities. Additionally, we assess the quality of students’ creativity by administering figural creativity tests (TTCT-Figural). Thus, this study addressed the following research questions: What is the VTTCI model? Is there a significant difference between the creativity scores of students in the treatment group and those in the control group that relates to the use of VTTCI? What practical implications regarding VTTCI use can be issued to educators?
Methods

Participants

The participants in this study were 262 seventh-grade students (boys = 121, 46.2%, and girls = 141, 53.8%) in South Korea. The students ranged in age from 13 to 14 years old. The target school was located in a city that the Korean government recently built as an administrative capital city to facilitate balanced regional development. All seventh-grade students in this school participated in the study. The school is a newly built high-tech school in a suburban area. The treatment group (n = 137) consisted of 73 girls (53.3%) and 64 boys (46.7%) from six classrooms. The control group (n = 125) consisted of 68 girls (54.4%) and 57 boys (45.6%) from six classrooms.

We examined students’ motivation in terms of their self-efficacy, intrinsic value and technology use to gauge the groups’ equivalence before the treatment, as individual creativity is influenced by motivation factors that arise from interest, enjoyment, satisfaction, and deep involvement in the work (Amabile, 1985; Amabile, 1993). People are most creative when they feel intrinsically motivated rather than externally motivated. We adopted the self-efficacy and intrinsic value subscales of the Motivated Strategies Learning Questionnaire (Pintrich & De Groot, 1990) to examine the groups’ equivalence. Students were instructed to respond to the items on a 5-point Likert scale (1 = almost never to 5 = almost always). The self-efficacy scale (α = .911) was created by taking the mean score of the students’ responses to nine items concerning their beliefs about their classwork performance. The intrinsic value scale (α = .866) was created by taking the mean score of the students’ responses to eight items concerning their intrinsic interest in and the perceived importance of course work, their preferences regarding challenges and their mastery goals (Pintrich & De Groot, 1990). Both scales showed no statistically significant differences between the treatment and control groups (ps > .05).

To examine the technology fluency of the students in both groups, we adopted five items related to middle school students’ possession of technology (devices, computers and smart phones), technological fluency (skills in using computers and painting tools) and information searching experience. The results indicated that there was no significant difference between the treatment and control groups (ps > .05) in terms of technology possession and technological fluency; however, there was a difference in information searching experience (p < .01). Because information searching was not a required skill for the classroom activity targeted in this study, we assumed no significant difference between the treatment and control groups.

Measure

To measure students’ creativity, the Korean version of the TTCT-Figural, (Torrance, 1990; Torrance, 1998; Kim, 2010a; Kim, 2010b) was administered. The TTCT is the most widely used and accepted method of measuring creativity (Cramond, Matthews-Morgan, Torrance, & Zuo, 1999). The Korean version of the TTCT is a well-developed instrument that has been widely used in experimental studies for over ten years. The TTCT asks students to perform tasks in which they need to complete a series of drawings using lines to tell a story. Students must also name their drawings. The TTCT-Figural consists of three incomplete pictures that the students complete individually to assess different aspects of creative functioning. The TTCT has five norm-referenced criteria for measuring creativity, including fluency (the number of relevant ideas), originality (the number of statistically infrequent ideas), elaboration (the number of added ideas), the abstractness of titles (the degree of abstract thinking), and the resistance to premature closure (the degree of psychological openness). The TTCT-Figural is considered a reliable and valid measurement of creativity (Kim, 2006; Torrance, 1998; Torrance & Wu, 1981). In addition to the six subscale scores, the “Creativity Index” (CI) can be used as a composite score for the TTCT. Raw scores are converted into standard scores, and the standard scores for each of the five norm-referenced measures are averaged. The frequency of creative strengths is added to the averaged standard scores to yield a CI, an overall indicator of creative potential (Torrance, 1998).

Classroom interaction system

The middle school that participated in this study opened in 2012 as a high-tech school in a new suburban town in South Korea. Each classroom has an electronic board with multiple input tools (camera, digital visualizer,
microphone, and student tablets), an Android-based 7-inch tablet for each student and teacher, a personal computer as a classroom server, an interaction management solution (iKaist™ Schoolbox, http://www.i-kaist.com), high-speed internet access with wireless hubs, access to private cloud servers in the school district, and a radio frequency identification (RFID) reader. The interaction management solution supports real-time file transfers between students’ tablets and the teacher’s electronic board and between individual students’ tablets (see Figure 1). Students and teachers can transfer instructional materials, worksheets, images, and video files in real time and share the pictures or notes that the students create. Teachers can also monitor all tablet activities and control the screens. For this study, the students used their tablets to share and monitor learner-generated drawings during classroom activities. Students and teachers focused on using the interaction management solution with their individual tablets.

Figure 1. Classroom interaction system’s architecture, consisting of tablet computers and an interactive electronic board

Model of Visual Thinking through Tablet-based Classroom Interaction (VTTCI)

Research on the effectiveness of methods or programs in developing creativity most frequently takes the form of interventions conducted with adults. By contrast, programs that stimulate students’ creativity are often based on creative activity in language or artistic activities in group or individual work (Dziedziewicz, Oledzka, & Karwowski, 2013). Various types of creativity training for children have been confirmed as effective (Ma, 2006). According to the meta-analysis of Scott, Leritz, and Mumford (2004), creativity training for children tends to be in the areas of analogies, open idea production, interactive idea production, creative processes, imagery, computer-based production, structured idea production, analytical thinking, critical/creative thinking, situated idea production, and conceptual combination.

In this study, we developed a creative thinking training model aimed at fostering middle school students’ visual representation skills to communicate with others using drawings and the classroom interaction system. Although applicable regardless of the class duration and size, the model was particularly designed to function within a short class period (40–50 minutes) and a medium-sized class (20–30 students) and to increase learners’ creative potential, particularly their visual creativity. This study posits that highly interactive classroom environments that use innovative technologies can provide various opportunities for students to engage in classroom activities. An environment that fosters deeper bonds among students and between students and their teachers supports greater engagement in classroom activities. In traditional middle school settings, each class period lasts 45 minutes, and teachers must plan lessons that fit within this time constraint. Thus, the traditional classroom approach, which inherently offers little time, does not provide ample opportunities for students to interact with teachers or with their peers during classroom activities. By contrast, technologies that support efficient classroom activities can establish highly interactive classroom environments. During many classroom activities, students do not have the opportunity to share and adapt their learning. Using technologies, for instance, mirroring tables, transferring files in real time, and projecting instructional objects, to facilitate classroom interaction can offer innovative and creative experiences to students.

The developed model focuses on identifying problems, understanding relationships, generating ideas, developing visual representations, and communicating concepts visually with teachers and other students. Principally, students understand the problems and relationships before they begin drawing. In this model, students are asked to represent their understanding or solutions through drawings and short written descriptions about their drawings. Students share
their drawings with teachers and classmates after representing the problem statement with visual elements and illustrated relationships. Then, students elaborate on their drawings and interpretations. During these activities, classroom technologies such as cameras (to take pictures of the drawings) and tablet computers (to render the drawings) are often used. The model utilizes a classroom activity that includes five stages: introduction, observation, representation, sharing and communication, and internalization.

(1) During the introduction stage, the teacher provides instructions for the learner-generated drawing activity. The teacher explains and demonstrates how to use the specific tools, particularly the tablet computers and the classroom interaction system (see Figure 2).

(2) During the observation stage, students are asked to read and decipher the designated reading and figures from the textbook, which includes stories with several themes, such as sports, cooking, and traditional festivals. Then, the teacher contextualizes the stories and provides additional reading resources to help students identify the learning topic. Meanwhile, students are allowed time to observe and think about the given topic and the context. Students are then asked to identify the key elements of the given texts and the ways in which the elements relate to one another. Finally, students are asked to graphically represent the key elements and their relationships on a tablet screen or on paper. In the meantime, the teacher tells the students that there are many possible ways to visualize the relationships and encourages them to come up with a creative idea.

(3) During the representation stage, students draw visual elements and individually render the relationships that connect these elements. The teacher emphasizes the students’ creative ways of representing the elements and, when needed, provides additional instructions on how certain elements might be modified or better represented, reminding students of the requirements of this relational rendering. Students freely draw to communicate their ideas with their classmates and with the teacher, and the teacher does not limit students’ drawing techniques or renderings of the relationships among the elements. At the beginning of these activities, students are allowed to practice formulating keywords and explanatory labeling for sample messages. Students are encouraged to use clear and concise labels and words in their visual representations to transfer the messages that they intend to communicate to their classmates. When students complete their initial attempts at visual representation, the teacher provides feedback, often using other students’ drawings from previous classes as examples that better achieve the assignment’s goals. For the actual task of drawing, students are given the option to use either paper and pencil or the drawing function on their tablet computers (see Figure 3).

(4) During the sharing stage, students are asked to share their drawings with either the entire classroom or a group using the classroom interaction system, which connects their tablets with an interactive white board. Students then refine their illustrations based on other students’ comments and the teacher’s feedback. If the lesson can be
structured as a block-of-time lesson, the writing activity can be planned using individual drawings that are uploaded to the learning management system (see Figure 4).

Figure 4. Sharing drawings transferred from students’ tablet computers

![Sharing drawings](image)

Figure 5. Uploading students’ drawings to the learning management system

![Uploading drawings](image)

(5) During the final stage, i.e., internalization, the teacher leads the students in a reflection exercise. The teacher asks students to discuss how each of the drawings is similar to, or different from, their own drawings and how each drawing makes sense of or addresses any challenges in terms of the elements identified and the relationships among the elements in the context of the given text. If the students complete the drawings, the teacher can summarize the results during the class after the final versions are transferred from the students’ tablet computers. Using the tablet computers, students upload their final versions to the learning management system or the classroom board to receive additional feedback from their classmates or the teacher (see Figure 5).

Procedure

The treatment and control groups had the same course content, textbook, class hours, instructional schedule, and course instructor. The treatment group included students from six classes who were treated with the instructional model with the six learner-generated collaborative drawing lessons. The control group, consisting of students from six other classes, took the same course but were not treated with the VTTCI model. The control group was taught the same topic using traditional methods in the classroom interaction system. The control group was not informed that there was another group participating in the experiment. The two groups were taught by the same teacher for the duration of the course and were taught according to the same scope and sequence. At the end of the semester, in December 2012, students completed the TTCT-Figural. TTCT-Figural batteries were conducted during one testing session. Trained test scorers reviewed the TTCT-Figural tests, and the scores were calculated according to the instructions of Kim (2010a; 2010b) and the Manual for Scoring and Interpreting Results (Torrance, 1990). Initially, 270 students responded to both a questionnaire and the creativity test. However, 8 invalid responses were excluded, resulting in a total of 262 complete and valid responses that were subjected to data analysis.

To examine the effectiveness of the VTTCI instructional model, we compared the average creativity scores of the treatment and control groups. The statistics shown in Table 2 indicate the differences between the groups. To investigate group differences with respect to psychometrically determined figural creativity after the instructional model with learner-generated drawings, a multivariate analysis of variance (MANOVA) was conducted. The data

214
were analyzed using the general linear model in IBM SPSS software for the dependent variables of Fluency, Originality, Abstractness of Titles, Elaboration, and Resistance to Closure.

A normal quantile-quantile plot was used to check the normality assumption of the data. Box’s M test of the assumption of homogeneity of covariance resulted in a file-to-reject decision, $M = 13.162, F(15, 267286.99) = .86, p = .611$, indicating that the assumption was not violated. The independent variable was the presence or absence of an instructional model in the classroom.

**Results**

This section presents the results of the descriptive statistics and group comparisons for the 262 students who participated in the drawing activities.

The descriptive statistics and correlations between variables are shown, with the inputs divided into the treatment and control groups. Table 1 presents the correlations between the study variables in both the treatment and control groups. Most of the correlation coefficients between the variables are significant at the .01 level of significance, except the correlations between Fluency and Elaboration (not significant at the .05 alpha level) and between Originality and Resistance to Closure (not significant at the .05 alpha level) in both groups. Notably, the highest correlation coefficients are those between Elaboration and Resistance to Closure in both the treatment group ($r = .72, p < .001$) and the control group ($r = .66, p < .001$). The effect sizes, Cohen’s $d$, range from 0.28 to 0.37.

<table>
<thead>
<tr>
<th>Table 1. Descriptive statistics and correlations between the creativity subscales</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M</strong></td>
</tr>
<tr>
<td>1. Fluency</td>
</tr>
<tr>
<td>2. Originality</td>
</tr>
<tr>
<td>3. Abstractness of Titles</td>
</tr>
<tr>
<td>4. Elaboration</td>
</tr>
<tr>
<td>5. Resistance to Closure</td>
</tr>
</tbody>
</table>

*Note.* The correlations for the treatment group ($n = 137$) are below the diagonal; the correlations for the control group ($n = 125$) are above the diagonal. Titles = Abstractness of Titles; Closure = Resistance to Premature Closure. **$p < .01$; ***$p < .001$.**

**Differences in the creativity between the groups**

To examine the effectiveness of drawing-based instruction, we attempted to compare the average creativity scores between the treatment and control groups. The results are presented in Table 2. The data show that, compared with the control group, the treatment group (VTTCI) has a significantly higher creativity index mean score (97.17 vs. 92.90), a wider standard deviation (13.93 vs. 14.17) and a smaller range value. The results of the MANOVA indicate a single main effect of the instructional model, rejecting the null hypothesis of multivariate equality of the means over all groups. The MANOVA shows a significant overall treatment effect according to Hotelling’s $T$ criterion, $F(5, 256) = 5.23, p < .01$. There is a significant difference in the creativity scores between the treatment group, which adopted the VTTCI model, and the control group, which did not adopt the model.

An examination of the subsequent univariate tests for the dependent variables indicates that the instructional model has a significant effect on Originality, Abstractness of Titles, and Elaboration and does not have a significant effect on Fluency or Resistance to Premature Closure. The treatment and control groups ($M = 119.60, SD = 24.09$ and $M = 115.09, SD = 25.28$, respectively, $t = 1.48, p > .05$) do not show a significant difference in their Fluency subscale scores. The ceiling effect seems to suggest that most students obtained high Fluency scores on the TTCT. Additionally, the two groups’ levels of drawing skills may be similar because the VTTCI model did not require strong drawing skills and did not allocate practice time in the sessions other than the first introduction class. The Originality subscale scores of the treatment and control groups ($M = 106.17, SD = 17.66$ and $M = 101.56, SD = 19.53$, respectively; $t = 2.01, p < .05$) differ significantly. The effect size is .24 (according to Glass’s $\Delta$), indicating that the model’s effectiveness in stimulating students’ originality of thinking ranges between low and moderate.
levels. The treatment and control groups \((M = 74.56, SD = 25.49 \text{ and } M = 67.54, SD = 27.60, \text{ respectively, } t = 2.14, p < .05)\) show significant differences in their Abstractness of Title subscale scores. The effect size of .25 is moderate (Glass’s \(\Delta = .25\)); therefore, the model’s effectiveness in stimulating students’ abstractness of titles ranges between low and moderate levels. The Elaboration subscale scores of the treatment and control groups \((M = 97.68, SD = 21.53 \text{ and } M = 92.08, SD = 18.55, \text{ respectively, } t = 2.25, p < .05)\) also differ significantly. The effect size is \(\Delta = .30\), indicating that the model’s effectiveness in stimulating students’ elaboration of creative thinking ranges between low and moderate levels. There was no significant difference between the treatment and control groups’ Resistance to (Premature) Closure subscale scores \((M = 88.09, SD = 21.50 \text{ and } M = 88.18, SD = 23.18, \text{ respectively, } t = -.04, p > .05)\).

**Table 2. Means and standard deviations of the TTCT-Figural subscales**

<table>
<thead>
<tr>
<th></th>
<th>Treatment ((n = 137))</th>
<th>Control ((n = 125))</th>
<th>(t)</th>
<th>(p)</th>
<th>(F) ((df = 5, 256))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>M = 119.60, SD = 24.09</td>
<td>M = 115.09, SD = 25.28</td>
<td>1.48</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Originality</td>
<td>M = 106.17, SD = 17.66</td>
<td>M = 101.56, SD = 19.53</td>
<td>2.01</td>
<td>&lt; .05</td>
<td></td>
</tr>
<tr>
<td>Abstractness of titles</td>
<td>M = 74.56, SD = 25.49</td>
<td>M = 67.54, SD = 27.60</td>
<td>2.14</td>
<td>&lt; .05</td>
<td>3.228**</td>
</tr>
<tr>
<td>Elaboration</td>
<td>M = 97.68, SD = 21.53</td>
<td>M = 92.08, SD = 18.55</td>
<td>2.25</td>
<td>&lt; .05</td>
<td></td>
</tr>
<tr>
<td>Resistance to closure</td>
<td>M = 88.09, SD = 21.50</td>
<td>M = 88.18, SD = 23.18</td>
<td>&lt; .04</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Creativity score</td>
<td>M = 97.17, SD = 13.93</td>
<td>M = 92.90, SD = 14.17</td>
<td>2.45</td>
<td>&lt; .05</td>
<td></td>
</tr>
</tbody>
</table>

*Note. \(p < .01; N = 262. \) Variable ranges: Fluency = 40–149; Originality = 40–154; Elaboration = 40–160; Abstractness of Titles = 40–160; and Resistance to Premature Closure = 40–160; ns = not significant.*

**Discussion**

The current study sought to examine the effects and benefits of the VTTCI model on students’ creativity after one semester of use. The advances in digital technology and a rapidly evolving technological ecology have the potential to dramatically change teaching and learning. This study developed an instructional model, VTTCI, for generating drawings to enhance students’ creative experience in high-tech interactive classrooms.

The treatment group revealed a higher creativity score than the control group. Notably, the scores on three subscales, Originality, Elaboration, and Abstractness of Title, were significantly higher in the treatment group than in the control group. There are multiple explanations for these higher scores. First, students who engaged in active drawing imagined the context that surrounded their texts; these students also demonstrated generally deeper information processing because they had to express their understanding and reflections in their final products (i.e., drawings, titles, and labels) and during the sharing process. Second, students in the treatment group scored significantly higher on Originality, Abstractness of Title, and Elaboration. These results reflect the particular instructional model for learner-generated drawings in which teachers do not limit students’ drawing techniques or control their renderings, which produces a higher concentration of creativity. Students also freely drew to communicate their ideas with others. Likewise, after reviewing their peers’ drawings, students learned more effective and unique ways to express meaning and understanding. These creative expressions might have further influenced them and helped them elaborate on their learning. In addition, Abstractness of Title encouraged students to be clear and concise to convey their meaning to their classmates using labels or words. At the beginning of these activities, students did not produce effective outcomes to represent the problems that teachers gave them, such as appropriate explanatory labels for their drawings (e.g., the labels were overly long or absent). Additionally, students required practice to formulate keywords to represent and label the target objects or stories so that others could understand their drawings. Thus, these activities enhanced students’ scores on Abstractness of Titles, regardless of whether the title was a concrete one-word label for the picture or was so abstract that one had to view both the picture and title to gain a complete understanding. Students provided more clearly focused titles and labels for their fourth and fifth drawings. Thus, we suggest that students learned how to more clearly label their work using clear and concise language.

This finding implies that students increased their level of comprehension in the process of transforming verbal information into a mental model. This process can be augmented by students’ presentations and communication with
their peers and teachers (Leutner, Leopold, & Sumfleth, 2009). Teachers need to guide their students to develop their own visualizing strategies. Students can be trained to visualize textual content and to develop additional cognitive learning strategies, such as text highlighting or concept mapping. However, when students’ drawings depend too much on texts, students repeatedly check the text and the picture to ensure a better understanding of the text. In the cognitive process, students unnecessarily expend their cognitive resources, which leads to an increased cognitive load. Because the results of an increased cognitive load can hinder student comprehension and learning, teachers need to guide their students in this process (Leutner et al., 2009).

A highly interactive learning environment supports effective and efficient learning, as students can find clues to enhance future outcomes and comprehension using drawing-related approaches and techniques. In contrast to traditional classrooms, high-tech interactive classrooms are built so that students can formulate their ideas and share them with others in the classroom. Such classrooms can serve as successful learning environments that support students’ creative experiences and unlock their imaginative potential.

This study posits that highly interactive classroom environments that use innovative technologies can provide various opportunities for students. An environment that fosters a deeper bond among students and between students and their teachers supports more engagement in classroom activities. In traditional middle school settings, each class period lasts 45 minutes, and teachers must plan their lessons to fit within that time constraint. Thus, the traditional classroom approach, which inherently offers less time, does not provide ample opportunities for students to interact with their teachers or peers during classroom activities.

Additionally, students can benefit from sharing their work, interacting on assignments, exchanging thoughts with their peers, and collaborating in the classroom. Technologies that support efficient classroom activities can be used to establish highly interactive classroom environments. During many classroom activities, students do not frequently have opportunities to share and evolve their learning. Classroom interaction using technologies, for example, mirroring tables, transferring files in real time, and projecting instructional objects, can offer students innovative and creative experiences.

**Implications and limitations**

In this paper, we have developed an instructional model, called VTTCI, for generating drawings to enhance students’ creative experiences in high-tech interactive classrooms. Our results contribute to the existing research on learning activities that foster learners’ creativity through drawings; our study also suggests an innovative method of enhancing classroom interactivity using high-tech systems. This study has produced important findings that teachers and administrators in school districts can consider in building more interactive classrooms with current innovative technologies. We believe that the VTTCI model can be applied in various courses and adapted for incorporation with tablet-based interactive classrooms and alternative technologies.

This study developed an instructional model in which students seemed to be creative collaborators in a classroom setting with technology integration. The instructional model suggested in this study should be further tested using samples from various instructional contexts, such as different class sizes, cultures, socioeconomic backgrounds, education levels, and technological environments. Additionally, future studies should utilize randomized conditions to extract precise data. Students’ performances may be influenced by academic motivation, self-regulation, artistic abilities, and study skills, particularly note-taking skills. Further studies can examine other determinants that enhance student performance.

Although the study’s findings contribute to our knowledge of the effects of technology integration on creativity in schools, there are several limitations that must be considered when planning additional research. First, the findings should be interpreted to explain learners’ characteristics after they complete learner-generated drawings in the process of comprehending instructional texts. The question of whether the results can be generalized to students’ learning achievement and overall learning, in terms of the use of an instructional model that includes drawing activities, requires additional research. Additionally, the sample in this study was taken from a middle school in
South Korea only six months after the school first opened. The climate of a new school possibly hindered the collection of more stable results.

Acknowledgements

The authors are grateful to the editor and the anonymous reviewers for their valuable comments and suggestions. The authors would thank the Korea Educational and Research Information Service for the permission to use the creativity data (TTCT-Figural) collected at the school for this study.

References


Effectiveness of a Learner-Directed Model for e-Learning

Stella Lee¹, Trevor Barker² and Vivekanandan Suresh Kumar¹*

¹Faculty of Science & Technology, Athabasca University, Canada // ²School of Computer Science, Hertfordshire University, United Kingdom // stellaylee@gmail.com // T.1.Barker@herts.ac.uk // vivek@athabascau.ca
*Corresponding author

(Submitted January 14, 2015; Revised April 9, 2015; Accepted November 3, 2015)

ABSTRACT
It is a hard task to strike a balance between extents of control a learner exercises and the amount of guidance, active or passive, afforded by the learning environment to guide, support, and motivate the learner. Adaptive systems strive to find the right balance in a spectrum that spans between self-control and system-guidance. They also concern a smoother shifting of the balance point during learning episodes in light of competing requirements from learning goals, learner capacity, instructional affordances, and educational theories, among others. This research investigates one of the extremes of this spectrum, where learners actively assume control and take responsibility for their own learning, while catering to individual preferences with little or no guidance from the e-learning environment. In this study, one unit material from an online Introduction to Java Programming course has been redesigned based on the proposed Learner-Directed Model for the experimental design study. The model is developed based on the exploration of two educational theories — Experiential-Learning Theory (ELT) and Self-Regulated Learning (SRL) Theory. The study involved a total of 35 participants (N = 35) divided randomly into one Experimental Group and one Control Group. They were assigned to either a Learner-Directed Model (Experimental Group) or a linear model (Control Group). Pre/post tests, survey, follow-up interview as well as log file analysis were instruments used for assessing students’ domain knowledge, meta-knowledge, and their attitudes for their overall learning experience. The results of the study have revealed that there is a statistically significant higher level of overall learning experience and better learning attitudes compared to Control Group students who studied with e-learning components that are linear in nature and are without explicit associations with educational theories.

Keywords
Learner-directed model, Self-regulated learning, Learning preferences, Instructional design, Learning design

Introduction
In the space of the past decade, e-learning has gone from being a supplementary form of learning to an increasingly staple part of higher education and industry training. Nowadays, learners are free to enroll in a Massive Open Online Course (MOOC) with diverse topics ranging from The Science of Gastronomy to Jazz Improvisation. The growth of portable digital devices, from tablets, smart phones, touch screen mini-laptops to e-readers further solidify the ubiquitous learn-as-you-see-fit mentality. Yet, the demand for personalized learning is not adequately supported by current technology or practices (Johnson et al., 2013). As students’ tastes for online pedagogy become more sophisticated, they will increasingly demand learning that provides learner-directed choices and control. As the e-learning service provider sphere becomes more competitive, competition will drive the change on a more customized education model that cater to each student’s unique learning needs. One-size-fits-all e-learning has declined in popularity — it was a solid base to start just as Web 1.0 had its place in popularizing information dissemination electronically in a networked environment. Now, that method is less effective, and in some ways, it is irrelevant to the changing demographics of fragmented global learners. Just as the travelers can opt to drive to their destination instead of journey by train; as self-directed, autonomous adult learners, they need the options to negotiate modes that best support each individual’s learning needs, be it the amount of content, the types of learning activities, or the form of devices used to access the material. Thus, the level of control a learner has depends on how much he/she is willing to trade off between having to make independent decisions at a granular level in one end of the spectrum (thus having a greater control) and allowing the system to take care of the bulk of the decision-making process (thus having less control) on the other end. It is a hypothesis of this research that having an optimal level of learner control is essential in creating a positive learner-directed learning experience. In particular, the aim of this research is to support learners to assume active control and take responsibility for their own learning. As an extension of that, it aims to augment the learning of computer programming in a way that would motivate and engage learners, so as to counter the high attrition rate in this domain, especially in an asynchronous online environment.
To create a positive learner-directed learning experience and provide optimal levels of learner control, an alternative approach to learning computer programming online called a Learner-Directed Model was developed for this study. This model is grounded in the integration of two established education theories that are learner-centric. The two theories are Experiential Learning Theory (ELT) and a modified version of Self-Regulated Learning (SRL) Theory, which this research called Self-Directed Regulated Learning (SDRL). ELT was used as a guideline for the domain-knowledge content with the four learning modes. SDRL was used as a guideline for the meta-knowledge (i.e., study skills) content that was integrated and contextualized into the model. This model caters to learner control with the consideration of freedoms of pace, content, and media, and it includes domain knowledge learning activities as well as meta-cognitive skills’ interface.

The hypotheses of this research are as follows:

- Online computer systems that are designed based on learning style theories and educational theories are beneficial to learners
- Online computer systems that employed unobtrusive learner support for meta-cognitive activities such as study skills are beneficial to learners
- Online computer systems developed in the above ways could improve learning for students in terms of learner control, attitudes, and learner experience (LX)

Based on the above stated hypotheses, research questions were formulated. They are as follows:

- Which learning style theory and education theory would be useful to develop an e-learning system?
- How could an e-learning system be developed based on the learning style and education theories?
- What is the effectiveness of an e-learning system based on a model of learning style and education theories?
  - What would the effect of such a system be on the performance of learners?
  - What would be the attitudes and learner experience (LX) of learners who would be using such a system?

### Related work

**Experiential Learning Theory (ELT)**

Developed by David Kolb, Experiential Learning Theory (ELT) suggests that learning requires polar opposite abilities; thus it creates conflict and learners are forced to choose which set of abilities to use in a certain learning situation in order to resolve the conflict. For example, it is not possible to drive a car and analyze a driver’s manual at the same time, we need to choose which approach to take; as a result, over time, we develop a preferred way of choosing (Kolb, Boyatzis, & Mainemelis, 2000). While ELT recognizes that learners might have their preferences in how they learn, it emphasizes the importance of a well-rounded learning experience. Kolb (1984) stated that effective learners need four types of abilities to learn: concrete experience (CE); reflective observation (RO); abstract conceptualizations (AC); and active experimentations (AE). Figure 1 below illustrates the four types of learning abilities as well as how a learner can progress through the experiential learning cycle: experience is translated through reflection into concepts, which are in turn used as guides for active experimentation and the choice of new experience. Kolb (1984) stated that a learner could begin the learning cycle at any one of the four modes, but that learning should be carried on as a continuous spiral. As a result, knowledge is constructed through the creative tension among the four modes and learners will be exposed to all aspects of learning: experiencing, reflecting, thinking, and acting.

Furthermore, there are four dominant learning styles that are associated with these modes. The four learning styles are: Converging (AC and AE focus); Diverging (CE and RO focus); Assimilating (AC and RO focus); and Accommodating (CE and AE focus). Converging is interested in practical applications of concepts and theories. A learner who has this preference is good at problem solving, decision making, and practical application of ideas. Diverging is interested in observation and collection of a wide range of information. A learner leaning toward this preference is more concerned with ideas and abstract concepts than with people. Accommodating is interested in hands-on experiences. A learner with this preference likes doing things, carrying out plans, and the trial-and-error method.
As this research is grounded in individualized instruction, it furthered Kolb’s take on distinctive learning preferences in order to provide options on learning activities accordingly. The learning orientations Kolb proposed provide a framework for learning design, and it promotes holistic learning in the context of a self-paced online learning environment. The four modes in the learning cycle have been used to design task-level content in an online Java programming course in the research design. However, it is important to note that it is not the researchers’ intention to prescribe and match a learner with a particular learning preference, as every learner uses a mix of learning styles, and not just one.

**Self-Regulated Learning (SRL) Theory**

As this research is interested in the person-environment interaction in academic learning online, another learning theory that helps facilitate individuals in monitoring their learning process and to act accordingly to gain some control over them in academic setting is Self-Regulated Learning Theory (Dinsmore, Alexander, & Loughlin, 2008). Self-regulated learning (SRL) refers to the setting of one’s goals in relation to learning and ensuring that the goals set are met (Boekaerts & Corno, 2005; Butler & Winne, 1995; Perry, Philips, & Hutchinson, 2006; Winne & Perry, 2000; Zimmerman, 1990).

SRL theory states that learners not only need to regulate their performance, but also how they learn. Literature shows that self-regulation can positively affect learners’ achievement (Azevedo, 2005). SRL is “an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behaviour” (Pintrich, 2000). It is a deliberate, judgmental, and adaptive process (Zimmerman, 1990) where feedback is an “inherent catalyst” (Butler & Winne, 1995). In addition, Hadwin, Wozney, and Ponton (2005) revealed that “ordinary” collaboration (as found in traditional online forums and chats) is insufficient to transform learners’ everyday “reflection” into productive “regulation.”

What sets self-regulated learners apart is their awareness of when they know a skill or fact and when they do not, at a meta-knowledge level — i.e., based on what they know or do not know, they plan, set goals, organize, self-monitor, and self-evaluate throughout their studies (Zimmerman, 1990). In addition, they are able to evaluate and trade-off between small-scale tactics and overall strategies, and are able to predict how each can support their learning progress toward the goals they pre-selected (Winne, 1995).

Despite the benefits of self-regulation, there are many reasons why learners choose not to self-regulate their own learning. For example, some learners might feel that planning, evaluating, monitoring, and evaluating their learning process takes too much time and they are not willing to invest the resources in that particular context (Boekaerts, 1999). Boekaerts (1999) noted that SRL has a bidirectional relationship with learning environments. She asserted that learning environments could be used to promote self-regulatory skills and act as facilitators for learning new self-regulation strategies. The authors would like to extend this line of argument to online learning environments where a system can be designed to introduce, facilitate, scaffold, and provide options for self-regulation process and
strategies in an unobtrusive way. Indeed, Winne (1995) suggested that self-regulation is common, and that learners should develop an effective means to self-direct their learning, so it is recommended “the classical topic of instructional design could be revitalized as a scientific approach to accommodating and changing learners’ knowledge” (Winne, 1995). It is with this foundation that this research model design is based on: The learner-directed flexible system wherein students are presented options and self-regulated strategies to select and engage in.

Learner-Directed Model

The concept behind a flexible Learner-Directed Model for this research is grounded on learner control and learner experience. As a starting point, this research will employ the ELT stages of learning cycle in creating a learning design framework for presenting the domain knowledge, namely, introductory computer programming in Java. As a constructivist theory of learning, ELT fulfils many of the criteria for building a learner-directed learning model. It is a learner-centred and multilinear model as it caters to different learning styles; it is consistent with how people learn, grow, and develop (Kolb et al., 2000). More importantly, this is an inclusive model of adult learning that intends to explain the complexities of and differences between adult learners within a single framework (Oxendine, Robinson, & Willson, 2004).

Using SRL as a complementary theory for the design of a Learner-Directed Model fits well with the requirements for the current model design. It supports learner control, promotes autonomous learning, and enhances the overall learner experience. The type of self-regulated learning approach this study will use for the design of meta-cognitive activity support is called Self-Directed Regulated Learning (SDRL), a term coined by the authors.

Self-Directed Regulated Learning (SDRL)

At one end of the spectrum, self-reflection could be as simple as a thinking process made self-aware, the intangible “ah ha” moment of a conceptual breakthrough; at the other end, self-regulation could be more tangible as a system or a tutor can observe what students do after they self-reflect. For example, in debugging a piece of programming code, a student can self-reflect on errors identified by the compiler at the end of each compile of the program being developed. The system can track/record the number of errors and warnings faced by the student at the end of each compile. The system can also classify the types of errors encountered by a single student over multiple sessions of program development. Looking at this list of errors is an act of self-reflection. However, self-regulation takes it one step further. Students may try to identify most common errors and warnings they faced, take notes on how they resolved these common errors and warnings, and refer to these notes when they encounter these common errors and warnings when writing another program. This “proactive self-reflection” is what the authors identified as self-directed regulated learning (SDRL) because at this stage, this research does not plan on tracking the end results of using the study skill tools, it is rather the provision of these “how to” study skill guides and tools embedded as options for students to assist them in becoming better self-regulated, self-reflected learners that this research is interested in.

The difference between SDRL and SRL is that SDRL aims at “regulation” in a more implicit manner than SRL. That is, in SRL, regulation is more explicit, whereby one can directly and independently measure the degree of regulation being imparted by the system as well as measure the degree of regulation being absorbed or applied by a student. In SDRL, regulation is being implicit in the design and interactions, measurements related to regulation are inherently intertwined with other components of the system, and hence it is hard to measure the degree of regulation in a direct and independent manner.

Model design

One unit (Unit 3 – Control Structure) of an introductory to programming course material was being redesigned to fit with the Learner-Directed Model. The content was based on COMP268 - Introduction to Java Programming course currently offered by the School of Computing and Information Systems (SCIS) at Athabasca University in Canada. This course was chosen as the domain due to the fact that this course has one of the highest attrition rates within the school. Common complaints among students are that the course does not engage them and there is a lack of
motivation as a result. In addition, the complexity of learning computer programming further discourages students from completing their online study.

After the redesign, the module used in this research study was composed of the five following learning concepts:
- Overview of Intro to Programming
- If-Else Statement
- Loops
- Break, Continue, and Return Statement
- Switch, and Try-Catch-Finally Statement

The ELT learning cycle was mapped on top of a circular wheel interface with four quadrants representing the four learning stages. Each stage was linked to a learning mode based on Kolb’s experiential learning cycle: concrete experience (CE), reflective observation (RO), abstract conceptualization (AC), and active experimentation (AE). The four learning modes corresponding to Kolb’s four stages are: watching, discussing, conceptualizing, and trying out (Figure 2).

The four main types of activities, one representative of each aspect of the learning cycle are:
- Concrete Experience – Watching (viewing tutorial videos)
- Reflective Observation – Discussing (posting to discussion forums)
- Abstract Conceptualization – Conceptualizing (concept mapping and mental model building)
- Active Experimentation – Trying Out (writing code in Java)

As Kolb stated that learning can begin at any stage in the cycle, similarly, this learning design follows the idea that learners have different ways of learning, having an option to choose which way is best for them as to where the starting point and end point are, and when they are ready to leave the learning material. The five learning concepts all start with the same interface that presents the four modes of learning. A pre-test is given at the beginning of the course to assess their prior knowledge in computer programming, as well as to gauge their prior competency in study skills. Figure 3 indicates what learners assigned to the Experimental Group will see when they enter the course home page and select “Learning Concept 1.”

For learners who are interested in learning by observation, they can select the “watching” mode, which will provide a series of tutorial videos on how to write and edit a basic program in Java (Figure 4). For others who prefer to engage in the reflective process in the form of discussion forum posting, they can select the “discussing” mode. It will take the learners to the page with learning activities that are designed for social interaction, dialogues, and reflections.
This interface design encourages the learner to explore and access more than one learning activity, but they might or might not choose to go through them all; the decision is theirs to make. If at the end of any chosen learning activity, the learners feel that they have a good grasp of the material, they can opt to skip the rest of the cycle and go to the next topic. A post-test is available at the end of the module for learners to self-assess their knowledge both at the domain and meta-cognitive levels. In other words, learners can test how well they have learned a certain programming concept in Java as well as whether the associated key study skills are helpful to them in regulating their own learning and deploying them as learning strategies. The post-test for the domain-knowledge portion asked programming-related questions and is objective in nature; the meta-cognitive skill portion of the test is in the form of self-reporting and is subjective in nature. Altogether, it provides a high-level view of competency on both domain and meta-cognitive levels of learning.

As supplementary learning material, a Self-Directed Regulated Learning (SDRL) theory based interface titled “Key Study Skills” is available on the upper right-hand corner of the webpage to assist with learning about meta-cognitive knowledge. As mentioned in Chapter 1, learning computer programming doesn’t rely on a single skill; learners can benefit from mastering a set of meta-cognitive skills to optimize learning (Azevedo & Cromley, 2004; Winne, 1995). The way to provide study skills in this research is a “built-in” approach (Wingate, 2006). As opposed to the “bolt-on” approach (Bennett et al., 2000) where study skills support is treated as an external element independent of the subject matter, this approach integrates study skills in a contextually appropriate manner.
Topical study skills for studying introductory computer programming has been selected in these four areas: note taking, communications, conceptualizing, and problem solving skills. Twenty related study skills — five successive skills (to scaffold) for each of the four learning modes — have been developed. Each study skill is paired with a tool for learners to try out and practice these skills. For example, within the “Watching” activity, the key study to emphasize is “note taking.” Learners can choose to access this supplementary material at any given time during their study. The design consideration is that the key study skills appear as an unobtrusive interface that sits on the upper-right-hand-corner of the main content. Learners can choose to engage and learn how the “note taking” skill helps with the “watching” activity (i.e., how to take better notes while watching a tutorial video online and what are the appropriate tools to use for note taking), or they can just ignore it and carry on with watching the tutorial video on their own. Figure 5 shows that the Note Taking skill is being presented with Evernote, and note-taking software. The rest of the study skills are paired with the following: communications – Moodle discussion forum; conceptualizing – FreeMap (mind mapping tool); and problem solving – BlueJ (an integrated development environment tool).

**Experimental design and methodology**

**Method**

This study involved 35 undergraduate student volunteers from Athabasca University to study and evaluate one unit of adapted content material from *Introduction to Java Programming* (COMP 268). The volunteers were assigned to either an Experimental Group or a Control Group during this six-month period. Participants agreeing to the consent form were invited to log into Moodle, a learning management system (LMS) currently used at Athabasca University. Learner-Model site was assigned to an Experimental Group of 18, while static and linear content without study skills support was being presented to a Control Group of 17. Six participants were selected randomly from the two groups (three from each group) for the follow-up semi-structured interviews to collect qualitative feedback about the study. Table 1 summarizes the participants’ demographic information for both Experimental and Control Groups:

<table>
<thead>
<tr>
<th>Table 1. Demographic information for experimental and control groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental group</strong></td>
</tr>
<tr>
<td>Total number of participants</td>
</tr>
<tr>
<td>Age range</td>
</tr>
<tr>
<td>Gender</td>
</tr>
</tbody>
</table>

Students who were assigned to the Experimental Group were directed to the Learner-Directed Model site called “Control Structures (Experimental)” while the Control Group students would study the linear-model site called
“Control Structures (Control).” For both the Control and the Experimental Groups, participants could only access the material after taking the pre-test.

For the Experimental Group, when students clicked on the Learning Activities, they were taken to the next screen that presented them with a flexible learning interface in the shape of a circle with four quadrants. At this stage, students were presented with different options for their learning activities: watching, discussing, conceptualizing, and trying out. Alternatively, they could skip ahead to another learning concept or go directly to the post-test if they were confident of passing this module without going through any learning material at all.

As for the Control Group, the view on the module and the interface was rather different. The design of the Control Group module mirrored the typical design of a current online offering for computer programming courses. Unlike the Experimental Group, the Control Group did not have a flexible circular interface in quadrants. Instead, a linear presentation of content was made available for learners to access. Five learning concepts were provided, the same learning content as the Experimental Group. Once the participants were done with one learning concept, it will link to the next learning concept, and this continued on until all five learning concepts were covered. After the learners were finished with all five learning concepts, a post-test was required for them to assess their competency on this module. Even though learners were not required to go through the learning concepts in a sequential order, it wasn’t made explicit to them. The learning activities were the same as the ones in the Experimental Group, but for the Control Group, learners were not made aware of the different learning preferences, thus, there was no differentiation of the Watching, Discussing, Conceptualizing, and Trying Out activities.

It is important to point out that this research is limited to the examination of the effect on adult learners on domain-knowledge, meta-cognitive knowledge, and their overall learning experience (LX) based on data collected quantitatively online and self-reporting from the participants with qualitative methods such as interviews and survey. The research will be limited to the adaptation of one unit from an introduction to programming course (COMP268) typically available for first or second year students at Athabasca University in Alberta, Canada. The course is part of a regular offering at the School of Computing and Information Systems (SCIS), Faculty of Science and Technology, one of the largest faculties at Athabasca University.

Data collection

Five data collection instruments have been used for this research study to collect quantitative and qualitative data:

- Pre-test/Post-test — pre-test and post-test have been conducted to measure whether the two groups were initially the same in terms of their prior subject — level knowledge and meta-level knowledge.
- Log file data — log file data has been collected to track the time spent on domain area activities and the study skills activities within the Moodle website for both groups.
- Survey — at the end of the module, a survey was provided to measure learners’ attitudes toward study skills and tools provided in the module, their perceived ease of use and satisfaction, system usability, learners’ experience, and perceived controllability. Questions on how the learners used the study skills tools as well as open-ended text questions were also included in the survey.
- Follow-up interviews — semi-structured interviews with six (three from Control Group and three from Experimental Group) randomly selected participants were conducted. The purpose of the interviews was to verify and extend information obtained from the survey, particularly in discovering more details about the learners’ attitudes and impression about the experiment. The interviews also gave learners a chance to elaborate on the survey results.

Results and discussions

Pre-post test results

Quantitative and qualitative data were collected and the findings are presented and analyzed in this section. For quantitative data, statistical tests were performed and displayed in Tables 2 and 3. They contain descriptive statistics for each variable corresponding to Experimental Group (N = 18) and Control Group (N = 17). The tables present the arithmetic means and standard deviation.
The null hypothesis employed in this research was that any difference in mean performances and attitudes between the experimental and control conditions was due to chance alone. In order to test this, an independent-sample t-test was conducted to compare domain knowledge pre-tests and post-tests, study skills usage pre-tests and post-tests and survey between the Experimental Group and the Control Group. The results of this analysis are shown in Table 4 below.

Table 2. Descriptive statistics for the experimental group

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Test (domain knowledge)</td>
<td>18</td>
<td>15.7</td>
<td>2.7</td>
</tr>
<tr>
<td>Post-Test (domain knowledge)</td>
<td>18</td>
<td>15.8</td>
<td>2.7</td>
</tr>
<tr>
<td>Pre-Test (study skills)</td>
<td>18</td>
<td>33.9</td>
<td>12.9</td>
</tr>
<tr>
<td>Post-Test (study skills)</td>
<td>18</td>
<td>35.4</td>
<td>13.6</td>
</tr>
<tr>
<td>Time spent on study skills (minutes)</td>
<td>18</td>
<td>9.2</td>
<td>22.8</td>
</tr>
<tr>
<td>Time spent on domain (minutes)</td>
<td>18</td>
<td>428.9</td>
<td>907.1</td>
</tr>
<tr>
<td>Total time (minutes)</td>
<td>18</td>
<td>438.1</td>
<td>929.9</td>
</tr>
</tbody>
</table>

Table 3. Descriptive statistics for the control group

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Test (domain knowledge)</td>
<td>17</td>
<td>14</td>
<td>3.7</td>
</tr>
<tr>
<td>Post-Test (domain knowledge)</td>
<td>17</td>
<td>14.5</td>
<td>3.1</td>
</tr>
<tr>
<td>Pre-Test (study skills)</td>
<td>17</td>
<td>32.8</td>
<td>13.1</td>
</tr>
<tr>
<td>Post-Test (study skills)</td>
<td>17</td>
<td>36.8</td>
<td>3.4</td>
</tr>
<tr>
<td>Time spent on domain (minutes)</td>
<td>17</td>
<td>160</td>
<td>163.6</td>
</tr>
</tbody>
</table>

Table 4. Independent samples t-test

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>t</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test (domain knowledge)</td>
<td>1.57</td>
<td>1.54</td>
<td>33</td>
<td>0.13^2</td>
</tr>
<tr>
<td>Post-test (domain knowledge)</td>
<td>0.17</td>
<td>1.30</td>
<td>33</td>
<td>0.10^1</td>
</tr>
<tr>
<td>Pre-test (study skills)</td>
<td>0.07</td>
<td>0.27</td>
<td>33</td>
<td>0.79^2</td>
</tr>
<tr>
<td>Post-test (study skills)</td>
<td>5.59</td>
<td>-0.39</td>
<td>33</td>
<td>0.35^1</td>
</tr>
<tr>
<td>Total time spent</td>
<td>4.1</td>
<td>1.01</td>
<td>33</td>
<td>0.15^1</td>
</tr>
<tr>
<td>Survey</td>
<td>0.40</td>
<td>2.32</td>
<td>33</td>
<td>0.02^1</td>
</tr>
</tbody>
</table>

Note. ^1 = 1 tailed; ^2 = 2 tailed.

The finding shows that there was no significant difference in performance between the Experimental Group and the Control Group over the course of this experiment (p > 0.05). The results of the analysis also show that there was no difference between the performance of the Experimental Group and the Control Group on pre-test and post-test for both domain knowledge and study skills (p > 0.05). One tailed statistics were employed in post-test comparisons as it was predicted that the Experimental Group would perform better than the Control Group. The results of the comparison of the survey results between the Experimental and Control Group show a significant difference in their responses (p < 0.05). The null hypothesis is therefore rejected and any differences in the means of attitude measurements are due to the effect of the experimental condition. There were no significant differences between the Experimental Group and the Control Group in terms of total time spent studying the domain material (p > 0.05), though the Experimental Group spent more time in total as they spent time on study skills as well as for domain knowledge. The lack of difference between the two groups on time spent on domain material is interpreted that even given a flexible approach to learning where learners can skip learning material, they choose to spend a similar amount of time in studying the material.

Survey results

The end-of-the-module survey has a total of 28 questions for the Control Group, with additional three questions for the Experimental Group for a total of 31 questions. The additional questions concern the use of the supplementary study skills and use of tools that were only available to the Experimental Group. To ensure high construct validity, eight survey items had been adapted from Davis (1989) and Unger and Chandler (2009) for measuring perceived ease of use and user satisfaction. For system usability, six items were based on the work of Koohang and Du Plessis
(2004) while three items derived from Smulders (2003) for learning experience. As for perceived controllability, four items were reworded from Liu (2003); finally, two open-ended questions “what do you think are the best features of the module, and why?” and “what features of this module did you think should be improved, and why?” were posted as an attempt to identify ways to improve the product and were based on the idea of Tullis and Albert (2008).

Table 5 summarizes the survey results of the survey with the following topics: Ease of Use and Use Satisfaction; System Usability; Learning Experience; and Perceived Controllability.

<table>
<thead>
<tr>
<th></th>
<th>Group</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Std. error mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of Use &amp; User Satisfaction</td>
<td>Experimental</td>
<td>26.89</td>
<td>3.45</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>26.06</td>
<td>4.23</td>
<td>1.03</td>
</tr>
<tr>
<td>System Usability</td>
<td>Experimental</td>
<td>24.33</td>
<td>4.12</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>21.41</td>
<td>3.14</td>
<td>0.76</td>
</tr>
<tr>
<td>Learning Experience</td>
<td>Experimental</td>
<td>14.11</td>
<td>2.95</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>12.89</td>
<td>1.96</td>
<td>0.48</td>
</tr>
<tr>
<td>Perceived Controllability</td>
<td>Experimental</td>
<td>17.61</td>
<td>2.87</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>15.41</td>
<td>2.92</td>
<td>0.71</td>
</tr>
</tbody>
</table>

It is interesting to note that in all cases the mean of the Experimental Group was higher than the mean of the Control Group. In order to test the significance of these possible differences in the means, an independent t-test was performed. One-tailed statistics were used as it was hypothesised that the Experimental Group would have a higher rating than the Control Group. Table 6 below shows the results of the test performed on the data related to the four usability components.

<table>
<thead>
<tr>
<th></th>
<th>Group</th>
<th>T</th>
<th>df</th>
<th>Sig. (one-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of Use &amp; User Satisfaction</td>
<td>0.41</td>
<td>0.64</td>
<td>33</td>
<td>0.13</td>
</tr>
<tr>
<td>System Usability</td>
<td>1.38</td>
<td>2.35</td>
<td>33</td>
<td>0.005</td>
</tr>
<tr>
<td>Learning Experience</td>
<td>1.19</td>
<td>1.44</td>
<td>33</td>
<td>0.04</td>
</tr>
<tr>
<td>Perceived Controllability</td>
<td>0.02</td>
<td>2.25</td>
<td>33</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Significant differences were found in the System Usability measure \((p < 0.01)\), Learning Experience \((p < 0.05)\) as well as for Perceived Controllability \((p = 0.01)\). These results indicate that in terms of these three usability measures, the increased ratings observed for the Experimental Group as compared to the Control Group were not due to chance alone, but to the influence of the independent variable, experimental condition. The result of the t-test comparing the means of the Experimental and Control Groups for Ease of Use and User Satisfaction was not significant, though it is fair to say that the difference, although small, was in the expected direction. Only 13% of the time would this result be expected to be due to chance alone. It is likely that the high quality of learning materials used in both Control and Experimental Groups may have produced a ceiling effect, both groups having similar levels of satisfaction.

All in all, the experiment was able to show that there was significance in total time spent studying. The Experimental Group spent longer time on the material than the Control Group. However, there was no significant difference in performance between the Experimental Group and the Control Group over the course of this experiment \((p > 0.05)\). The results of the analysis also show that there was no difference between the performance of the Experimental Group and the Control Group on pre-test and post-test for both domain knowledge and study skills \((p > 0.05)\). For the survey, the comparison results between the Experimental and Control Group show a significant difference in their responses to attitude questions \((p < 0.05)\). In all cases for the survey results, the mean scores for the Experimental Group were higher than the Control Group. Significant differences were found for the System Usability measure \((p < 0.01)\), Learning Experience \((p < 0.05)\) as well as for Perceived Controllability \((p = 0.01)\).
Interview results

The aim of the interviews for the purpose of this research is to support the results of the survey and give participants a chance to elaborate their survey answers. It is an opportunity to learn more about the participants’ attitudes and impressions of the modules, and to gain further insights on the differences between the Control Group’s and the Experimental Group’s learner experience. The interviewing questions were validated by an external user experience design professional with over 10 years of experience in user interviews in Edmonton, Canada. The results of the interviews have supported the survey data in which it indicated that there was a difference in perceived controllability and system usability between the Control Group and the Experimental Group. The interviews provided more details on the differences and also some of the similarities between the learner experiences of using the two modules – one for the Control Group and one for the Experimental Group.

The overall findings of the interviews can be summarised as the following:

- The Experimental Group had to deal with an initial learning curve of adapting to the new interface; however, once they were familiar with it, they thought it helped with the flow of the learning material and provided flexibility.
- The interface design for the Experimental Group seemed to promote thinking in learning styles and preferences. It also appeared to stimulate thinking in terms of how computer programming for online learning can be presented/designed differently. More comments about how to improve the interface design stemmed from the Experimental Group rather than from the Control Group.
- It was evident that the Experimental Group used more constructive and positive descriptive terms when describing their learning experience. For examples, terms such as “feeling in control,” “enjoy the flexibility,” “freedom to pick or skip,” “give me an opportunity to use different ways of learning,” “it has a flow,” etc. all pointed to a stronger sense of learner satisfaction in their learning experience.
- The Control Group provided more comments on how to improve or present the content differently while the Experimental Group addressed both the content and the interaction design issues. It suggested that by presenting a different/unusual interface for e-learning, it prompted students to think beyond what the possibilities are for improving their online learning experience.

In conclusion, the findings of the interviews conveyed that the Experimental Group had an overall more positive and constructive learning experience as opposed to the Control Group. They paid attention to both the content and the interaction design, and in general felt in control and satisfied with their learning. This finding is in support of the results from the end-of-module survey.

Conclusion

This study developed an alternative approach to learning computer programming online called a Learner-Directed Model in order to provide a positive learning experience and optimal levels of learner control. The design is grounded in the integration of two established learner-centric education theories. The two theories are Experiential Learning Theory (ELT) and a modified version of Self-Regulated Learning (SRL) Theory, which this paper’s authors called Self-Directed Regulated Learning (SDRL). ELT was used as a guideline for the domain-knowledge content with the four learning modes. SDRL was used as a guideline for the meta-knowledge (i.e., study skills) content that was integrated and contextualized into the four learning modes. This model caters to learner control with the consideration of freedoms of pace, content, and media, and it includes domain knowledge learning activities as well as meta-cognitive skills’ interface. Furthermore, this model takes on a constructive approach to learning. Constructive learning is based on students’ active participation in problem solving and critical thinking regarding a learning activity, which they find relevant and engaging (Masuyama, 2005). This experimental study with 35 voluntary participants (N = 35) indicated that there is statistically significant difference between the survey results for the Experimental Group and the Control Group. The Experimental Group reported a higher level of overall learning experience and better attitudes in general. However, there was no statistically significant difference existing between the two groups on the domain and meta-level knowledge improvement.

These experimental results have important implications for designing the next generation of e-learning systems as the Learner-Directed Model enables us to present an alternative to the existing instructional design framework. It goes beyond presenting and delivering online material to a theory-centric, user-centred, and Learner-Directed Model for
asynchronous learning. This study highlighted a shift in underpinning learning design philosophy, especially one for learning about computer programming online. It also served as a departure from the view on student modeling using adaptive systems, to distinguish the difference between adaptive system and adaptable system with the aim to optimize learner direction and learner control.

Acknowledgements

The authors acknowledge the support of NSERC, iCORE, Xerox, and the research-related gift funding by Mr. A. Markin.

References


Predictive Effects of the Quality of Online Peer-Feedback Provided and Received on Primary School Students’ Quality of Question-Generation

Fu-Yun Yu¹ and Chun-Ping Wu²*

¹Institute of Education, National Cheng Kung University, Taiwan // ²Department of Educational Technology, TamKang University, Taiwan // fuyun.ncku@gmail.com // cpwu303@gmail.com

*Corresponding author

(Submitted February 11, 2015; Revised September 6, 2015; Accepted October 20, 2015)

ABSTRACT

The research objectives of this study were to examine the individual and combined predictive effects of the quality of online peer-feedback provided and received on primary school students’ quality of question-generation. A correlational study was adopted, and performance data from 213 fifth-grade students engaged in online question-generation and peer assessment for six weeks were analyzed using hierarchical multiple regression, with the dependent variable of scores on question-generation and independent variables of scores on peer-feedback provided and received. The results from the two-step hierarchical regression analysis indicated that the quality of peer-feedback provided and received, respectively, predicted students’ quality of question-generation. Furthermore, the results from the three-step hierarchical regression analysis showed that the quality of peer-feedback provided and received in combination also predicted students’ quality of question-generation. Details of the significance of this study are provided, as well as suggestions for instructional implementations.

Keywords

Peer-feedback provided, Peer-feedback received, Online learning activity, Peer-assessment, Question-generation

Introduction

Student question-generation as a contemporary approach to teaching and learning, and areas in need of future work

Student question-generation (SQG) engages students in reflecting on a recent learning experience, and constructing questions around areas they deem personally relevant and important for self- or peer-assessment. Its pedagogical significance and value have been well established empirically. Specifically, based on the results of a meta-analysis of 109 empirical studies on SQG conducted in a wide variety of disciplines and at all levels of schooling (with the primary level taking up nearly two-thirds of the studies), there is wide support for its positive effects on a variety of learning outcomes (e.g., academic achievement, attitudes toward learning) (Yu, 2012).

Despite SQG’s solid empirical support, and sound theoretical foundations on the theories of self-regulated learning, constructivism, and self-determination, several factors affect its wider adoption in classrooms. In particular, studies show that primary school students have concerns about their capability and performance as related to SQG (Yu & Liu, 2005). There is thus a need to examine how to ease students’ concerns and provide adequate scaffolding for SQG by taking advantage of peer-assessment (PA), as this approach not only allows students to receive more timely feedback in a large class, but also encourages them to keep examining the quality criteria of the expected learning outcomes (Topping & Ehly, 1998).

Peer-assessment as a generative learning approach and the current research gaps

The cognitive conflict, social constructivism, and social learning theories can help elucidate why the various cognitive processes brought about by PA (including self-evaluation, self-correction, self-adjustment, and self-reflection through giving, receiving and responding to comments) promote students’ performance and cognitive growth (van Gennip, Segers & Tillema, 2010). Indeed, a growing body of empirical evidence since the 1990’s has shown PA to promote students’ critical thinking, cognitive development, and performance (Nelson & Schunn, 2009; Topping, 2010; van Gennip et al., 2010).
Efforts have been devoted to refine PA designs by identifying beneficial elements of PA tasks, such as the provision of clear and pre-specified criteria for objective comments and training (van Zundert, Sluijsmans, & van Merriënboer, 2010), the offering of both appropriate affective and cognitive feedback (Cho & Cho, 2011; Nelson & Schunn, 2009), and the inclusion of both quantitative and descriptive feedback (Topping, 2010; Yu & Wu, 2013). However, there are still a number of gaps in the related literature worthy of further investigation. In particular, few studies examine the effects of PA from both perspectives (i.e., the assessors and the assessed), or have clearly differentiated the effects of assessing peers from those of being assessed by peers (Topping, 2010; van Zundert et al., 2010). Moreover, those studies that have done so have all involved college or secondary school students. For instance, the results from Li, Liu and Steckelberg (2010), which involved undergraduate students engaged in computer-related projects, showed that the quality of comments given significantly predicted the performance of assessors, whereas the quality of received comments did not. The results from Cho and Cho (2011), which examined undergraduate students writing a science report on physics, similarly found that giving comments positively influenced the manner in which the assessors revised their own writing, although the effects of receiving peer comments were limited. Finally, Lu and Zhang (2012) and Lu and Law (2012) both examined secondary students and indicated that the cognitive feedback that was given strongly predicted how the assessors performed in their final projects, but was not related to the performance of the assessed student.

Topping (2010) noted that little research about PA was conducted in primary schools, and it was also suggested by van Zundert et al. (2010) that research on PA applied in contexts other than higher education should be a focus of future work, to extend the generalizability of PA. Since primary school students have the capability needed to engage in and benefit from PA (Hwang, Hung, & Chen, 2014; Yu, Liu, & Chan, 2005), by recognizing this gap in the literature the first focus of the current study is the respective effects of receiving and providing peer-feedback among primary school students.

Cognitive demands involved in providing and receiving feedback during peer-assessment: cognitive load perspective

It is generally believed that students can benefit from serving as both the assessor and the assessed in a PA context (Kollar & Fischer, 2010; Topping, 2010; van Gennip et al., 2010). On the one hand, the process of commenting on the strengths and weaknesses of assessed products may prompt the assessor to examine the evaluation criteria, and to develop knowledge of what constitutes good work and what needs to be avoided. This information can be further adopted by the assessor to monitor and regulate learning (Cho & Cho, 2011). On the other hand, those that are assessed can receive feedback from the assessor in a timely and personalized fashion, and make revisions to further enhance the quality of their work.

Despite the fact that simultaneously playing the roles of the feedback provider and receiver may yield some gains, playing both roles imposes cognitive loads on students. In short, while serving as the assessor, students have to study the assessed work first and then evaluate its quality against a set of criteria, followed by the construction of descriptive comments in a way that the assessed can sympathize with and benefit from. Similarly, when serving the role of the assessed, students have to process the comments provided, assess the validity and usefulness of each suggestion, and decide whether and how to integrate these to enhance their current work (Kollar & Fischer, 2010). As noted by cognitive load theorists (e.g., Paas, Renkl, & Sweller, 2004), any task composed of sub-tasks, each involving the activation and use of complex processes, demands cognitive capacity from the learner. If such tasks exceed the cognitive capacity of the learner, then this will cause a state of cognitive overload, which will lead to diminished learning effects.

Although there are studies that have students play only one role (e.g., Lundstrom & Baker, 2009), and others in which students play both roles, to the best of the authors’ knowledge, no works have produced empirical data substantiating any added benefits from students playing both roles during PA. Therefore, the second focus of this study was to examine the combined predictive effect of both aspects involved in PA while playing both roles (i.e., peer-feedback received and provided) on the quality of students’ question-generation. The combined effects of giving and receiving feedback on learning need to be substantiated to warrant the further use of such practices, especially for primary school students who, despite having reached Piaget’s (1926) formal operational stage of being able to manage the logical use of symbols related to abstract concepts, are usually still very limited in term of cognitive capacity when compared to older students.
Research questions

The following two research questions are examined in this work:

- Will the quality of peer-feedback provided and received, respectively, significantly predict primary school students’ quality of question-generation?
- Will the quality of peer-feedback provided and received, in combination, significantly predict primary school students’ quality of question-generation?

Method

Participants

Two hundred and thirteen fifth-grade students (10-11 years old) from eight intact classes of one public primary school in Taiwan participated in this study. With the current emphasis on a competency-based curriculum targeting core competencies, including problem-solving, creativity, ICT skills, and independent thinking (Taiwan Elementary and Educator Community, 2014), the introduction of SQG and PA was thus well received by the participating school.

Research design and implementation procedures

The correlational research design method was adopted to guide the data collection and analysis. Hierarchical multiple regression was used to detect the effects of predictor variables on a criterion variable, with the criterion variable being the students’ scores on the quality of question-generation, and predictors the scores on the quality of peer-feedback provided and received. Specifically, two-step hierarchical regression was employed to examine the predictive effect of peer-feedback provided and received, respectively, on the quality of question-generation, and three-step hierarchical regression to examine the predictive effect of peer-feedback provided and received, in combination, on the quality of question-generation, with significantly diminished learning effects connoting cognitive overload on the part of the learner, as suggested by cognitive load theory.

This study lasted for six consecutive weeks. Participants were informed that the online SQG and PA activity was intended to support their learning of two science units: (1) heat transfer and (2) sound and musical instruments. Since true/false and multiple-choice questions are among the most frequently encountered question types in primary schools in Taiwan, they were chosen for the focal online activities.

Each week, students headed to a computer lab to participate in a 40-minute online learning activity led by the implementer, after attending three instructional sessions on science taught by their respective science teachers in class. Training sessions that have been shown to be adequate for preparing participants of the same grade level to possess the fundamental skills required for question-generation and PA (Yu, Liu, & Chan, 2005) were scheduled in the first two weeks, before the subsequent four weeks of implementation. The experimental procedure is shown in Figure 1, and explained below.

A training session on generating the focal question types was arranged in the first week. Briefly, with reference to Rosenshine, Meister, and Chapman (1996), effective instructional elements supporting question-generation, namely the criteria and models of question-generation, were introduced (see “variables and measures” sub-section). In addition, based on the literature on test construction (Haladyna & Downing, 1989; Osterlind, 1998), basic principles for item writing and operational procedures for question-generation were illustrated in the Question-Authoring and Reasoning Knowledge System (QuARKS), followed by hands-on practice with the system.

At the beginning of the second session, whole-class feedback on students’ question-generation performance was given to highlight good question-generation practices. PA training was then given, in which models of constructive feedback with reference to the devised scoring scheme (see “variables and measures” sub-section) and built-in criteria in the online PA form (see “online learning system” sub-section) were explained. Moreover, the operational procedures for PA and details of how to access peer-feedback in QuARKS were explained before the students practiced giving peer-feedback using the online PA form, and then receiving peer-feedback via the notification system.
system. The remaining time was devoted to individual hands-on practice of question-generation and PA using the system.

Figure 1. Experimental procedures

Starting from the third week, as a routine, whole-class feedback on good question-generation and PA practices was highlighted using the work of 3-5 students from the previous session. Afterwards, students were given 20 minutes to individually generate at least one question for each of the two chosen question types in accordance with the instructional content covered that week. Then, with the system’s embedded automatic random assignment, students were directed to choose and assess at least two questions from a pool of eight peer-generated questions for each chosen question type for the remaining time (i.e., about 15 minutes).

Online learning system

The QuARDS online learning environment equipped with the usual advantages associated with computer and networked technologies (Yu, 2009) was adopted to support the learning activities carried out in this study. In brief, QuARDS allows students to contribute and benefit from the process of constructing question items of different types and media formats. Also included is a PA component to support question-generation activities. For PA, in particular, with reference to a set of built-in criteria (the bottom-left portion of Figure 2) and the scoring scheme for the quality of PA (see “variables and measures” sub-section for detail), the assessor first completes two rating scales with regard to: (1) the overall quality of the assessed question, and (2) the recommendation for the examined question to be included in the online drill-and-practice item bank (the top portion of Figure 2). The assessor then provides descriptive comments by typing in suggestions in a designated feedback space (the bottom-right portion of Figure 2).
To ensure the fluidity of the PA process, a notification system was put in place to automatically alert the question-author of any assessment updates. Specifically, an assessment is completed, the author of the question is alerted instantly by a blinking red icon on the screen. By clicking on the icon and the “proceed” button placed next to the notification, the author is transferred directly to the focal question, where its assessment can be viewed, and appropriate actions can be taken, such as making revisions.

![Image](image.png)

*Figure 2. Assessment form for the assessor to give feedback to the author*

**Variables and measures**

Three variables were examined in this study: quality of peer-feedback provided, peer-feedback received, and question-generation. First, for the assessment of quality of peer-feedback provided, all the descriptive comments an assessor provided to the respective assessed items were analyzed against a pre-defined scheme with scores of 0 to 4 (the scheme is discussed in later section of this section) by a rater (having experience with both the focal science learning material and analysis procedure). The scores gained on each item in the same session were summed up (i.e., sessional score), and then divided by the number of implementation sessions (i.e., 4) as the average sessional quality of peer-feedback provided. For illustration purposes, as shown in Table 1, assessor A assessed a total of 10 questions during the four implementation sessions. The scores gained by each item in each session were summed up (i.e., 2+3+7+9=21), and then divided by 4 (i.e., 21/4= 5.25) as his or her average sessional quality of peer-feedback provided during the activity.

<table>
<thead>
<tr>
<th>Implementation #1</th>
<th>Implementation #2</th>
<th>Implementation #3</th>
<th>Implementation #4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>Score</td>
<td>Item</td>
<td>Score</td>
</tr>
<tr>
<td>Assessor A</td>
<td>1-1</td>
<td>2-1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1-2</td>
<td>2-2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3-3</td>
<td>2</td>
<td>4-3</td>
</tr>
<tr>
<td>Sessional score</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Average sessional score</td>
<td>5.25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Second, for the quality of peer-feedback received assessment, all the descriptive feedback an author received from different assessors with regard to a specific question item was analyzed against the same scheme by the same rater, and averaged to yield feedback-received per item score (e.g., as shown in Table 2, (1+2)/2=1.5 for item 1-1, (2+0)/2=1 for item 2-1). The average scores gained on each item in the same session were then summed up (i.e., sessional score) across all sessions, and divided by the number of implementation sessions as the average sessional quality of peer-feedback received. For illustration purposes, as shown in Table 2, author A composed a total of 8 questions and received peer-feedback on 6 (i.e., no feedback received for items 1-2 and 2-2). For implementation session #1, author A received feedback from assessors B and C on item 1-1, but no feedback for item 1-2, and thus, the sessional score for implementation #1 is 1.5 (i.e., (1+2)/2=1.5). The average sessional score was then 2.5 ((1.5+1+3.5+4)/4= 2.5).

With reference to Nelson and Schunn’s (2009) study on feedback and Yelon’s (1996) open communication instructional principle with regard to providing feedback for students’ work, the quality of online peer-feedback provided and received was evaluated using the same scheme in terms of the following five discrete levels: meaningless comments (0); general comments (1); specific comments, where strengths and weakness are identified
(2); identification and explanations for improvement (3); and explicit suggestions for further refinement of questions (4). For example, as shown in Table 3, for question #3, although specific comments (i.e., two strengths and one weakness) were provided, it did not offer any explanation for question-refinement, which were offered for questions #1 and #2. Also, for question #2, not only were areas for improvement identified and explained, but explicit suggestions were offered, although these were not for question #1. As such, the quality of peer-feedback provided for question #3 was scored “2,” whereas for question #2 was scored “4,” and that for question #1 was scored “3,” based on the 5-level scoring scheme. The higher the score is, the better quality the comment provided and received is. High quality comments tend to be more detailed, specific and instructive, as attested by Nelson and Schunn (2009), who indicated that feedback of this nature would lead to better performance.

Table 2. An illustrative example of scores on quality of peer-feedback received

<table>
<thead>
<tr>
<th>Implementation #1 (2/1b)</th>
<th>Implementation #2 (2/1b)</th>
<th>Implementation #3 (2/2b)</th>
<th>Implementation #4 (2/2b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author</td>
<td>Assessor</td>
<td>Score</td>
<td>Author</td>
</tr>
<tr>
<td>A</td>
<td>1-1</td>
<td>B</td>
<td>2-1</td>
</tr>
<tr>
<td></td>
<td>1-2</td>
<td>C</td>
<td>x</td>
</tr>
<tr>
<td>Sessional score</td>
<td>1.5</td>
<td></td>
<td>1.5</td>
</tr>
<tr>
<td>Average sessional score</td>
<td>2.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.  
A Number of questions generated by Author A; B Number of questions being assessed by peers; C The feedback was rated “0”; D No feedback received from peers.

Third, with regard to the quality of question-generation assessment, and in reference to Torrance’s creativity index (i.e., fluency, flexibility, elaboration, originality) (1974), the cognitive levels proposed by King (1992), and the questions generated by students, a set of criteria was developed and operationally defined. Each question that the students generated during the question-generation activities was analyzed and scored against these criteria by the same experienced and qualified rater. Specifically, each question was graded along the following six dimensions:

- Fluency (0–3) assesses the correctness of wording and punctuation, clarity of meaning and logic, and relevancy of the constructed question.
- Flexibility (0–2) gauges the interconnectedness between the currently covered topic/unit and prior topics/units, and any self-derived examples.
- Elaboration (0–2) assesses the refinement of the questions in terms of creating scenarios for the question and using multimedia files to enhance understanding of the question.
- Originality (0–2) examines the uniqueness of a specific question as compared to those of peers in terms of innovative ways of formulating questions and content/ideas.
- Cognitive level (0–3) evaluates the cognitive levels demanded of the question-author, and these can be in terms of fact, comprehension or integration. Fact stresses the verbatim nature of questions from the learned materials, whereas comprehension indicates that students use their own words to define or describe the learned content. Integration shows that a link has been built across topics/units, and that explanations have been provided to build connections.
- Importance (0–1) evaluates the importance of the concepts assessed in the constructed question.

For example, as shown in Table 3, because all three example questions addressed the “correctness” and “relevance” elements, and the meaning of question #2 would be strengthened if the word “only” was included in the question (as suggested by one peer-assessor), it was scored “2” on the fluency dimension (due to not receiving a score on the “clarity” element), while questions #1 and #3 scored “3.” Besides, as all three questions tested the currently covered topic/unit, but did not refer to concepts in prior topics/units, and only question #1 included self-derived examples, question #1 was scored “1” on the flexibility dimension, and questions #2 and #3 received a score of “0.” Furthermore, as all questions #1–3 included self-created scenarios, each gained one point on the “scenario” element. Note that even though question #1 included a picture in one of the options, because it did not help to enhance the understanding of the item, as defined by the scoring scheme, no extra point was given for the “multimedia” element. Moreover, with more than one correct way to respond to the question (i.e., how to change the pitch of a guitar), the
combination approach used in question #3 was viewed as unique, as it was not seen in a question produced by any other student. As a result, among the three examples, question #3 is the only one that gained a score with regard to “innovative ways of formulating questions.” Finally, as all questions touched on important concepts in the study material and were paraphrased, but none of them included any explanations with regard to connections across topics/units, all questions scored “1” on the importance dimension and “2” on the cognitive level dimension (i.e., comprehension). All scores gained on each of the dimensions per question per week were summed up and then divided by the number of implementation weeks (i.e., 4) as the average weekly quality of question-generation. The higher this score is, the better the overall quality of the student-generated questions.

To show the relation between the rater, the assessor, and the assessed, a figure was created with the inclusion of the criteria used by the respective parties/roles for evaluating the quality of question-generation and PA (see Figure 3). Finally, to establish inter-rater reliability, 10 pieces of student work (i.e., questions generated and comments received with regard to these) were randomly selected from each of the 8 participating classes from 2 out of the 4 implementation weeks. The samples \((N=160)\) were evaluated by another independent rater, who was trained on the devised scheme and criteria before proceeding with the task. The results of the Pearson correlation between the two raters indicated that the coefficients were satisfactory \((r=.94, p < .01\) for PA and \(r=.73, p < .01\) for question-generation).

| Table 3. Examples of student question-generation and PA with reference to the respective scoring schemes |
|--------------------------------------------------|---------------------------------|---------------------------------|----------------------------------|
| SQG examples | Quality of question-generation | PA examples | Quality of PA comments |
| Question #1: Spiderman sees people selling things at the night-market. He hopes to buy an item that conducts heat the fastest. Which of the following items should he buy? | • Fluency: correctness (1), clarity (1), the relevancy (1) | + The question is set in a context (night-market) with an interesting comic book character. | Level 3: Identification and explanations for improvement |
| (A) A steel spoon | • Flexibility: self-derived examples (1) | + ‘Underline’ was used to highlight the main concept being tested in the question. | |
| (B) Bamboo chopsticks | • Elaboration: self-created scenario (1) | + All four options are items made of different materials with different levels of heat conduction. The question generated can assess students’ level of understanding of the study content (i.e., heat conduction). | |
| (C) Paper | • Originality (0) | - Only the correct option has a picture in it. It may give the answer away. | |
| (D) Plastic | • Cognitive level: comprehension (2) | + A question with a meaningful scenario | |
| Answer key: (A) | • Importance (1) | - A heater was the example included in the textbook. Other appliances frequently encountered in daily life (e.g., toasters) should be used to evaluate students’ level of understanding (not memorization) on this topic. Based on the answer key provided by the author, because both radiation and heat convection play a role in how heaters work, the question should be revised | Level 4: Explicit suggestions for further refinement of questions |
| Question #2: John uses a heater in his room in winter. His mom says that the heater works on radiation. Answer key: False | Fluency: correctness (1), the relevancy (1) | + A question with a meaningful scenario | |
| | Flexibility (0) | - A heater was the example included in the textbook. Other appliances frequently encountered in daily life (e.g., toasters) should be used to evaluate students’ level of understanding (not memorization) on this topic. Based on the answer key provided by the author, because both radiation and heat convection play a role in how heaters work, the question should be revised | |
| | Elaboration: self-created scenarios (1) | | |
| | Originality (0) | | |
| | Cognitive level: comprehension (2) | | |
| | Importance (1) | | |
Question #3: Mary is practicing on her guitar, but finds the pitch is off. What can she do to change the pitch of her guitar?
1. Change the tightness of the strings
2. Change the sequence of the strings
3. Change the thickness of the strings
4. Change the color of the strings
5. Change the length of the strings

(A) 123
(B) 125
(C) 135
(D) 124

Answer key: C

Note. *+ = Identified as strengths by a peer-assessor; - = Identified as weaknesses by a peer-assessor.

Figure 1. Relations between the rater, the assessor, and the assessed
Results

The correlations among variables, as well as the means and standard deviations of the examined variables, are listed in Table 4. As the assumptions with regard to multicollinearity and independent errors were not violated, as indicated by correlation between predictors (r = .13, p > .05), and all VIF (>1) and Durbin-Watson values (= 1.79), thus we can proceed to the hierarchical regression analyses.

Table 4. Correlations between examined variables and descriptive statistics (N = 213)

<table>
<thead>
<tr>
<th></th>
<th>Pre-assessment I^a</th>
<th>Pre-assessment II^b</th>
<th>Quality of peer-feedback received</th>
<th>Quality of peer-feedback provided</th>
<th>Quality of question-generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-assessment I</td>
<td>1</td>
<td>.32**</td>
<td>.30**</td>
<td>.43**</td>
<td>.56**</td>
</tr>
<tr>
<td>Pre-assessment II</td>
<td>.32**</td>
<td>1</td>
<td>.15*</td>
<td>.46**</td>
<td>.40*</td>
</tr>
<tr>
<td>Quality of peer-feedback received</td>
<td>1</td>
<td>.15*</td>
<td>1</td>
<td>.37**</td>
<td></td>
</tr>
<tr>
<td>Quality of peer-feedback provided</td>
<td></td>
<td></td>
<td>1</td>
<td>.54**</td>
<td></td>
</tr>
<tr>
<td>Quality of question-generation</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>5.06 (2.98)</td>
<td>3.82 (3.11)</td>
<td>6.62 (3.75)</td>
<td>5.36 (3.02)</td>
<td>5.75 (2.27)</td>
</tr>
<tr>
<td>Maximum scores</td>
<td>14</td>
<td>19</td>
<td>19.38</td>
<td>20</td>
<td>13.25</td>
</tr>
<tr>
<td>Minimum scores</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.75</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note. ^p < .05; **p < .01. ^The quality of question-generation at the training; ^b The quality of peer-feedback provided at the training; ^c indicates that the quality was rated to be of no quality based on the scoring scheme.

The predictive effect of the quality of provided and received online peer-feedback on the quality of question-generation

As shown in Table 4, the results from correlation analysis found significant relationships among the criterion variable, predictors, and students’ pre-existing abilities (i.e., quality of students’ question-generation and feedback provided during the training sessions). Thus, both pre-existing abilities were used as pre-assessments and entered into the regression model as step 1 to control for their effects before entering the predictor variables in the subsequent step(s).

The two-step hierarchical regression results showed that the quality of peer-feedback provided significantly predicted the quality of question-generation of the assessor (β = .32, p < .01), and thus students providing high quality feedback to their peers would be more likely to have high-quality questions.

Similarly, the quality of peer-feedback received, after controlling for the pre-assessment effect, had a significant predictive effect on the quality of question-generation of the assessed (β = .21, p < .01). In other words, this indicated that students receiving more specific and higher quality online feedback from the assessor would be more likely to have better quality questions.

The combined predictive effect of the quality of provided and received online peer-feedback on the quality of question-generation

A three-step hierarchical regression was conducted. As shown in Table 5, the quality of peer-feedback provided predicted a significant proportion of the variance with regard to the quality of question-generation of the assessor (R^2 = .44, F = 54.66, p < .01). In model 3, adding the variable of peer-feedback received significantly enhanced R^2 (R^2 change = .04, F = 16.53, p < .01), thus indicating that, together, the quality of online peer-feedback provided (β_{FP} = 0.33, p < .01) and received (β_{FR} = 0.21, p < .01) significantly predicted the quality of question-generation of the students who simultaneously played the role of both the assessor and assessed. In other words,
students’ engagement in both roles was more likely to produce high-quality questions than when playing only one role.

Table 5. Hierarchical regression analyses of the quality of peer-feedback with regard to predicting the quality of question-generation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th></th>
<th></th>
<th>Model 2</th>
<th></th>
<th></th>
<th>Model 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SEB</td>
<td>β</td>
<td>B</td>
<td>SEB</td>
<td>β</td>
<td>B</td>
<td>SEB</td>
</tr>
<tr>
<td>Constant</td>
<td>3.21</td>
<td>.26</td>
<td>.48**</td>
<td>2.62</td>
<td>.27</td>
<td>.38**</td>
<td>2.02</td>
<td>.30</td>
</tr>
<tr>
<td>Pre-assessment I</td>
<td>.37</td>
<td>.04</td>
<td>.48**</td>
<td>.29</td>
<td>.04</td>
<td>.38**</td>
<td>.25</td>
<td>.04</td>
</tr>
<tr>
<td>Pre-assessment II</td>
<td>.18</td>
<td>.04</td>
<td>.24**</td>
<td>.09</td>
<td>.04</td>
<td>.13</td>
<td>.08</td>
<td>.04</td>
</tr>
<tr>
<td>FP</td>
<td>.24</td>
<td>.05</td>
<td>.32**</td>
<td>.25</td>
<td>.05</td>
<td>.33**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td></td>
<td></td>
<td></td>
<td>.13</td>
<td>.03</td>
<td>.21**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>.37</td>
<td></td>
<td>.44</td>
<td>.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>60.76</td>
<td></td>
<td>54.66</td>
<td>48.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R² change</td>
<td></td>
<td>.07</td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F (R² change)</td>
<td></td>
<td>27.26**</td>
<td></td>
<td>16.53**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Dependent variable: question-generation; FP denotes peer-feedback provided to the assessed; FR denotes peer-feedback received from the assessor; *p < .05, **p < .01.

Discussions and conclusions

Several important findings were obtained from this work. First, the quality of online peer-feedback provided to the assessed was found to have predictive effects on the quality of the assessor’s question-generation. In other words, an assessor who provided higher quality online feedback to his/her peers tended to produce better quality questions. Students who engaged in cognitive processes requiring observation of peer-produced artifacts, assessing their quality based on a set of criteria, and providing high quality feedback that assisted their peers to improve the quality of their work, were found to have greater improvements in their own work. This finding supports the results of studies such as Cho and Cho (2011), Li et al. (2010), Lu and Law (2012), and Lu and Zhang (2012), and further substantiates the learning benefits of providing peer-feedback.

Secondly, the quality of online feedback received from the assessor was also found to significantly predict the assessed students’ quality of question-generation. When examined from the perspective of the assessed, students receiving higher quality online feedback appeared to then produce better questions. This result not only confirms what most researchers intuitively expect from PA (i.e., the positive effects of receiving peer comments) (Cho & Cho, 2011; Nelson & Schunn, 2009; Topping, 2010), although no supportive empirical evidence has been reported in the literature (e.g., Cho & Cho, 2011; Li et al., 2010; Lu & Law, 2012; Lu & Zhang, 2012), but also supports the findings of Nelson and Schunn (2009), suggesting that higher quality feedback (i.e., identification of problems along with clear suggestions for revision) can improve the quality of students’ work.

Third, the combined predictive effects of online feedback provided to the assessed students and received from their peer-assessor on the quality of question-generation were confirmed in this study. As shown in the current study, students providing and receiving higher quality peer-feedback while serving as both the assessor and assessed was found to be related to a higher quality of question-generation as compared to students only engaging in one role. On the basis of this result, it may be inferred that the cognitive overload that the authors were initially concerned about was not present in this study. To elaborate, as implied by cognitive load theorists, if the tasks involved in playing both roles exceed the cognitive capacity of the learner, this will cause a state of cognitive overload, leading to diminished learning effects. Nevertheless, as shown in the results of the three-step hierarchical regression, student performance at question-generation did not decrease, but instead was significantly enhanced when compared to that seen when the students only provided feedback. Possible reasons for why playing both roles simultaneously did not result in diminished learning effects, as implied by the cognitive load theory, are offered.

As noted in the Introduction section, by playing the role of the assessor, students are encouraged to make objective judgments on the quality of the work they are viewing, and give constructive comments about it. On the other hand, when playing the role of the assessed, the comments provided by the assessor need to be studied, assessed, and integrated, if deemed appropriate. On the surface, the total number of tasks that needs to be attended to exceeds that involved in playing any one role (i.e., feedback-receiver or provider), and thus cognitive load theory suggests that the
students may be overwhelmed by simultaneously processing all of the information elements associated with the learning tasks (Paas et al., 2004). Nevertheless, if analyzed more closely it can be seen that both roles revolve around the set of criteria that are used for assessing the quality of question-generation.

To elaborate, efforts at evaluating the quality of an item and constructing descriptive comments as feedback would be most productive if they are centered on the set of criteria for the quality of question-generation. Similarly, processing and assessing the validity and usefulness of each piece of feedback that is provided would be easier if these criteria are also considered. Moreover, the provision of feedback requires the assessors to understand and apply the criteria for the quality of question-generation, and also helps the assessors to incorporate the information elements addressed in the criteria in a schema. With reference to the cognitive load theory that suggests students’ schema availability influences the cognitive load imposed by a task (Paas et al., 2004), the schema of the criteria constructed during the feedback provision task may help to reduce the cognitive demand of the feedback reception task. Therefore, the tasks involved in being both a feedback-receiver and provider seem to be mutually beneficial, with students gaining fluency and mastery each time they engage in this practice. As such, even though each of the tasks to be executed involves the activation and use of complex processes, as the essential concepts are more or less the same and help schema formation, it may be that its total cognitive demand did not exceed the cognitive capacity of the learners to the extent of negatively affecting student performance.

Finally, the significant results from the three-step hierarchical regression further indicates that primary school students are capable of accomplishing the tasks associated with being a feedback-receiver and provider simultaneously, and can benefit from doing this.

**Scientific significance and implications of the study**

This study contributes to the literature on PA and has the following empirical and methodological significance. First of all, related studies have mostly drawn their inferences based on participants’ subjective perceptions of the learning improvements obtained as a result of feedback given by peers, as well as the participants’ attitudes towards PA (Wen, Tsai & Chang, 2006). This study moves beyond such research by adopting objective measures to correlate the quality of PA with the quality of the work that is produced.

Secondly, by adopting the perspectives of both the assessor and assessed, and by differentiating both the effects of providing and receiving peer-feedback on the quality of produced work, their predictive effects, both individually and in combination, with regard to elementary students’ quality of question-generation were empirically examined and supported.

Thirdly, because of the validation of these combined predictive effects, a long-held belief and practice associated with PA, based on the idea that learners benefit by playing the roles of both the assessor and assessed, was empirically substantiated for the first time.

Two suggestions are provided for instructional implementations. First, because the individual and combined effects of the quality of online peer-feedback provided and received were substantiated in the current study, ways to enhance the quality of peer-feedback should be explored in order to achieve the best results. Considering that feedback is important and is also a learned skill (Nelson & Schunn, 2009), possible ways to promote the provision and reception of high quality peer-feedback may include: context-specific assessment criteria regarding the target activity, training sessions with regard to the provision of quality feedback in accordance with a set of criteria, and whole-class feedback provided by the instructor highlighting exemplary PA practices, among others.

Second, instructors concerned about the possible cognitive overload that may result from students’ playing multiple roles (i.e., the assessor and assessed) and completing multiple learning tasks (i.e., producing work, providing and receiving feedback) during PA can rest assured, as the findings of this study indicate that primary school students can benefit from engaging in all the processes associated with the feedback-provider and receiver. As such, instructors should provide students with opportunities to both assess peer-generated work and receive peer-provided comments.
Limitations of this study

The current study examined the individual and combined predictive effects of online peer-feedback provided and received on fifth-grade students’ quality of question-generation. As noted earlier in this work, with the support of networked technologies it is possible to easily carry out automatic, random assignment of students’ work to be assessed and peer-feedback to be received. Another important point in the research design of this study is that fifth graders were recruited to participate for six consecutive weeks, and the assessors were given some freedom to choose with regard to which items to assess. As such, the generalizability of the results of this study to other contexts involving different age groups, or the employment of different media or work assignments, which may not be functionally equipped to support fluid interaction in PA, should be exercised with caution. Finally, although the students’ pre-existing abilities in the focal tasks (i.e., question-generation and providing feedback) were controlled by including them in the regression models tested, the readers are advised to note that this study involved intact classes and differences among classes and between genders were not accounted for, and that the state of cognitive overload was inferred by statistical testing rather than empirical testing in this study.

References


A Survey of Students’ Experiences on Collaborative Virtual Learning Activities Based on Five-Stage Model

M. Kemal Karaman¹ and Sevil Orhan Özen²*

¹New Media, Usak University, Turkey // ²Computer Education and Instructional Technologies, Usak University, Turkey // kemalkaraman@usak.edu.tr // sevil.orhan@usak.edu.tr

*Corresponding author

(Submitted March 4, 2015; Revised September 30, 2015; Accepted December 22, 2015)

ABSTRACT

In this study, we aimed to design collaborative virtual learning (CVL) activities by using a five-stage model (FSM) and survey of students’ experiences. The study group consisted of 14 voluntary students in the Turkish Teaching Department. In this case study, data were collected through observations, recordings in Second Life (SL) and interviews. For the data analysis part, descriptive analysis based on research questions and document analyses were used. In this context, the research is significant to regenerate and demonstrate SL how to use the FSM for the design of collaborative virtual activities.

Keywords

Second life, Virtual campus, Virtual learning, Collaborative learning, Five-stage model

Introduction

Low cost personal computers, advances in wide band web systems, wireless computing, and sound and video technologies supported virtual reality technology development in the 2000s and show an increase in the number of members using virtual environments or virtual worlds. Therefore, technological advances play a vital role in the spread of virtual worlds, for which many definitions exist in the literature (Andreas, Tsiatsos, Terzidou, Pomportsis, 2010; Çukurbaş, 2012; Yalcinalp, Sen, Kocer & Koroglu, 2012). In a general sense, a virtual-world enables the students to live a series of experiences by means of an avatar in a technological environment, which can potentially trigger authentic learning activities, because it is a representation of the real environment (Crisp, Hillier & Joarder, 2010). In virtual worlds, students can walk, run, teleport or fly from one place to another and they can build, purchase, rent or sell a three-dimensional virtual object or land. In this context, the students can interact with each other and other virtual objects in many virtual worlds such as OpenSim, Active Worlds, Multiverse and There. However, one of the prominent virtual worlds is Second Life (SL) (Kohler, Fueller, Matzler & Stieger, 2011; Kotsilieris & Dimopoulou, 2014). It is known amongst many users as the most developed social virtual world platform, and has become the subject of many studies (Crisp, Hillier & Joarder, 2010; Çukurbaş, 2012; Wang & Burton, 2013; Yalcinalp et al., 2012). In fact, many universities or educational institutions such as Texas State, Sheffield, Middle East Technical and Istanbul Universities have built education islands under the name of virtual campus in SL and have carried out education activities.

SL provides very strong group communication tools for students in order to establish social networks and groups (Kohler, et al., 2011). The students can assign roles within the group using avatars, and can also use multi-communication channels for laughing, dancing and applauding. Those who are in a different position geographically can send immediate messages (IM) to a group orally or in writing in order to receive assistance from other online students. The students in SL are able to take a role actively, rather than passively (Andreas et al., 2010). Another benefit of SL is that the barriers between the students and tutors are broken down with the use of SL. This is because the students interact with tutors and other students in SL more comfortably while collaborating during group activities (Hawkridge & Wheeler, 2010). For example, recently, the use of 3D worlds has allowed distance tutors to create activities in which students can collaboratively learn concepts by playing different roles (Bravo & García-Magariño, 2015). SL allows tutors to create collaborative learning environments (Loureiro & Bettencourt, 2014; Sutcliffe & Alrayes, 2012) because it contains many collaborative learning tools. The tutors can develop a SL environment by supporting collaborative learning and organize discussion groups and virtual conferences by using these tools (Çukurbaş, 2012; Loureiro & Bettencourt, 2014). In collaborative learning activities, the students share a common goal, depend on each other and are mutually responsible for their successes or failures (Mavridis, Konstantinidis & Tsiatsos, 2012). However, it is not a straightforward process to perform learning activities in collaborative virtual learning (CVL) environment. It is a highly complex process, especially when there is a lack of
appropriate methodologies (Bravo & García-Magariño, 2015). Thus, the tutors need a methodology or a model to conduct CVL activities in the common framework. In this study, we used a development model to plan CVL activities in SL.

The need for a model

Literature offers various models of online learning and teaching for development (Burgess, Slate, Rojas-LeBouef & LaPrairie, 2010; Puzziferro & Shelton, 2014). An example of such literature is Salmon’s five-stage model [FSM]. FSM is learning and teaching model assisting educators to design collaborative activities in SL (Salmon, Nie & Edirisingha, 2010). Furthermore, FSM is a development model made of five stages. The model was tested and developed by Salmon in the 1990s for teaching and distance learning for blended learning environments (Salmon et al., 2010). FSM as a learning and teaching model was published in Salmon’s study for the first time in 2000 (Figure 1). Salmon (2002) designed learning activities within the group and defined e-tivities for each stage of the model.

![Figure 1. FSM of teaching and learning online (Salmon, 2000)](image)

The model was adapted for education-training activities in the SL environment by Salmon and others (2010) (Figure 2). It was used for training in SL and tested in three different case studies carried out by archaeology, digital photography, media and communication students. Ultimately, it was discovered that each stage provided a different learning opportunity.
The first stage, which is called access and motivation, provides basic preconditions for effective participation of individuals. The second stage, which is online socialization, is the formation of virtual profile of learners in an online environment, and this stage also enables other learners to form communication. In the third stage, called information exchange, learners share beneficial information and thoughts related to learning content and tasks with other learners. The fourth stage, called knowledge construction, allows learners to undertake more complex tasks and form discussions. Interactions begin to contain more collaboration in this stage. In the last stage, known as the development stage, learners seek the means to benefit more from the system, and wish to obtain assistance in order to achieve their learning tasks, which allow them to apply and transfer the experiences they lived and the things they learned in the online environment. In each stage of the model, interaction and the learning level among the learners increase incrementally while passing to an upper stage (Salmon et al., 2010, p. 171).

Designing and implementing CVL activities through the use of a development model and virtual worlds require preparation, time and planning; also, one needs to promote and maintain the students’ participation in collaborative learning environments (Loureiro & Bettencourt, 2014). In this sense, the educators can benefit from a development model such as FSM to design CVL activities (Çukurbuş, 2012). FSM is a very practical, easy and useful model, because it allows for preparation, time, planning and maintaining participation for design and implementation of the CVL activities. Also, it has been adopted and adapted to different contexts by many educators who teach online. Thus, we aimed to design CVL activities based on FSM and survey students’ experiences using it with the aim of identifying the variables that may influence knowledge sharing in designed CVL by using a development model. The FSM was only used as a guide for the creation of CVL activities and SL was used as a learning tool for this aim. We wrote the following research questions to investigate this aim:

- What are students’ perceptions on CVL environment in SL?
- How do students describe their experiences on CVL activities in SL?
- How does CVL activities in SL based on FSM influence knowledge sharing?
Method

A case study was used as one of the qualitative research patterns in the study. The case study is searching an event, activity or a continuing process in detail using social units including one or several persons or classes, school and neighbours (Fraenkel, Wallen & Hyun, 2011, p. 440). Criterion sampling was used as one of non-probability sampling selection techniques for the students. It is a criterion that all the students should be voluntary for this study and furthermore all the necessary equipment for connection to SL should be ready for them whenever they want. Participants consisted of 14 voluntary students, nine girls and five boys, in second grade at the Turkish Teaching Department. The students’ experiences were surveyed. Descriptive analysis based on research questions and document analyses were used to analyse data collected from virtual world snapshots in Second Life, observations and interviews.

Data triangulation was made according to the research purpose by means of more than one data collection tools. The researchers took part actively in all processes and guided the students. The opinions of different experts conducting studies in SL were used in the interview form that consisted of 14 questions. The researchers interviewed each student directly after all stages of implementation were completed. Following the interviews, transcripts and notes taken during observations were read in detail so that a coding system could be created. The themes created from the descriptive analysis were categorized by research questions and the quotations were used for the data presentation.

The implementation

The study lasted for six weeks with groups determined by the students to conduct CVL. Each group determined a topic under the theme of virtual campus architecture. They visited places several striking in SL environment called virtual tours and took snapshots of various designs that they wanted to see in the virtual campus architecture such as buildings, entertainment places and conference areas. Then, they created a virtual presentation of an individual product and a virtual cube as a common product by using snapshots of their virtual tours based on their topics (Table 1).

Table 1. The tasks and study topics of group numbers

<table>
<thead>
<tr>
<th>Group number</th>
<th>Study topic</th>
<th>Tasks for virtual tours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Architecture and travel: Various countries</td>
<td>Monuments Universities Towers Buildings</td>
</tr>
<tr>
<td>2</td>
<td>Entertainment places</td>
<td>Play locations Beaches Entertainment places Parks, gardens Hotels</td>
</tr>
<tr>
<td>3</td>
<td>Exhibition areas</td>
<td>Nature images Animal images Cities Art galleries Portraits</td>
</tr>
</tbody>
</table>

The group studies as CVL activities in SL were designed weekly based on FSM stages and were detailed (Table 2). Furthermore, Sutcliffe and Alrayes (2012) argue that SL can be used to prepare virtual gatherings and presentations. The students followed all these processes such as gathering times, announcements of events and conferences in SL by means of a closed Facebook group formed by the researcher. In this way, the researchers and students continued to communicate in SL as well as in Facebook, but Facebook was only used to support communication and the announcement of the event.

According to Table 2, the students met twice with researchers in the laboratory. The purpose of the first gathering was to provide information about SL. During the second gathering they logged into the SL in the laboratory. Each
student used a computer and headset during this process. And then, students logged into SL three times as a group and twice individually from any computer in any location they found.

<table>
<thead>
<tr>
<th>Week</th>
<th>FSM stages</th>
<th>Process</th>
<th>Environment</th>
<th>Mean time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 wk</td>
<td>Access and motivation</td>
<td>Orientation</td>
<td>Laboratory</td>
<td>one and half an hour</td>
</tr>
<tr>
<td>2 wk</td>
<td>Online socializing</td>
<td>Virtual conference</td>
<td>Laboratory-SL</td>
<td>30 min. and 1 hour</td>
</tr>
<tr>
<td>3 wk</td>
<td>Information exchange</td>
<td>Group gatherings</td>
<td>SL</td>
<td>from 20 to 40 min.</td>
</tr>
<tr>
<td>4 wk</td>
<td>Knowledge construction</td>
<td>Individual virtual tours and products</td>
<td>SL</td>
<td>from 30 to 50 min</td>
</tr>
<tr>
<td>5 wk</td>
<td></td>
<td>Group gatherings and common products</td>
<td>SL</td>
<td>from 20 to 40 min</td>
</tr>
<tr>
<td>6 wk</td>
<td>Development</td>
<td>Group gatherings</td>
<td>SL</td>
<td>from 20 to 40 min</td>
</tr>
</tbody>
</table>

**Access and motivation**

As mentioned before, the researchers held an orientation gathering in the laboratory during the first week and at the first stage. They informed the students about basic avatar movements such as sitting, walking, running and flying. They also told the students how to download SL and various other essential information such as interface information. The orientation process lasted for one and half an hour.

**Online socializing**

The students met with the researchers in the laboratory during the second week and at the second stage to attend a virtual conference about various islands and communities that could be visited during virtual tours. The virtual conference, announced in the Facebook group, was given orally in a conference area in Infolit iSchool Island by an educator conducting studies in SL at the Sheffield University, which lasted 30 minutes (Figure 3).

Before the conference, the students practiced such processes as adding friends, accepting friendships, and viewing online persons nearby, teleporting to different places where other avatars live, joining groups, the customization of avatars and as well as virtual objects. This process lasted for one hour and the online socializing stage lasted for one and half an hour in total.

**Information exchange**

The groups held their first virtual gatherings in the third week and at the third stage. The researchers attended the virtual gatherings via SL. Each group chose a group name, a group leader and a study topic for virtual campus architecture. They also shared tasks during the virtual gatherings. The researchers educated the students how to use
snapshot feature and presentation tools of SL and how to build a virtual cube and various locations in SL. Also, the group members exchanged information with each other by brainstorming about their study topic. The mean time for group gathering varied from 20 to 40 minutes.

Knowledge construction

The groups carried out their virtual tours and took snapshots of SL during the fourth and fifth weeks at the fourth stage of FSM. They created individual products by using both snapshot feature and presentation tools. The students individually made inputs twice in the SL at this stage. The amount of time it took for the virtual tours varied from 30 to 50 minutes.

Subsequently, the virtual tours and individual products were collected again. The groups examined members’ individual products and brainstormed about which snapshots were to be transferred both to cube faces and to reflect the study topic selected by the group. The groups built the virtual cubes from their snapshots images as common products in SL. These gatherings lasted about from 10 to 40 minutes depending on ability of the group. Figure 4 shows both presentation and virtual cubes made by the students.

Development

The students and the researchers attended the third gathering at the first stage and brainstormed about group studies and their CVL experiences. Also, the researchers informed them how to sell their products by using Linden dollars in SL. If the researchers bought the group products, the copyrights would be paid for them to be used in virtual campus architecture. The students were informed beforehand about the usage of the Linden dollar in real life as an exchange rate. This encouraged the students to make the connection between the real life and Second Life and to discover how to use SL for different opportunities. The mean time for group gathering varied from 20 to 40 minutes.

Findings

Data sets were analysed and encoded according to research questions. Table 3 shows the categories, themes and codes obtained from interviews. Quotations from the expressions of the students were used to present the findings. These were expressed by coding as [K1, K2…] and also supported by the observations of the researchers.

According to Table 3, Perception of CVL activities in SL is the first category consisted of two themes, the environment used for the CVL activities and the usage areas of this environment for the first research question. Experiences of CVL activities in SL are the second category consisting of two themes, the positive and negative experiences of CVL activities for the second research question. And the last category is Effect of knowledge sharing on CVL activities in SL, consisting of three themes, the effect of experiences, the effect of avatar selections and the effect of environment for the third research question.
Table 3. Categories, themes and codes

<table>
<thead>
<tr>
<th>Categories</th>
<th>Themes</th>
<th>Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceptions of CVL activities in SL</td>
<td>Environment used</td>
<td>representing real environment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>different environment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>unlimited environment</td>
</tr>
<tr>
<td>Usage areas of SL</td>
<td></td>
<td>language education</td>
</tr>
<tr>
<td></td>
<td></td>
<td>virtual festivals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>association among universities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>higher education level</td>
</tr>
<tr>
<td></td>
<td></td>
<td>interest in technology</td>
</tr>
<tr>
<td>Experiences of CVL activities in SL</td>
<td>Positive experiences</td>
<td>entertaining</td>
</tr>
<tr>
<td></td>
<td></td>
<td>arousing curiosity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>unprecedented experiences</td>
</tr>
<tr>
<td></td>
<td></td>
<td>actual experiences</td>
</tr>
<tr>
<td></td>
<td></td>
<td>easy to use</td>
</tr>
<tr>
<td></td>
<td></td>
<td>presence</td>
</tr>
<tr>
<td></td>
<td>Negative experiences</td>
<td>boring alone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>complicated</td>
</tr>
<tr>
<td></td>
<td></td>
<td>immersion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>difficult in practice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sense of being lost</td>
</tr>
<tr>
<td>Effect of knowledge sharing on CVL activities in SL</td>
<td>Effect of experiences</td>
<td>language deficiency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>technical deficiency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ignoring of tasks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>time-consuming</td>
</tr>
<tr>
<td></td>
<td></td>
<td>indecision of common runtime</td>
</tr>
<tr>
<td></td>
<td>Effect of avatar selections</td>
<td>like themselves</td>
</tr>
<tr>
<td></td>
<td></td>
<td>like a student</td>
</tr>
<tr>
<td></td>
<td></td>
<td>nonhuman avatars</td>
</tr>
<tr>
<td></td>
<td></td>
<td>their mood</td>
</tr>
<tr>
<td></td>
<td></td>
<td>their taste</td>
</tr>
<tr>
<td></td>
<td>Effect of environment</td>
<td>improve self-confidence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increase motivation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increase productivity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increase creativity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increase interaction</td>
</tr>
</tbody>
</table>

Perception of CVL activities in SL

According to Table 3, two themes were obtained in this category from the interview data: environment used and areas of usage. The students indicated their perceptions under the theme of environment used through three codes including representing real environment, different and unlimited environment with the following expressions:

Most of the objects were close to reality, sir; three-dimensional states of the drawings were not far from real life and were not too simple, they were quite professional... The conference given to us in SL had affected me because the instructor giving the conference was far away and we listened to him in a virtual environment comfortably. It was a very different and comfortable environment in terms of traditional ways [K4].

Traveling on a helicopter was really fun. It was easy to use and made our work quite fun... It gives you an extraordinary feeling. There are worlds and persons different than the real world but these are live worlds......It is a nice feeling that you are able to do many things you cannot do in real life. You can do things that are difficult to do in real life or things that are impossible to do with a few buttons [K8].
Changing the places of huge cubes without getting help from anyone feels different. Besides, going from one place to another by teleporting, flying is a very nice feeling. You can go anywhere and whenever you want in SL, and see and learn, but there is a limit in traditional education [K11].

In addition, they indicated their perceptions under the theme usage areas of SL through five codes including language education, virtual festivals, and association among universities, higher education level and interest in technology with the following expressions:

Those who want to learn a language (English), can talk to foreigners and improve themselves vocally or silently…. I build a house on my own and travel, I talk to different people. Especially foreigners... For example, the language is important to me. Those who want to learn a language (English), can talk to foreigners and improve themselves vocally or silently [K7].

We could find a fellow university from another country, and our president and dean could give a seminar there for us.... There could be more students. Spring festival, etc. could be organized there [K10].

SL can address any individual interested in technology for the target group; however, it can’t be applied to a group apart from the university level. A student expressed this situation as follows:

I think that the students to be selected in this matter should be interested in computers (I am not talking about Facebook, etc., they should be really interested in computers) or students in the computer and technology departments should be selected. I think that people who are going to deal with this work without enjoying it will get bored quickly. In the end, it is a technological virtual environment; it doesn’t matter if it is SL or a different environment. I think that it is an environment which will contribute to anyone who is interested in technology even if scarcely... I don’t think that this environment can be applied to any children before the university level. It has a really difficult and complex system. I think that it will be too hard for a high school level person [K1].

Experiences on CVL in SL

According to Table 3, two themes were obtained in this category from the interview data: positive experiences and negative experiences. The students indicated their positive experiences through six codes including entertaining, arousing curiosity, unprecedented and actual experience that could not be done in real life, presence (feeling like they are really there), and easy to use (feeling familiar with the internet and computer) with the following expressions:

Traveling on a helicopter was really fun. It was easy to use and made our work quite fun ... It gives you an extraordinary feeling. There are worlds and persons different than the real world but these are live worlds .......It is a nice feeling that you are able to do many things you cannot do in real life. You can do things that are difficult to do in real life or things that are impossible to do with a few buttons [K8].

You feel like you are over there and next to each other; it is different and more useful than normal virtual world sites [K6].

In fact I did not have much difficulty because I have been using computers and internet for years [K3].

Moreover, they indicated their negative experiences through five codes, including difficult in practice, boring alone, complicated, immersion and sense of being lost. In this regard, two students expressed that they find SL complicated because they were experiencing such an environment for the first time; four students indicated that they had difficulty at first but got used to it over time and three students stated negative feelings because of being lost. Also two students stated that SL disconnects you from real life with its immersion. The students expressed their negative experiences as follows:

It was a little hard to learn basic skills at the beginning, short ways of flying and teleport were a little tough. Of course, walking was the hardest [K2].

I also had difficulty when I first used it and had trouble in sitting and getting up. I had to sit down places which I am not supposed to sit in at all and started to fly and could not land [K9].
The adaptation process was hard at the beginning, because it was an environment I haven’t seen before. Therefore, I had problems in comprehending and moving to practicing later on. I got lost in most of the places and could not reach the places I wanted [K4].

It can carry you away from life really. I personally would not want to come from a world where I am thinner and taller [K8].

According to the researcher’s notes on snapshot, it was observed that an advisor in SL comforted the students and facilitated their tasks; the students were apprehensive in the beginning since they were experiencing an environment that they didn’t know. However, according to observations on the students’ experience, it can be said that the students improved their motivation as they practiced more in SL.

Effect of knowledge sharing on CVL activities in SL

According to Table 3, three themes were obtained in this category from the interview data: effect of experiences, avatar selections, and environment. The students indicated effect of their experiences on knowledge sharing through six codes, including language deficiency, technical deficiency, ignoring of tasks, time-consuming and indecision of common runtime. In this regard, some students indicated that SL is boring and time-consuming alone, and six students expressed that there was neglect of tasks, including reaching a decision for the gathering of the group at the common runtime. These situations create a problem with sharing knowledge in the group. Those are reflected in students’ comments as follows:

I had difficulty a little since I don’t know a foreign language...They were all foreigners, we said hi, ran away, you know about the language problem. I think they sent something [K8].

The play requires high levels as a graphic card. It causes freezing in computers with poor hardware [K1].

In a work made in collaboration, even if one person neglects their tasks, the group becomes negative [K5].

You should have a long time on an individual basis; it is not easy to learn almost everything for SL [K6].

Unfortunately my friends did not make time for it, I participated on my own....I had difficulty in doing many subjects on my own, but if my group friends had entered regularly, we could have created a more pleasant and successful environment [K1].

The students indicated effect of avatar selections on knowledge sharing through five codes, including an avatar which looks like themselves or a student or a nonhuman being and which is suitable for his/her mood and his/her taste. When selecting the avatar in SL, seven students preferred avatars resembling themselves and other students preferred avatars like a student or a nonhuman being. It was also stated that they form their avatar profile by making changes to avatars according to their mood at that moment or according to their tastes. According to the researcher’s notes on snapshot, it was observed that the same avatars were selected and therefore they made changes to their avatars when they noticed this situation. Also it was observed that the students used SL as a socializing tool for gathering different persons by using their avatar. In this context, the students have cared to use different avatars in the online community for knowledge sharing.

Furthermore, the students indicated effect of environment on knowledge sharing through five codes, including improvement of self-confidence, productivity, creativity and enhancement of motivation and interaction with the following expressions. In this regard the students expressed they improved their productivity, creativity and self-confidence, gaining different perspectives because they shared knowledge without interruption. Their comments are stated as follows:

We had a continuous opinion exchange. It was important to care for each other’s opinions [K2].

Visiting a place you want and creating some things makes you have creative ideas [K5].

255
It focuses the students on learning; the students are able to do some things themselves. SL is a plus in comparison to traditional education [K6].

Your self-reliance improves, first of all. I generally acted with my group, but reaching decisions alone enables one to think reasonably and plan the reason and result of the event. If I am alone in a place I am not familiar with, I can decide what to do [K11].

Also, the students expressed that working in SL as a group was motivating because they completed tasks in a short time through helping each other. Some students indicated that SL was effective in improving national and international interaction taking place between people from different nationalities.

Getting recognition about the work you were going to do as you share opinions with the group motivates you even further [K5].

There were cubes we needed to form in the campus. We were supposed to take pictures and paste them to each face of the cubes. We attempted to hose cubes individually first but this had lasted nearly two weeks and we still could not finish it. Then we entered all together as a group and did that, and sir, what we could not do within two weeks, we finished within one or two hours [K10].

When you get together, everything is easier, and you get information from one another and refer [K11].

We are informed about the opinions of others and quality of communication improves [K9].

I met many different people. I added one from Facebook already from Peru [K3].

It is very effective and pleasant to socialize and meet new people...In addition, when you work as a group with your classmates, interaction improves [K4]

In normal conditions, it is difficult to communicate with an instructor in Ankara; it would be troublesome for him or her to give us a lecture. But in SL we can communicate with many instructors in the country and world...I can say that it is effective for our socialization in terms of gathering different people [K9].

According to the researcher’s notes on snapshot, it was observed that the students made suggestions to one another about their tasks in the virtual gatherings and shared virtual objects. Also it was seen that two groups had difficulty getting together for their gathering, therefore one of the groups did not complete their tasks, while the other group, which consisted of the students with knowledge in the group, completed the common tasks. However, it was witnessed that the group members completed the tasks in a shorter time when they got together in virtual gatherings.

Conclusion

In this study, we aimed to design collaborative virtual learning (CVL) activities in SL by using a five-stage model (FSM) and survey of students’ experiences. The findings were accessed by using descriptive analysis based on research questions and document analysis for data obtained through observations, recordings in SL and interviews.

The students worked individually in one part of the implementation process and they made an effort to work as a group for a common purpose in the other part. In virtual gatherings, the students exchanged information to reach the common purpose in the framework of counseling the researcher, and helped one another, and formed a common product in SL and made changes to improve the product by discussions. Finally, the students tried to reveal their perceptions, experiences of CVL activities and the factors affecting their knowledge sharing. In this context, it was found that the students encounter a variety of problems in the process. Some of these problems are technical deficiency such as walking, flying and teleporting in the basic moves of SL. However, it can be said that the students can connect to SL successfully and access the study environment easily through the orientation where basic skills are provided. And also, they obtained basic information through consultancy given by the researcher. Subsequently, they practiced individually in the progressing stages. So, Çukurbaş, Bezir and Karamete (2011) in SL, revealed the
importance of an advisor present in the environment, and Salmon, Nie and Edirisingha (2010), also, indicated that the role of the advisor is significant.

Other striking findings shows the fact that students think SL environment as boring while they are alone in the environment, and that they feel as if they were lost and that they think SL tears them apart from real life. In addition, they tried to establish a virtual profile through their own avatar chosen and they attempted to communicate with other avatars so as to socialize and share knowledge in the SL environment. The students had difficulty shaping and dressing their avatars in the SL environment and this was mentioned in the researcher observation notes, and the findings also showed that the students who selected the same avatar tried to change their selection. The fact that the students attempted to change their avatars due to being identical with other avatars may show that they internalized the environment and associated it with real life. There are similar findings in the study of Çukurbaş (2012). Apart from these, the students had difficulty communicating with other avatars because of foreign language problems.

Furthermore, it was observed that there were difficulties with the common product of only one group. The difficulties the students experience such as time consuming, indecision of common runtime and ignoring of tasks in knowledge sharing might be the reasons of this problem. Therefore individual responsibility, participation of the advisor and solutions to the difficulties mentioned above have prominent role in completing the CVL activities in SL during group work. Despite all these difficulties experienced by students, it could be said that the students gained positive opinions about SL such as entertainment and curiosity by gaining experience in SL environment in general and by the elimination of technical deficiencies that they experience. It could be stated that students’ motivation, self-confidence, productivity and creativity improved by noticing opportunities of SL such as representation of real environments, presence, gaining actual and unprecedented experiences. Moreover, the students made suggestions for the usage of SL in education. The students found SL beneficial for forming collaborations with universities from different cultures without recognizing any time and place limit and especially as a vital tool for the development of language education. In addition, the students expressed that the target group addressed by SL must be individuals interested in technology and virtual environments such as SL, and that it is not appropriate for individuals below the university level. The studies in Turkey were mostly about university students for the usage of SL for educational purposes (Bulu, 2012). However, it is possible to see studies conducted among people at different ages for SL in the literature in recent years (Bezir & Baran, 2014). In fact, SL had limited the usage of certain words in SL by persons in 13-15 age range and access to certain adult regions by age validation in 2005, but it closed this limitation in 2010 (Harrison, 2010). Moreover, there are various virtual worlds similar to SL for age groups in different levels (Quest Atlantis, Minecraft, etc.). In essence, the students associated their experiences in virtual campus with real life, and developed opinions about effective use of SL. In addition, the students supported the idea of continuing virtual campus with more students or volunteers. Some students thought that the use of virtual environments such as SL would be beneficial in the future for their own students because we are in the age of technology, and some thought that such technologies would contribute to university development. If the virtual campus established in the research scope continues, the products generated by the students will be important for being a guiding spirit.

Loureiro and Bettencourt (2014) conclude that the learners in virtual environments tend to feel more confident, open, creative and participatory to learning. The results in this study are somewhat similar. Although it was a small group we can say that the students have a more open and free attitude because they can complete their tasks and communicate with friends or other people in different places by virtual conferences and virtual gatherings in these environments whenever they want. So, the students stated that they are able to go wherever they want and learn about those places, without being exposed to boundaries of traditional education. When we consider all the mentioned problems, it is pointed out that the problems encountered in the use of educational SL in the last decade haven’t changed. Those problems which are still unsolved today might be the reason for the loss of attractiveness of the SL in the educational researches over time. Well what can be done to solve these problems?

In this context, it is suggested that the similar studies may be conducted on different subject topics and the common difficulties encountered in those studies may be compared with each other in order to achieve a solution to these problems. Another solution may be the blending of traditional education with education opportunities presented by SL environment. It is also proposed that technical problems be dealt with before putting the SL into application. Using a model as FSM while designing learning activities in SL may improve perspectives of educator and instructional designers. So in this process, successfully completion of each stage in this model is of great importance to achieve applicable results. All the problems encountered in the process should be minimized. In particular, the impact of knowledge sharing should be evaluated using quantitative, qualitative, and mixed methods.
Acknowledgements

The research was co-funded by Usak University Scientific Research Projects Coordination Unit (Project No: 2012/SOSB004).

References


Effects of Sharing Clickers in an Active Learning Environment

Todd Daniel¹ and Kristin Tivener²*
¹RStats Institute, Missouri State University, Springfield, MO, USA // ²Faculty in the Department of Sports Medicine and Athletic Training, Missouri State University, Springfield, MO, USA // ToddDaniel@MissouriState.edu // KTivener@MissouriState.edu

*Corresponding author

(Submitted April 15, 2015; Revised September 21, 2015; Accepted November 11, 2015)

ABSTRACT

Scientific research into learning enhancement gained by the use of clickers in active classrooms has largely focused on the use of individual clickers. In this study, we compared the learning experiences of participants in active learning groups in which an entire small group shared a single clicker to groups in which each member of the group had his or her own clicker. Our findings indicate that when clickers are used as part of an active learning activity, asking students to reach consensus in small groups is an effective way of using clickers to achieve the positive outcomes associated with active learning.

Keywords

Clickers, Active learning, Interactivity, Multidisciplinary

Introduction

Active learning has attracted dedicated advocates among faculty as research into this teaching pedagogy continues to demonstrate positive outcomes (Prince, 2004). Active learning describes any instructional method that engages students in the learning process through meaningful and thought-provoking activities (Goodsell, 1992). The theory underlying active learning is that students’ transition from being passive recipients of lecture-based information to becoming actively engaged in their own learning through a cooperative process between themselves and their instructor.

Engagement in active learning could be achieved through methods such as collaborative work, group discussions, case studies, or directed small-group activities. Clickers are educational technology tools that are commonly used to facilitate students’ interaction with an instructor in real time. In the following section we review literature on clickers and their use in active learning environments. The literature informs our presentation of six hypotheses related to potential effects on student participation when a group of students, working collaboratively, are forced to share a single clicker. Other studies have established the effectiveness of clickers in improving knowledge acquisition (Stowell & Nelson, 2007; Tivener & Hetzler, 2015). The present study adds to the literature by examining the experiences of participants during the active learning process.

Background

In recent years, the use of educational technology in higher education classrooms has grown as a strategy for increasing engagement in an active learning environment (Brians, Dounoucos, & Frye, 2012). At the forefront of these educational technologies is the clicker. Companies that market clickers call them student response stations, audience response systems, interactive voting systems, and electronic voting systems (Cain & Robinson, 2008). However, the word clicker is the simplest and most accessible nomenclature for a hand-held device that allows students to interact electronically with an instructor by instantly responding to an instructor-generated question commonly projected on a screen through Microsoft PowerPoint. Clickers may be specialized individual keypads or be installed through an application on a personal cell phone.

Clickers are commonly used in large-enrollment classrooms to permit students to contribute their individual opinions or respond to questions during class. In one study (Sternberger, 2012), 72 students enrolled in an undergraduate healthcare course responded to a 22-item questionnaire at the conclusion of a semester-long course that integrated clicker questions into lectures. Students in this study reported that the clickers significantly improved their...
Clickers have also been demonstrated to improve the interactivity in the classroom and to support two-way interaction between the students and instructor (Liu, Liang, Wang, & Chan, 2003). Measuring students’ attitudes about a class, Eastman, Iyer, and Eastman, (2011) reported that students believed that clickers prompted them be more attentive in class, increased their comfort level with participation, and created a more enjoyable classroom environment. Additional positive outcomes with clickers include improved learning outcomes (Mayer et al., 2009; Terenzini, Cabrera, Colbeck, Parente, & Bjorklund, 2001), increased class attendance, (Mastoridis & Kladidis, 2010), and greater positive emotional response to the course experience (Stowell & Nelson, 2007).

Currently, the research on clickers has been limited to research designs in which each student had a clicker to him or herself, and to designs in which clickers were one component of an active learning environment. For instance, one study (Knight & Wood, 2005) compared two sections of an undergraduate introductory biology course, taught by the same instructor, with approximately 400 students each. In the first section, course content was delivered through verbal lecture presentations and in the second section, the format was more interactive and included collaborative work in student groups, use of in-class formative assessment, individual clickers, and group discussion. Students in the more interactive section showed significantly higher learning gains; however, it was impossible to assess whether there was a single component of the active learning format that contributed most to these positive outcomes. When multiple active components – such as small group collaboration and large group discussion – are included in the course, the impact of an individual clicker may be overshadowed by the other learning modalities.

Several other studies (Martyn, 2007; DeBourgh, 2008; Mayer et al., 2009; Armbruster, Patel, Johnson, & Weiss, 2009) reported positive outcomes on student knowledge, interactivity, engagement, and performance in instructional designs with multiple forms of active learning, including the use of an individual clicker. Each of these studies, however, was agnostic as to which form of active learning most affected the positive outcomes.

Tivener and Hetzler (2015) studied the effect of individual clickers on knowledge and interactivity among athletic training students in a clinical skills class. The class incorporated small group discussion, hands-on skill application, and group case analyses. While they found significant gains for interactivity among students, the researchers questioned whether those gains resulted from use of the clickers or as a response to the transition from a passive lecture format to the active learning environment.

The present study was prompted by the uncertainty, expressed in multiple studies, whether clickers are an effective tool when combined with other active learning strategies. This study examined student experiences with clickers in a multidisciplinary forum when the active learning format included both small group collaboration and large group discussion. The purposes of this study were to (a) determine how the degree of individual interactivity in a large group discussion was affected when a small group used a single clicker to provide collaborative group responses; (b) assess whether student participation in small group collaborative table discussion was adversely affected when a group of students shared a single clicker; and, (c) assess student preferences for individual versus group clicker use.

Overall, the aim of the study was to determine whether the positive outcomes associated with active learning were similar when an entire small group shared a single clicker as to when each person used an individual clicker. We advanced six hypotheses to inform our research:

Hypothesis 1: Requiring a group of students to share a single clicker will not impair group member’s awareness of other group members’ opinions.

Hypothesis 2: Requiring a group of students to share a single clicker might inhibit individual participation.

Hypothesis 3: Requiring a group of students to share a single clicker will inhibit individual students’ contributions to the group.

Hypothesis 4: Requiring a group of students to share a single clicker will reduce students’ perception that the group clicker response accurately reflected their own opinion.

Hypothesis 5: Requiring a group of students to share a single clicker will reduce interactivity in the group.

Hypothesis 6: All participants will prefer to have their own clicker rather than share a clicker with the group.
Methods

Participants

Participants in this study were recruited during a multidisciplinary educational forum hosted by the College of Health and Human Services at a 4-year public university in the Midwest. Of the 204 forum attendees, 82.8% were female ($n = 169$), mostly undergraduate junior or senior status ($n = 164$, 80.4%), with a median age of 21 years old. Students were informed that data would be collected about their participation and were given the opportunity to opt out; all of the students elected to participate in the multidisciplinary forum and use clickers to record their answers. Participants were randomly assigned into individual clicker status (each participant having her own clicker) or group clicker status (up to 8 participants sharing a single clicker at their table). This arrangement resulted in 39 students assigned to a small group in which they had an individual clicker and 165 students assigned to a small group in which they would share a clicker among them.

Procedure design

Thirty tables were set up for the multidisciplinary forum. Before the forum began, researchers randomly selected 5 tables to be individual clicker tables and the remaining 25 were group clicker tables. Participants were randomly assigned to a table number upon their arrival at the forum. Students in the group clicker condition shared a single clicker (Turning Technologies, Youngstown, OH) at their table, with a median of four participants per clicker; students in the individual clicker condition were each given their own clicker at their small table. Every participant received a paper survey upon which they recorded their individual responses during the study.

The two-and-a-half hour forum began with a 30-minute keynote presentation introducing the topic of multidisciplinary healthcare. The remainder of the forum comprised three rounds of case vignettes. Each round of vignettes consisted of a presentation about a patient-provider case that was acted out by students from the theater arts department, followed by a 15-minute discussion at each table about questions raised by the vignette. Immediately following the 15-minute collaborative small table discussion, participants responded to three questions displayed on a PowerPoint presented to all participants in the forum.

The first two questions asked about the ease of creating a team care plan for the patient in the vignette and were presented with a 5-point Likert-type response. The third question was multiple-choice and asked participants to identify what they thought was the most essential component of creating a team care plan for that vignette. Participants were given 60 seconds to respond to each question and were given a verbal prompt before moving to the next question to assure that all participants were given adequate time to answer. The same set of three questions was posed after each of the three vignettes.

Participants in the individual clicker condition ($n = 39$) provided their responses using their own individual clickers. Participants in the group clicker condition ($n = 165$) provided a group-consensus response using the single clicker on the table. Next, the forum speaker led a 10-minute large group (all forum) discussion about the vignette, and the round concluded with each participant providing written responses to four questions on the Individual Interactivity, Perception, and Preference (IIPP) survey at their table. Following all three rounds of vignettes, at the conclusion of the forum, all participants completed the remaining 11 questions on the written IIPP Survey.

Instrument

We developed and utilized the IIPP survey to assess participant interactivity, perceptions, and preferences for this study. The IIPP survey consisted of 4 demographic questions followed by 15 questions answered on a Likert-type scale ranging from 1 to 5 on which 1 = strongly disagree and 5 = strongly agree. The first four questions concerned the student’s experience during each vignette and were answered three times, once after each vignette. The last 11 questions were answered once at the end of the forum and asked about overall experience at the forum and preferences regarding clickers. All questions on the IIPP survey were examined for content validity by a panel of six experts in survey design. Based on the feedback from the panel, no modifications to the original question set were deemed necessary. The complete IIPP survey may be found in Appendix A.
Data analysis

Responses from the IIPP survey instrument were recorded in the IBM SPSS Statistical Package for Windows (Version 20; SPSS) for statistical analysis. Data were screened for accuracy, missing data, univariate outliers, and normality (Tabachnick, & Fidell, 2007). Cases with missing data were excluded pairwise (Byrne, 2009). No other violations to the assumptions about the data were found.

Power analysis

Because we expected that group clicker usage would affect some aspects of the participants’ experiences, but not other aspects, we conducted a power analysis using G*Power software (Faul, Erdfelder, Buchner, & Lang, 2009) to determine the power of our study to distinguish between significant and non-significant findings and to minimize the chance for a Type II error if findings were non-significant.

Shapiro and Gordon (2012) reported effect sizes between $d = .38$ and $.77$, and $\eta^2$ between $.20$ and $.40$ for their analysis of clicker-assisted learning outcomes. Using their findings as an estimate of clicker effect sizes, we conducted a power analysis looking for $d = 0.35$ effect size. We set the probability value for significance of all tests at $\alpha = .05$, for an ANOVA repeated measures, within-between interaction, $\beta/\alpha$ ratio = 1, 2 groups, 3 repeated measures, no nonsphericity correction (later confirmed through examination of Mauchly’s Test of Sphericity on all F tests), with $n = 204$ participants: power was $1.00$.

To confirm the power of the study, we then decreased the desired effect size to $d = 0.10$ found that this very small effect size could be detected with a power of .92. This indicated that we had sufficient power with our sample size to detect the expected effect sizes at a power of $100\%$ and very small effect sizes at over $92\%$ power (Cohen, 1988), meaning that statistically non-significant results more likely reflected no difference than a Type II error. We later confirmed that the observed effect sizes for our repeated measures ANOVAs were between $\eta^2_p = .01$ to $.001$, (equivalent to $d = .20$ to .06, using the formula $d = 2*(\text{sqrt(eta}^2/ (1 - \text{eta}^2))$); Cohen (1988, p. 276) further supporting the contention that our non-significant findings accurately detected when there were no group differences between individual and group clicker users.

Results

Our first four research hypotheses dealt with participants’ experience in the small table discussion. The fifth and sixth hypotheses dealt with participants’ experience in the large forum. We began with hypotheses for which we expected to find no statistically significant differences and then tested hypotheses for which we thought that statistically significant differences would exist. Although two of the questions put to students used the phrase “…in the forum” in the question, both the timing of the responses immediately following the small table group discussions with clicker voting, and the explicit directions from the researcher asking students to reflect upon their small table experience, made it clear that students were to respond to their contemporary experience rather than their overall experience in the forum when answering these questions.

Hypothesis 1: Group clicker usage does not impair awareness of group opinions

Based on active learning theory (Smith, Wood, Krauter, & Knight, 2011), we expected that improvement in learning among students occurs because of the peer discussion component of the small table active learning experience; therefore, we hypothesized that group clicker usage would not affect awareness of group opinions because all members of the group would participate in the small table discussion, regardless of whether they had an individual or group clicker.

Participants were asked to respond to the statement “Using a clicker made me more aware of my peers’ opinions and attitudes” at three times during the seminar immediately following their small table group discussions. We hypothesized that participants in the group clicker condition would be equally aware of the opinions of other group members as individual clicker participants because all groups had to provide responses to the larger forum following
small table discussion. The average ratings of awareness were slightly higher for the group clicker condition, but the differences were not statistically significant and the effect size was small, $F(1,201) = .452, p = .637$, ns, $\eta^2_p = .004$. Hypothesis 1 was supported. Based on the results of power analysis, we concluded that this statistically non-significant result was not the Type II error and reflected our research expectation that participants’ interaction during their table discussions was not affected by having to share a single clicker in a group.

**Hypothesis 2: Group clicker usage inhibits individual participation**

Stowell and Nelson (2007) found that among participants in an introductory psychology class, clicker participation was higher than hand-raising participation. Martyn (2007) found that clicker groups were not statistically significantly different than discussion groups (although clicker groups had consistently higher scores). In keeping with these ambiguous findings, we were divided on this hypothesis with one author (KT) hypothesizing that participants in the group clicker condition would rate their level of participation lower than the individual clicker participants, and the other author (TD) hypothesizing that those differences would not exist.

Participants were asked to respond to the statement “Using a clicker helped me participate in the forum” at three times immediately following their small table group discussions. We discovered that there were no significant differences among or between group and individual participants. Using a multivariate repeated measures ANOVA, no significant differences were found for reports of individual participation between the individual clicker condition and group clicker condition, $F(1,201) = 1.29, p = .277$, ns, $\eta^2_p = .013$. Referring to our power analysis and the small effect size from this finding, we concluded that these results were not a Type II error but reflected the reality that level of participation in small group discussions does not differ as a function of sharing a single clicker within a group.

**Hypothesis 3: Group clicker usage inhibits individual contributions**

Studies using individual clickers consistently report an increase in positive experiences as a result of using the clickers. Trees and Jackson (2007) found that 1500 undergrad students over 7 large-enrollment courses individual clickers contributed to reports of greater feelings of involvement in the class. Duncan (2006) reported that when clickers were used within large lecture classes, students subjectively reported it made the class feel smaller. Additionally, high satisfaction with individual clickers has been reported because of the ability to provide immediate feedback during the learning process. DeBourgh (2008) found that undergraduate nursing students reported high levels of satisfaction with clicker technology and felt that clicker use provided them with individual feedback from their class instructor. We hypothesized that participants with individual clickers would feel that their contribution was more important than participants sharing a single clicker per group, because group clicker participants would lose their feeling of connection to the larger forum experience that was shared by participants with individual clickers.

Participants were asked to respond to the statement “Using a clicker made my input an important part of the forum” at three times immediately following their small table group discussions. Using a multivariate repeated measures ANOVA, no significant differences were found for perception of individual contribution between the individual clicker condition and group clicker condition, $F(1,201) = .174, p = .177$, ns, $\eta^2_p = .017$. Hypothesis 3 was not supported. This finding suggests that students’ perceptions of making an individual contribution are not reduced when they share a clicker with other group members.

**Hypothesis 4: Group clicker usage reduces representation of opinion**

Participants were asked to respond to the statement “My/the group clicker response reflected my personal opinion” at three times during the seminar immediately following their small table group discussions. We hypothesized that participants in the individual clicker condition would rate their clicker response as precisely reflecting their own opinions whereas participants in the group clicker condition would rate the group consensus response as being less reflective of their individual opinions. Surprisingly, not only were group clicker participants responses not significantly lower than individual clicker participants, both groups rated their participation above 4 on a scale of 1 to
5 and the marginal means were not different: individual \( M = 4.3, SD = .14 \), group \( M = 4.0, SD = .07 \), on a scale of 1 to 5.

The multivariate repeated measures ANOVA confirmed that ratings of opinion reflectiveness were not statistically significant for the individual clicker condition, \( F(1,201) = .048, p = .953, ns, \eta^2_p < .000 \). Participants’ perception of making an individual contribution were not reduced when participants shared a clicker with others in a group. This finding suggests that group-clicker participants worked towards consensus and tended to agree with the final clicker vote.

**Hypothesis 5: Group clicker usage reduces interactivity**

In a large introductory psychology class, Stowell and Nelson (2007) surveyed students’ perception of individual interactivity when using clickers compared to hand raising. Students reported significantly higher interactivity scores, both with other students and with the instructor, when using the clicker technology rather than responding to a question by raising their hand. Other studies (Bartsch & Murphy 2011; Bachman & Bachman, 2011) found similarly significant positive effects on student interactivity as a result of individual clicker use.

At the conclusion of the large forum, participants were asked to reflect upon their overall experiences in the large forum. These questions were different than the questions regarding experiences in the small table discussions. We constructed a variable called *interactivity* by combining the scores for five questions (Using the clickers increased my involvement in the forum. ... made my learning experience more enjoyable. ... helped me to stay interested in the forum. ... enhanced my interaction in the forum. ... helped me participate in the forum.) The variables measured in these five questions have been classified as components of the concept of *interactivity* described by Siau, Sheng, and Nah (2006). We conducted a reliability analysis on the five items and achieved a Cronbach’s Alpha of .93, with inter-item correlations ranging from \( r = .635 \) to \( r = .829 \).

We hypothesized that the individual clicker users would report greater interactivity in the large forum. We conducted an independent measures \( t \)-test, using bootstrapped confidence intervals, to look for group differences. Levene’s Test for Equality of Variances was not significant (\( F = .985, p = .322, ns \)) indication that the assumption of homogeneity of variance had not been violated. The interactivity variable did not differ significantly between the individual and group clicker conditions, \( t(202) = 1.87, p = .063, ns \). \( BCI = -.016, .602 \). Student reports of interactivity in the large forum did not differ regardless of whether students had individual clickers or shared a single clicker with a group. This finding supports the primacy of active learning pedagogy over simple clicker usage in improvement of student experiences in other studies of clicker benefits.

**Hypothesis 6: Participants prefer individual clickers**

Because of the positive perception of clickers reported in multiple publications (Hoffman & Goodwin, 2006; Martyn, 2007; Kenwright, 2009; Kay & Lesage, 2009), we theorized that all participants would want their own clicker. This assumption implied a hypothesis that the group participants would not differ statistically significantly from individual participants, in that both would report equally high preference for having their own clicker. Instead, we found that having had the group experience, group participants were statistically significantly lower in their desire for individual clickers, \( t(202) = 4.061, p < .000, BCI = .461, 1.332 \).

As expected, individual clicker participants rated their desire for individual clickers quite high, \( M = 4.10, SD = .995 \); group clicker participants, surprisingly, rated their desire for individual clickers significantly lower, \( M = 3.21, SD = 1.29 \). These findings imply that while students may say they prefer having individual clickers, once they experience using a single clicker to reflect group consensus, their preference for having an individual clicker will decrease. Hypothesis 6 was not supported.
Conclusion

Multiple studies about clicker effectiveness have concluded with a speculation that perhaps the results were not due to the clickers so much as to the active learning environment made possible by clickers (Martyn, 2007; DeBourgh, 2008; Mayer et al., 2009; & Armbruster, et al., 2009). Our findings support the contention that clicker technology is secondary to active learning pedagogy in the classroom experience. When clickers are used as part of an active learning activity, asking students to reach consensus in small groups is an effective way of using clickers to achieve the positive outcomes associated with active learning. This is good news to instructors and school systems who have a limited budget for clickers, or for instructors who want to encourage group collaboration. Our findings suggest that the outcome improvement noted by other researchers was due to the emphasis on active learning regardless of whether students responded individually or shared clickers within small groups.

The transition to an active learning environment is made possible by using clickers. Simply including hand-raising activities is not sufficient to achieve an interactive environment (Stowell & Nelson, 2007). However, despite the emphasis on individual clickers in previous research, our findings demonstrate that clickers need not be individualized to be beneficial. Interactivity in an active learning environment is facilitated by inclusion of clickers but is independent of the distributions of clickers in the classroom, at least when the active learning task requires reaching group consensus and the students share a single clicker in their group. Our findings generalize to small groups between four and eight students.

Suggestions for future research

In previous non-significant clicker group research, the authors noted that the direction of change was better for the clicker groups and speculated that the active learning environment was key to any positive outcomes (Armbruster et al., 2009; Tivener & Hetzler, 2015). Future iterations of clicker research should include a meta analysis to better understand the small effects for clicker usage and the larger effects for active learning pedagogy.

This study did not include an objective measure of student responses to assess learning outcomes. This study could have been strengthened by collecting data from individual and group tables on objective measures to determine whether the clicker responses were objectively correct. In conclusion, our findings in this study support the contention asking students to reach consensus in small groups using a shared clicker is an effective way of using clickers to achieve the positive outcomes associated with active learning. Clickers can provide an advantage to classrooms using active learning, but clickers alone will never be a replacement for engaging pedagogy.

References


Appendix A

INDIVIDUAL INTERACTIVITY, PERCEPTION, AND PREFERENCE (IIPP) SURVEY

Demographic Information Instructions: Please check the box that applies to you;

<table>
<thead>
<tr>
<th>Class standing:</th>
<th>Male</th>
<th>Female</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>□</td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>Undergraduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freshman</td>
<td>□</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sophomore</td>
<td></td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>Junior</td>
<td>□</td>
<td></td>
<td>□</td>
</tr>
<tr>
<td>Senior</td>
<td></td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Year</td>
<td>□</td>
<td></td>
<td>□</td>
</tr>
<tr>
<td>Second Year</td>
<td></td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>Third Year</td>
<td></td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td></td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>Faculty or Staff</td>
<td>□</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Declared major (or Department)

| □ Athletic Training | □ Biomedical Sciences | □ Communication Sciences and Disorders |
| □ Kinesiology       | □ Nursing             | □ Physical Therapy                      |
| □ Psychology        | □ Public Health       | □ Social Work                            |
| □                   | □ Other               |                                           |

Instructions: Circle your response

Using a clicker…

| Using the clickers increased my involvement in the forum. |
| Using the clickers made my learning experience more enjoyable. |
| Using the clickers helped me to stay interested in the forum. |
| Using the clickers enhanced my interaction in the forum. |
| Using the clickers helped me participate in the forum. |
| I found the clickers simple to use. |
| I was an active participant in the small table discussions. |
| I found the case studies to be very informative and relative to my area of study. |
| I would consider myself to have an outgoing personality (extrovert). |
| I would prefer to have my own clicker rather than use one for the entire small group. |
| I was very comfortable voicing my opinion in the small table discussion. |
| The key note speaker provided great information that was applicable to me and other healthcare workers. |
| The subject matter of the forum was very interesting to me. |

Question

<table>
<thead>
<tr>
<th>Question</th>
<th>1=Strongly Disagree</th>
<th>2=Moderately Disagree</th>
<th>3=Neutral</th>
<th>4=Moderately Agree</th>
<th>5=Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using the clickers increased my involvement in the forum.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Using the clickers made my learning experience more enjoyable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Using the clickers helped me to stay interested in the forum.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Using the clickers enhanced my interaction in the forum.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Using the clickers helped me participate in the forum.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I found the clickers simple to use.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I was an active participant in the small table discussions.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I found the case studies to be very informative and relative to my area of study.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I would consider myself to have an outgoing personality (extrovert).</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I would prefer to have my own clicker rather than use one for the entire small group.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I was very comfortable voicing my opinion in the small table discussion.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>The key note speaker provided great information that was applicable to me and other healthcare workers.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>The subject matter of the forum was very interesting to me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Measuring the Moral Reasoning Competencies of Service-Learning e-Tutors

Chih-Feng Chien¹, Ching-Jung Liao²*, Brent G. Walters³ and Ching-Yieh Lee⁴

¹Graduate School of Education, Chung Yuan Christian University, Taiwan // ²Department of Information Management, Chung Yuan Christian University, Taiwan // ³Department of Applied Linguistics and Language, Chung Yuan Christian University, Taiwan // ⁴Center for General Education, Chung Yuan Christian University, Taiwan // cchien@cycu.edu.tw // cjliao@cycu.edu.tw // brent@cycu.edu.tw // cylee@cycu.edu.tw

*Corresponding author

(Submitted April 21, 2015; Revised September 26, 2015; Accepted November 22, 2015)

ABSTRACT
As education has turned towards technology to provide academic support, the incidence of e-tutoring has grown due to decreasing educational budgets and as a potential remedy for the generational digital divide. However, many service-learning e-tutoring studies have focused on tutees’ academic achievement and tutors’ cognitive development rather than tutors’ moral development. The purpose of this study is to develop a literature-based instrument related to service-learning e-tutoring in order to explore college students’ development of moral reasoning competencies as a result of their service as online tutors for rural primary students. By reviewing the literature on service-learning and conducting interviews, content analysis is applied to identify the moral reasoning factors; then factor analysis is conducted to assess the reliability and validity of the moral reasoning questionnaire for e-tutoring. Five competencies, including moral character, problem solving, caring, empathy, and social interaction,—are recognized. Further study is needed to better ascertain the range of educational and personal gains in both tutor and tutee that result from service-learning e-tutoring.

Keywords
E-tutoring, Service-learning, Competencies, Instrument development

Introduction
Involving students in service-learning has been a long executed pedagogy on American campuses since the 19th century (Codispoti, 2004). By emphasizing student growth, service-learning is an instructional strategy to enrich students’ learning experience through service and to cause students to reflect on their personal experiences by enhancing their responsibility (Anderson, 1998). College students can provide service to enhance local elementary or secondary children’s learning achievement in the form of tutoring, (Carter Andrews, 2009), in which their own prosocial, moral, cognitive and personal development also increase.

Service-learning student teachers often help underperforming students in rural areas (Stachowski & Mahan, 1998). Children who lag behind on academic achievement are usually remedied by individual tutoring from parents or school teachers. For example, Flynn, Marquis, Paquet, Peeke, and Aubry (2012) conducted a study discovering that individual direct-instruction tutoring by parents improved children’s fundamental academic achievement. For students with a lack of resources, teachers and peers are usually the only ones available to play the role of tutor when economically stressed parents are forced to work long hours. Thus, service-learning tutoring may help to fill this gap.

McMann (1994) refers to traditional tutors as face-to-face mentors and e-tutors as computer-mediated mentors. Because traditional tutoring is hindered with the issue of distance, inefficiency, and expenditure, e-tutoring has come about by providing communication and interaction between tutor and tutee through the Internet. Although the educational role difference between traditional tutoring and their online counterpart is minimal (McMann, 1994), there are four differences between them as identified by Berge (as cited in McPherson & Nunes, 2004). E-tutoring plays (1) a pedagogical and intellectual role to provide probes and questions for discussion, conversation, debates, and summary; (2) a social role to involve tutee in a friendly social environment for interactive learning; (3) a managerial role to organize the learning schedule for tutees; and (4) a technical role to introduce tutee to a competent online learning environment.

Many studies have investigated the academic growth, regulatory development, and cognitive changes in preservice teacher tutors after they engaged in tutoring. For example, Shastri’s (1999) quasi-experimental study found that preservice teachers who tutored as part of a service-learning teacher education group performed significantly better than a non-service-learning group on their academic assignments. Guichon (2009) identified three competencies of
regulatory development which are not present in face-to-face tutoring. The three regulatory competencies for managing synchronous online tutoring for Master’s students in language teaching courses are: (1) socio-affective regulation to develop skills required for a learning community, (2) pedagogical regulation to provide feedback to facilitate content learning, and (3) multimedia regulation to apply technology supporting online interaction. Although tutors’ academic, cognitive and regulatory development has been researched considerably, most tutoring studies investigate the effects of tutoring on preservice teachers. There is a scarcity of research specifically focusing on service-learning tutors who are not preservice teachers. Moreover, the moral reasoning competencies of service-learning tutors have not been extensively studied.

Therefore, to investigate the research on service-learning e-tutor’s moral reasoning, a reliable and valid instrument was planned to be designed. The theoretical literature review and inquiry interviews proceeded at the same time. Our research group conducted an initial interview with 15 tutors and recognized that tutors tended to improve competencies such as moral character, responsibility, self-sacrifice, empathy, and equity after serving as an e-tutor. Thus, the purpose of this study is to develop a student-perceived and literature-based instrument on e-tutoring for exploring college students’ moral reasoning competencies after serving as service-learning e-tutors. The development of an instrument was initiated to be in accord with the most recent literature on service-learning. Meanwhile, an open-ended qualitative interview was conducted for withdrawing the moral competencies gained by student tutors. Then a pilot study was executed for validating the designed quantitative survey. A member-checking interview was administered to confirm the generated moral competencies. Finally a formal experiment was executed with a concluding questionnaire.

Literature review

Service-learning tutoring and moral development

Service-learning and the development of those who provide service-learning are topics that have been thoroughly researched over the past 20 years (Conway, Amel, & Gerwien, 2009; Yorio & Ye, 2012). Yet a major weakness lies in most research on service-learning in the field’s lack of theoretical rigor in study design and imprecise descriptive language (Yorio & Ye, 2012). This means that most studies on service-learning are difficult to replicate and the findings are likewise hard to generalize and compare. For example, the service-learning experience can be quite diverse in terms of the activity students undertake, the voluntary or involuntary nature of the service-learning, and the degree to which the service-learning is linked to the curriculum (Conway et al., 2009). With regards to the outcomes of service-learning on those providing the service, the framework used to capture the development change in the service providers has been roughly split between cognitive or academic development and “prosocial” development where “prosocial” can be anything from service providers’ civic literacy to moral development to time spent watching TV (Conway et al., 2009; Teo & Lim, 2009). The broad spectrum of “prosocial” outcomes and overall lack of guiding theory within service-learning research has meant that it is extremely difficult to generalize the findings on service providers’ moral development across studies and even when moral development is at the focus of the study (Bernacki & Jaeger, 2008). Because of these problems, this study attempts to focus on the outcome of volunteer service-learning tutors, specifically in the service of e-tutoring, and the service providers’ moral development in light of the standard models of moral development.

Moral development theories

The classically documented moral development theory initiated from Piaget (1932) and was later expanded into stages of moral reasoning by Kohlberg (1981). Subsequently Gilligan (1995), Kohlberg’s student, distinguished the difference between a feminine ethics of care and a feminist ethics of care in terms of moral reasoning. The feminine ethics of care is the personal obligations of selflessness and self-sacrifice, as well as an interpersonal relationship of self-development. The feminist ethics of care began with the connection among humans and then developed into resistant voices. McLeod-Sordjan (2014) concluded that, according to Gilligan’s criticism, Kohlberg's work is initialized from a male viewpoint and overlooks the female viewpoints of caring. Drawing heavily from Kohlberg's theory of moral development, Gilligan (1982) put forward two modes of ethical choice: (1) a responsibility mode focused on caring and relationship based on human needs; and (2) a rights mode focused on justice, equality and
individual rights based on moral principles. Gilligan (1982) described that care expresses interpersonal senses, such as empathy and connectedness. Gilligan (1982) defined that care is a spontaneously social relation.

Gilligan’s (1982) work was based on interviews of women which asked about their consideration of abortion from the viewpoint of survival, caring, and interdependence. Gilligan accentuated the perspectives of connective relationship, self-sacrifice, and caring. From this, empathy, moral character, and social interaction all are important for our study. Although many studies have addressed that Gilligan criticizes Kohlberg moral development theory in light of sexual discrimination, Jorgensen (2006) pointed out that Gilligan supports Kohlberg’s stage theory and her theory of moral development is a part of an expansion to Kohlberg’s theory. Rest, Narvaez, Bebeau and Thoma (1999) even concluded that the viability of Gilligan’s theory is to modify Kohlberg’s theory. Conrad and Hedin (1991) indicated that the service-learning participants’ moral reasoning can be typically measured by Kohlbergian moral theory despite Kohlberg’s approach’s neglect of cognitive issues such as altruism or responsibility. Thus our study initially premises that service-learning should include the moral reasoning factors covered in Gilligan and Kohlberg’s theory, those are responsibility, justice, empathy, caring, and social interaction.

E-tutoring studies

While Kohlberg and Gilligan established the foundations of moral development as the origin of service-learning, community tutoring is considered as one of the important parts of community engagement pedagogies. To overcome the obstacles of distance, inefficiency, and expenditure, online tutoring is the trend suitable for the students who grow up with ubiquitous internet accessibility. Flowers (2007) has defined e-tutoring as an individual learning support provided through the Internet. E-tutoring ranges from synchronous individual support to asynchronous message exchange between tutors and tutees (Childs, 2003).

Cavanaugh, Gillan, Kromrey, Hess and Blomeyer (2004) conducted a meta-analysis and found that students in online learning performed as well or better than their counterparts. Johnson and Bratt (2009) concluded Cavanaugh et al.’s study and emphasized that online technological relationship development (e.g., online feedback, email) is the important element for online learning between tutors and tutees. Johnson and Bratt’s conclusion responds to the factors Gilligan recognizes for moral development and service-learning, namely, interactive relationship.

Kopp, Matteucci and Tomasetto (2012) have posited that online collaboration fosters positive interaction between tutors and tutees. They investigated 76 tutors’ growth from engaging in e-tutoring and find that experienced e-tutors made more specific cognitive activities for online collaboration. By using cluster analysis, experienced e-tutors reported they intervened significantly more often than inexperienced ones regarding cognitive activities, social activities, and meta-cognitive activities. Their survey included five factors of e-tutoring—knowledge exchange, online discussion, argumentation, collaborative problem solving, and considering different perspectives (Kopp et al., 2012). Our study conducted inquiry interviews and discovered that the most significant concern for the e-tutors was problem solving. These results might be concluded by noting that the one-to-one online tutoring is more one-way teaching where collaborative learning rarely occurred during the service-learning tutoring.

Methodology

E-tutoring process

College students, who had been trained through a series of training workshops about online teaching, chose their own teaching materials and media based on the individual characteristics of their assigned tutees. A majority of students took advantage of PowerPoint with the integration of interactive whiteboards, videos, and animations. As a part of the project “Online Tutoring for After School’s Learning” funded by the Ministry of Education in Taiwan (see Figure 1), a number of Taiwanese universities have joined the project since 2006 to help bridge the digital divide and learning gap (ChanLin et al., 2012). The elementary- or secondary-school students (tutees) benefited from the advanced communication technology, and college students (tutors) benefitted from the experience of service-learning e-tutoring.
Participants

As a part of government-funded project, college students were recruited for an e-tutoring project to provide remediation for elementary-school students in remote areas. Participants were recruited to provide e-tutoring every semester, two times a week, one hour a time. The research team started collecting data from the spring semester to fall semester in 2014. Ninety tutors were recruited in the spring and as some left at the end of the spring semester because of personal reasons (e.g., graduation, time demands), 20 more were recruited in the following fall semester. Recruited from across all the departments of a private, comprehensive university in northern Taiwan, a majority of the tutors were undergraduate students and a minority of them were graduate students. Our research initiated inquiry interviews with 15 undergraduate students (7 males and 8 females) at the end of spring semester. As a questionnaire was developed based on the results of the interviews, 90 surveys were collected in the following couple of weeks. In the beginning of the fall semester, 10 tutors who remained in the project were interviewed again for member checking (Merriam, 2002), bringing ambiguous data interpretations to the participants for a clarification. The original questionnaire was modified based on the result of factor analysis, literature review, and member checking. Towards the end of fall semester, 77 surveys were collected for instrument reliability and validity.

E-tutoring system

The e-tutoring system used in this study is similar to the SLEIC² proposed by Liao, Yang, and Yang (2012), which integrates the concept of social learning 2.0, social software, cloud computing technology, Rich Internet Applications (RIAs) technology, and open APIs, to help the learners to use social networks while learning simultaneously. The framework of this e-tutoring system is shown in Figure 2. In this framework, the platform architecture contains the client user, external learning resources mashup, FOAF, and database access, which provides a centralized portal for users to operate collaborative real time social interaction mechanisms, such as video conferencing, electronic whiteboards, chat rooms, and Web 2.0 concepts including social networking services such as blog, YouTube, Facebook, and LINE.
Research instrument

The draft of interview questions was initiated by reviewing the relevant literature about service-learning, including the origin of service-learning, moral reasoning, and relevant e-tutoring literature. According to Kohlberg’s (1981) and Gilligan’s (1982; 1995) theories of moral development, our initial literature review distilled a number of factors relating to service-learning, such as responsibility, justice, empathy, caring, social interaction, and moral character. Other literature relating to e-tutoring extracted problem solving and interpersonal skills as the competencies e-tutors have gained (Kopp et al., 2012; McPherson, & Nunes, 2004).

An inquiry interview with 15 tutors was administered. The interview involved open-ended questions based on the literature reviewed as follows:

- What did you learn from service-learning e-tutoring?
- Did you bear any burden for providing e-tutoring service? If yes, what are they?
- Do you think it is necessary to provide service-learning e-tutoring because of social justice?
- Are there any moral lessons learned during e-tutoring?
- What skills or competencies did you learn from e-tutoring?

Based on the interview results and literature, the research team developed a questionnaire with 62 items on a five-point Likert Scale. The questionnaire was developed with the six most common factors, including moral character, caring, social interaction, responsibility, problem solving, and empathy. We decided to remove the factor of social justice because it was not revealed from the interview results. The following steps are the approach of member checking to confirm the initial data results including inquiry interviews and surveys. The half-structural member checking interview was guided according to the questions about probing factors, such as “what did you learn about problem solving skills during the service-learning e-tutoring?” As the questionnaire was modified according to the results of member checking interviews and factor analysis, 32 survey items were identified for further confirmation of reliability and validity.

Qualitative content analysis

Sarantakos (1993) describes inquiry interviewing as a data collection to probe initial information from interviewees. Before administering inquiry interviews, researchers read through the literature to locate the concepts of service-learning and e-tutoring. As a number of factors were identified from Kohlberg’s (1981) and Gilligan’s (1982; 1995)
moral development theory, as well as Kopp et al. (2012), Cavanaugh et al. (2004) and others’ studies, inquiry interviews were executed to confirm the probing factors. The interviews with 15 undergraduate students were administered by two authors towards the end of first semester. The interviews’ audio was recorded and verbatim transcribed by research assistants. The first two authors read through the transcripts and applied content analysis to allocate the interaction (Downe-Wamboldt, 1992). The transcript was selected for being abstracted and coded (Weber, 1990). Modifying the content analysis approaches of Granheim and Lundman (2004), and Mazaheri et al. (2013), textual units were analyzed and referred to meaning units, condensed units, sub-themes, and themes. Granheim and Lundman (2004) address that a meaning unit is the “constellation of words or statements that [are] related to the same central meaning” (p. 106). In other words, a phrase, sentence, and paragraph of interactions containing similar aspects among interviewees can be considered as meaning units. To shorten the text for data reduction (DeWalt & DeWalt, 2011), we executed the process of selecting, simplifying, and reducing a large amount of text for condensed meaning units. Similar condensed meaning units can be clustered based on theoretical assumptions from the literature, which are the codes or labels. A category contains a group of codes sharing a commonality (Krippendorff, 1980). By linking different categories, themes are developed as a concept in a qualitative approach.

Table 1. Examples of content analysis of interviews with 15 e-tutoring participants

<table>
<thead>
<tr>
<th>Meaning unit</th>
<th>Condensed meaning unit</th>
<th>Sub-theme</th>
<th>Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>When encountering problems on lesson preparation, I usually looked for solutions on the Internet. Then I went to ask one of my friends majoring in Chinese about how to teach Chinese online. When facing interpersonal issues, I usually asked professors how to break the ice. Some techniques I used were tongue twisters or watching a short video clip.</td>
<td>Looking for solutions on the Internet first, and then discussing with friends and professors.</td>
<td>Searching for solutions online and discuss with friends</td>
<td>Problem solving</td>
</tr>
<tr>
<td>I had a senior who majored in mathematics and did an internship in an elementary school. I usually asked him how to make tutees understand mathematical concepts. Other than that, I went online for more information about teaching mathematics.</td>
<td>Consulting a senior in a similar field, and then went online for further information.</td>
<td>Asking help from an experienced tutor</td>
<td></td>
</tr>
</tbody>
</table>

The interview transcripts were read through several times by the two authors separately. Then the results of content analysis were compared and discussed for agreement (Mazaheri et al., 2013). Table 1 is an example of content analysis using the approach of Granheim and Lundman (2004). Finally six themes related to e-tutoring were identified: moral character, problem solving, caring, social interaction, empathy, and responsibility. The theme of justice was left out because all participants, who volunteered for e-tutoring, never thought about social justice. Some stated that the government provided rural schools many resources, while others thought that they also learned a lot of other things besides social justice by being an e-tutor.

Table 2. Distribution of e-tutoring themes.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moral character</td>
<td>The moral development such as helping others, discernment, empathy, positive attitudes, and caring in e-tutoring.</td>
<td>66 (32.8%)</td>
</tr>
<tr>
<td>Problem solving</td>
<td>Working through detailed thinking and processing to reach a solution, such as searching for resources for teaching preparation in e-tutoring.</td>
<td>33 (16.4%)</td>
</tr>
<tr>
<td>Caring</td>
<td>The protective aspect of emotional expression among people with close relationship, such as enthusiasm, devotion, and positive emotions toward e-tutoring.</td>
<td>30 (14.9%)</td>
</tr>
<tr>
<td>Empathy</td>
<td>A caring orientation of a person via the feeling of compassion, such as showing consideration to tutee’s daily life in e-tutoring.</td>
<td>33 (16.4%)</td>
</tr>
<tr>
<td>Social Interaction</td>
<td>The communication skills, such as articulating, expressing, conveying their ideas, as well taking an initiative to talk with others in e-tutoring.</td>
<td>39 (19.4%)</td>
</tr>
</tbody>
</table>
The frequency of the themes was calculated by the first two authors to examine how tutors perceived their competencies after providing service-learning e-tutoring (see Table 2). The calculating disagreements between the two authors achieved a high inter-judge reliability value of 0.942. After a further discussion, a consensus was reached on the messages that were coded differently by the authors.

**Item analysis**

The initial questionnaire was created by analyzing the results of inquiry interviews and literature on service-learning. A pilot study was executed to assess 62 items created and discussed by two authors. With the limiting data of participants collected, a total of 90 tutors who had been providing service-learning e-tutoring for at least one semester participated in the pilot study. The purposes of the exploratory factor analysis (EFA) were twofold: (1) to verify the questionnaire validity, and (2) to access to the reliabilities of the questionnaire and each extracted factor (Field, 2009). After removing 30 items with low factor loading, a total of five factors (eigenvalue > 1) were extracted with 32 items remaining and this cumulatively explained 69.98% of total variance. Based on the theoretical framework of moral development (Kohlberg, 1981; Gilligan, 1982) and e-tutoring (Kopp et al., 2012), the first factor—moral character—includes 8 items ($\alpha = 0.907$), the second factor—problem solving—includes 8 items ($\alpha = 0.923$), the third factor—caring—includes 6 items ($\alpha = 0.904$), the fourth factor—empathy—includes 5 items ($\alpha = 0.920$), and the fifth factor—social interaction—includes 5 items ($\alpha = 0.895$). The Cronbach’s Alpha reliability of the questionnaire is 0.947 (see Table 3).

**Table 3. Factor analysis of e-tutoring competencies**

<table>
<thead>
<tr>
<th>Item</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q43</td>
<td>.764</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Moral character</td>
</tr>
<tr>
<td>Q46</td>
<td>.731</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.907</td>
</tr>
<tr>
<td>Q42</td>
<td>.729</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q44</td>
<td>.695</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q41</td>
<td>.676</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q38</td>
<td>.660</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q39</td>
<td>.634</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q45</td>
<td>.628</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q57</td>
<td>.834</td>
<td>.788</td>
<td></td>
<td></td>
<td></td>
<td>Problem solving</td>
</tr>
<tr>
<td>Q58</td>
<td>.786</td>
<td>.786</td>
<td></td>
<td></td>
<td></td>
<td>0.923</td>
</tr>
<tr>
<td>Q56</td>
<td>.741</td>
<td>.735</td>
<td>.684</td>
<td>.676</td>
<td>.578</td>
<td></td>
</tr>
<tr>
<td>Q55</td>
<td>.676</td>
<td>.578</td>
<td>.578</td>
<td>.578</td>
<td>.578</td>
<td></td>
</tr>
<tr>
<td>Q24</td>
<td>.862</td>
<td>.856</td>
<td>.805</td>
<td>.801</td>
<td>.739</td>
<td>Caring</td>
</tr>
<tr>
<td>Q26</td>
<td>.856</td>
<td>.805</td>
<td>.801</td>
<td>.801</td>
<td>.739</td>
<td></td>
</tr>
<tr>
<td>Q27</td>
<td>.805</td>
<td>.805</td>
<td>.805</td>
<td>.805</td>
<td>.739</td>
<td></td>
</tr>
<tr>
<td>Q22</td>
<td>.801</td>
<td>.801</td>
<td>.801</td>
<td>.801</td>
<td>.739</td>
<td></td>
</tr>
<tr>
<td>Q25</td>
<td>.739</td>
<td>.739</td>
<td>.739</td>
<td>.739</td>
<td>.739</td>
<td></td>
</tr>
<tr>
<td>Q23</td>
<td>.542</td>
<td>.542</td>
<td>.542</td>
<td>.542</td>
<td>.542</td>
<td></td>
</tr>
<tr>
<td>Q34</td>
<td>.802</td>
<td>.789</td>
<td>.742</td>
<td>.726</td>
<td>.659</td>
<td>Empathy</td>
</tr>
<tr>
<td>Q33</td>
<td>.789</td>
<td>.789</td>
<td>.789</td>
<td>.789</td>
<td>.789</td>
<td></td>
</tr>
<tr>
<td>Q35</td>
<td>.742</td>
<td>.742</td>
<td>.742</td>
<td>.742</td>
<td>.742</td>
<td></td>
</tr>
<tr>
<td>Q36</td>
<td>.726</td>
<td>.726</td>
<td>.726</td>
<td>.726</td>
<td>.726</td>
<td></td>
</tr>
<tr>
<td>Q37</td>
<td>.659</td>
<td>.659</td>
<td>.659</td>
<td>.659</td>
<td>.659</td>
<td></td>
</tr>
<tr>
<td>Q14</td>
<td>.888</td>
<td>.855</td>
<td>.811</td>
<td>.772</td>
<td>.664</td>
<td>Social interaction</td>
</tr>
<tr>
<td>Q13</td>
<td>.855</td>
<td>.855</td>
<td>.855</td>
<td>.855</td>
<td>.855</td>
<td></td>
</tr>
<tr>
<td>Q11</td>
<td>.811</td>
<td>.811</td>
<td>.811</td>
<td>.811</td>
<td>.811</td>
<td></td>
</tr>
<tr>
<td>Q12</td>
<td>.772</td>
<td>.772</td>
<td>.772</td>
<td>.772</td>
<td>.772</td>
<td></td>
</tr>
<tr>
<td>Q15</td>
<td>.664</td>
<td>.664</td>
<td>.664</td>
<td>.664</td>
<td>.664</td>
<td></td>
</tr>
</tbody>
</table>
Reliability of the instrument

In the beginning of the fall semester of 2014, fifteen students out of 90 were recruited for a focus-group interview to re-confirm the results of the inquiry interviews and surveys for member checking process. Five factors (moral character, problem solving, caring, empathy, and social interaction) were confirmed strongly by these students. Thus, we decided to administer this verifying survey with 32 items for the tutors in the fall semester of 2014. Seventy-eight tutors were recruited for this experiment. After a semester of e-tutoring, the Cronbach’s Alpha was computed to assess the questionnaire’s reliability. Results revealed that the questionnaire had relatively high reliability ($\alpha = 0.976$) in which the five factors moral character ($\alpha = 0.973$), problem solving ($\alpha = 0.945$), caring ($\alpha = 0.937$), empathy ($\alpha = 0.927$), and social interaction ($\alpha = 0.891$), all had high reliability respectively. Please see Table 4 for the items and reliabilities of the final questionnaire.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Items</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moral character</td>
<td>Q21. I agree with the importance of character for life.</td>
<td>.973</td>
</tr>
<tr>
<td></td>
<td>Q24. I agree with the importance of holding a positive attitude towards life.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q20. I agree that character education is very important to my present and future.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q22. I agree that moral character has a positive meaning to my life.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q19. By understanding the meaning of moral character, I have a higher expectation for my future and potential.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q17. A key to success is not a person’s intelligence but his/her moral character.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q18. The role models of moral character have their own ideal goals.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q23. A person with good character often inspires and affects my life.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q28. I usually verify and validate the causes of the problems that have occurred.</td>
<td>.960</td>
</tr>
<tr>
<td>Problem solving</td>
<td>Q29. I look at a problem from different perspectives in order to obtain deeper insights.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q27. I usually think carefully about who, what, when, where, and what on the causes of the problems that have occurred.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q31. I apply what I have learned to solve daily problems.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q30. I think about the problem’s solutions from different perspectives.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q33. I try to understand the nature of a problem by analyzing different aspects.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q35. I plan procedures of problem solving, and execute them systematically.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q32. I try to figure out the ways to solve problems from research literature or from successful people.</td>
<td></td>
</tr>
<tr>
<td>Caring</td>
<td>Q8. I love e-tutoring and want to participate more.</td>
<td>.937</td>
</tr>
<tr>
<td></td>
<td>Q10. I do not want e-tutoring to disappear from my life.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q11. I will be very happy participating in e-tutoring if there is an opportunity.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q6. I feel that the e-tutoring activity is very appealing to me.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q9. Despite my bad mood, I want to participate in the e-tutoring activity.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q7. I have a strongly positive emotion towards e-tutoring.</td>
<td></td>
</tr>
<tr>
<td>Empathy</td>
<td>Q13. I usually concentrate on listening to people during a conversation.</td>
<td>.908</td>
</tr>
<tr>
<td></td>
<td>Q12. I usually respect other’s feelings.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q15. I usually face people in a positive attitude when getting along with them.</td>
<td></td>
</tr>
<tr>
<td>Social interaction</td>
<td>Q4. I am able to articulate what I mean in a group.</td>
<td>.891</td>
</tr>
<tr>
<td></td>
<td>Q3. I can convey my own voices in a group interaction.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q1. I usually can express my thoughts clearly in a group.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q2. I can talk in a group discussion as naturally as I do with tutees.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q5. I often take an initiative to talk with others.</td>
<td></td>
</tr>
</tbody>
</table>

Discussion and conclusion

Kolberg’s (1981) and Gilligan’s (1982) moral reasoning theories are the theoretical foundation for service-learning. As tutoring has been developing a digital approach, some factors might be applicable for e-tutoring and some might not. The design of this mixed-method study attempts to triangulate the literature bases, the extraction of the key
points from qualitative interviews, and the assessment of the quantitative instrument. The development of the instrument used in this study needs continuous verification and experiment for additional service-learning e-tutoring moral reasoning to be discovered. Furthermore, the research design allows further reflective thinking and continuous timely modification among researchers and tutors when administering service-learning e-tutoring.

In the perspectives of college students’ moral reasoning competencies, we found that the moral foundations of service-learning, based on Kolberg’s (1981) and Gilligan’s (1982) theories of moral development, should include responsibility, justice, empathy, caring, social interaction, and moral character. Other e-tutoring related literature (Kopp et al., 2012) has been extracted into one factor—problem solving, which is also implemented in moral reasoning (Arbuthnot & Gordon, 1986). The content analysis of the initial interviews indicate that the effect of e-tutoring towards tutors is more similar to the female orientation towards moral development—the caring orientation—that Gilligan has advocated (1982; 1995). Contrarily, the factor of justice orientation Kohlberg (1981) promotes was not clearly manifested in the inquiry interview portion of this research. During the interviews, tutors pointed out that increasingly resources are pouring into rural schools so there does not seem a digital divide between rural and urban areas. We presumed this might be the reason that a justice orientation is hard to extract. Additionally, in our quantitative data collection of the pilot study, the factor loading of responsibility is lower than 0.5. This result might ascribe that in the training for e-tutoring, tutors should consider this activity as a duty or commitment, not merely a volunteer task. The majority of tutors never skipped tutoring except for sick leave.

Thus, our member checking and formal experiment concludes with five e-tutoring factors with relative high reliabilities (larger than 0.89) in our 32 item survey. The following is a detailed discussion for each of the factors.

Moral character

In the inquiry interviews, college students confirmed that moral characters are like helping others, discernment, empathy, positive attitudes, caring, listening and understanding. They considered that it is very important to learn about character in a person’s life. They also believed that a person’s moral characters are more important than his or her intelligence and that the key to success is having good character. Lai (a female interviewee) expressed that:

“Instead of saying I am sacrificing my time, I would rather say that I am enjoying my teaching. When the tutee was able to demonstrate the new concept, he felt happy, so did I. When he completed the problems I assigned correctly, I also sensed the self-fulfilment. I don’t think I am sacrificing my time and efforts.”

In ethics research, Hanson and Moore (2012) have hypothesized a conceptual model of an iceberg on moral development in the case of business students’ moral influence. They point out that the sense of moral consciousness and moral righteousness are below the surface—not managed by universities. On the contrast, the characters like caring, trust, and giving are above the surface—managed via school curricula and policies. Their study demonstrates that college students were not aware of their moral consciousness but did experience moral characters via school curricula. The service-learning e-tutoring is exactly a part of school curricula in higher education.

Problem solving

Kopp et al. (2012) concluded that the capability of problem solving is an important factor for collaborative members in an e-tutoring experience. In our inquiry interviews, students perceived a relatively high amount of problem-solving learning. This result was verified in quantitative data of the pilot study and the formal experiment. The questionnaire investigates college students’ ability of problem solving, such as validating the causes of an issue, seeing a problem from different perspectives, applying learned skills, planning procedures to solve problems, and brainstorming problem-solving approaches. College students implemented problem-solving skills in searching for resources for teaching preparation. Lee (a female interviewee) described that “when running into a problem, I would not get stuck and stop. I usually asked the experienced ones and went online to figure out the solutions.”
Caring

Caring is an important factor extracted from Gilligan and Attanucci’s (1988) morality of caring to moral education. Gilligan and Attanucci (1988) emphasized that caring is the approach women mainly apply to demonstrate morality and emotions, which is quite different from men’s moral expression—justice. Caring is the protective aspect of emotional expression among people with close relationships. This is similar to the factor of empathy in our study, discovered when a student is involved in e-tutoring. Gilligan (1982) accentuated that caring is similar to responsibility and connection in the third stage (post-conventional stage) of caring development she defines.

In e-tutoring, we consider caring as the enthusiasm tutors were eager to devote to this activity. Drawing from our quantitative data, the caring includes the enthusiasm to participate in e-tutoring, the joy felt during e-tutoring, and strong positive emotions toward e-tutoring. Liu (a female interviewee) expressed that:

“I am doing what I have to do. In the beginning, I felt it was difficult to get along with her (tutee), but I tried to care about her and talked something other than school works. She then started to share me the things in her mind. I felt that I am not only helping her, but also understanding myself more. An additional benefit is that I learned to share secrets with others, which I usually would not do.”

Thus, when tutors released their caring and enthusiasm to tutees, the reciprocal trust and relationship are developed which made tutor’s caring competency grow. According to the phone interview with the lead teachers of the tutees, some tutees felt tutors are like senior peers and less like teachers. The tutors’ involvement may have extended beyond the school work, which might have caused the tutees to feel as if the tutors were accompanying them along a process instead of simply focusing on helping them to reach a set goal.

Empathy

While our research defines caring as enthusiasm on e-tutoring, empathy is an expression of compassion by a person involved in an ethical situation (Glover, 2001). Empathy is an emotional expression in which the empathetic recipient is realized or sympathized (Davis, 1994). Our research found that empathy was present in e-tutoring when tutors not only lectured tutees on academic subjects, but also showed consideration to their daily life. In Gilligan’s (1982) theory of moral development, she highly accentuated the affective empathy that women pay more attention to. Our research also found this factor was a big part for service-learning e-tutoring in our study than was the justice orientation Kohlberg (1984) emphasized. Tang (a male interviewee) pointed out that:

“I served tutoring without thinking about social injustice. Instead, I felt tutees were taken care of heavily by the government. E-tutoring would never be conducted for urban kids. However, I feel tutees sometimes just need someone to listen, to respect, and to accompany them.”

Although the conventional wisdom holds that the tutees in our study lacked access to educational resources, the tutors believed that tutees needed more interpersonal support.

Social interaction

Social interaction has been long studied in peer tutoring, such as Vygotsky’s (1978) zone of proximal development. By receiving assistance from high-achieving peers or seniors, students’ potential and motivation are awakened. Social interaction during tutoring benefits tutors’ communication skills and tutees anxiety reduction and content transferring ability (Topping, 1996).

In our research, tutors affirmed that they have learned about articulating, expressing, conveying their ideas, as well taking an initiative for others. The results verify many studies about e-tutoring positively impacting on social interaction. Kopp et al. (2012) pointed out that the experiences of social interaction are useful for interpersonal influence that cognitively processes to prevent conflicts. One of our male interviewees Wang describes that:
“Before participating in e-tutoring, I was less confident to interact with children because of the experience with my younger sisters and brothers. I did not have patience to teach them. After the e-tutoring experience, I have learned to change my language to children’s level, and maybe told some of my stories to be closer to their daily life. I feel this approach was helpful, and I am not afraid of interacting with children anymore.”

The competency of social interaction, as Kopp et al. (2012) stressed, stands out as an important factor in e-tutoring as quantitative and qualitative data has confirmed. According to the phone interview with the lead teachers of the tutees, some tutees started opening their mind in terms of peer friendship after receiving e-tutoring.

Limitations

There are several limitations in this research which must be considered for future research execution. First, although tutoring has been long studied, many focus on cognitive and regulatory development particularly for traditional tutoring or peer tutoring. This study specifically emphasizes on service-learning tutoring in the perspective of moral reasoning due to the scarcity of relevant research. Thus, this study only applies to a limited area of e-tutoring specifically for service-learning. Second, for the quantitative data collection, only 90 participants responded for the pilot study and 77 participants were involved with the formal experiment. Although the reliability and validity are modest, the quantities for both data collections are fairly small. The purpose of this study is to pioneer a study focusing on tutor’s moral reasoning competence from the experience of e-tutoring. Our study will continue to proceed and the instrument will continue to be modified for appropriately measuring these factors we identified in the tutors. Third, although the tutors’ demographic background is diverse in terms of gender and major, the fact that they are from the same university may limit the generalizability of the research results. The diversity of universities (e.g., vocational and general universities, public and private universities, or universities in different regions) may produce different research results. To generalize the findings, more tutors from different universities should be recruited.

Conclusion

This study is significant in several aspects: (1) exploring service-learning e-tutoring experiences; (2) initiating an exploration on service-learning e-tutors’ moral development; and (3) developing a reliable and valid instrument for service-learning e-tutors’ moral reasoning. Although a number of limitations restrict the instrument’s reliability and validity, the literature review on service-learning and qualitative data collection establish initial research foundations of e-tutors’ moral development. As the Taiwan government has considerably funded e-tutoring in the past, the study suggests that governments should focus on the effects beyond tutors’ and tutees’ learning achievement. The competencies of moral reasoning, including moral characters, interpersonal relationships, problem solving, empathy, and caring, should be discussed and investigated among tutors and tutees. Further investigation, perhaps longitudinal in nature, of service-learning e-tutoring is needed in order to identify and measure the educational as well as personal gains made by both the tutor and tutee.

References


Inquiry Based-Computational Experiment, Acquisition of Threshold Concepts and Argumentation in Science and Mathematics Education

Sarantos Psycharis
Department of Education, ASPETE, Athens, Greece // spsytcharis@gmail.com

(Submitted May 1, 2015; Revised September 6, 2015; Accepted November 13, 2015)

ABSTRACT
Computational experiment approach considers models as the fundamental instructional units of Inquiry Based Science and Mathematics Education (IBSE) and STEM Education, where the model take the place of the “classical” experimental set-up and simulation replaces the experiment. Argumentation in IBSE and STEM education is related to the engagement of students/learners in a process where they make claims and use models and data to support their conjectures and justify or disprove their ideas. Threshold Concepts (TCs) are of particular interest to IBSE and STEM education, representing a “transformed way of understanding, or interpreting or viewing something without which the learner cannot progress.” The purpose of this study is to explore the effects of IBSE teaching approach on University students’: (a) argumentation; (b) involvement in the use of modelling indicators; and (c) acquisition of certain threshold concepts in Physics and Mathematics. 79 pre-service engineering school university students participated in this research and results indicate that the computational experiment can help students’ acquisition of threshold concepts and improve their level of argumentation as well as the use of modelling indicators.

Keywords
STEM Education, Computational experiment, Argumentation, Threshold concepts

Introduction

Inquiry based learning and modelling in science and mathematics education

Inquiry based learning has officially been promoted as a pedagogy for improving STEM learning in many countries (Bybee et al., 2008) and can be defined as “the deliberate process of diagnosing problems, planning experiments, distinguishing alternatives, setting up investigations, researching conjectures, sharing information, constructing models, forming coherent arguments, collecting and analyzing data” (Bell et al., 2004).

Significant parts of scientific research are carried out on models rather than on the real phenomena because by studying a model we can discover features of and ascertain facts about the system the model stands for. This cognitive function of models has been widely recognized in the literature, and some researchers even suggest that models give rise to a new form of reasoning, the so-called “model based reasoning” (Magnani & Nersessian, 2002) while modelling ability is also associated to model-based reasoning (Chittleborough & Treagust, 2007). It is well known that scientific theories are developed through a process of continuous elaboration and modification in which scientific models are developed and transformed to account for new phenomena that are uncovered. Similar processes are involved in students’ learning of STEM concepts when they develop conceptual models (e.g., Bell et al., 2010). In a similar fashion, inquiry based learning requires from students to make successive refinements to their mental models in order to transform them to conceptual models that align to scientific theories.

The Computational Experiment (CE)

Computational Science can be considered as a third independent scientific methodology (the other two are the theoretical science and the experimental science), has arisen over the last twenty years and shares characteristics with both theory and experiment, while it demands interdisciplinary skills in Science, Mathematics, Engineering and Computer Science. According to (Landau, Páez, & Bordeianu, 2008) Computational Science focuses on the form of a problem to solve, with the components that compose the solution separated according to the scientific problem-solving paradigm: (a) Problem (from science); (b) Modelling (Mathematical relations between selected entities and variables); (c) Simulation Method (time dependence of the state variables, discrete, continuous or stochastic processes like e.g., Monte Carlo simulation); (d) Development of the algorithm based on numerical analysis.
methods, (e) Implementation of the algorithm (using Java, Mathematica, Fortran etc.); and (f) Assessment and Visualization through exploration of the results and comparison with real data in authentic phenomena. In this framework, being able to transform a theory into an algorithm, requires significant theoretical insight, concrete Physical and Mathematical understanding as well as a mastery of the art of programming. In order to describe inquiry based learning as a search process, (Klahr & Dunbar, 1998) introduced two spaces, the hypothesis and the experimental spaces. (Psycharis, 2013) added one more space, the prediction space, in order to address the introduction of modelling and the comparison of data produced by the model with real data. In the “prediction space” the Computational Science methodology is implemented through the development of models of simulations that favor the so called “Computational Thinking (CT).” According to (Psycharis, 2013), the three spaces of the Computational Science methodology should include issues form (CT), namely: (a) logically organizing and analyzing data; (b) representation of data through abstractions such as models and simulations; and (c) algorithmic thinking.

In this context, the three spaces of the Computational Science methodology include:

- The hypotheses space, where the students, in cooperation with the teacher, decide, clarify and state the hypotheses of the problem to be studied, as well as the variables included in the problem and the possible relations between the variables.
- The experimental space, which includes the model and the method of simulation for the problems under study. In this space the learners are engaged in the scientific method writing models according to the variables they included and the interaction laws that govern the phenomenon. In this space students collect the data from their model and analyze them, while they try to connect them with the theory they have been taught.
- The prediction space, where the results, conclusions or solutions formulated in the experimental space, are checked through the analytical (Mathematical) solution as well as with data from the real world.

Bell et al. (2010), identified nine main science inquiry processes supported by different computer environments that could be used (IBSE) and STEM, namely: orienting and asking questions; generating hypotheses; planning; investigating; analyzing and interpreting; exploring and creating models; evaluating and concluding; communicating; predicting. The nine inquiry tools of (Bell et al., 2010) are closely related to the essential features of inquiry based learning (Asay & Orgill, 2010), namely: Question (the learner engages in scientifically oriented questions), Evidence (the learner gives priority to evidence), Analysis (the learner analyses evidence), Explain (the learner formulates explanations from evidence), Connect (the learner connects explanations to scientific knowledge and Communicate (the learner communicates and justifies explanations). Our learning and teaching approach integrates the inquiry based approach and (CE) through the relation of the (CE) spaces, namely the hypotheses space, the experimental space and the prediction space with the essential features of inquiry, the Computational Thinking(CT) features and the inquiry tools of (Bell et al., 2010).

In Table 1 below, we present the interrelation between the spaces of the (CE) and the features and the tools of inquiry.

<table>
<thead>
<tr>
<th>Spaces of the (CE) (Psycharis, 2013)</th>
<th>Essential features of inquiry</th>
<th>Inquiry tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypotheses space</td>
<td>Question</td>
<td>Orienting and asking questions; generating hypotheses</td>
</tr>
<tr>
<td></td>
<td>Evidence</td>
<td>Planning-Investigating</td>
</tr>
<tr>
<td></td>
<td>Analyze</td>
<td>Analysis and interpretation</td>
</tr>
<tr>
<td></td>
<td>Explain</td>
<td>Modelling</td>
</tr>
<tr>
<td>Experimental space</td>
<td>Connect</td>
<td>Conclusion-Evaluation-Prediction</td>
</tr>
<tr>
<td></td>
<td>Communicate</td>
<td></td>
</tr>
<tr>
<td>Prediction Space</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Modelling indicators

Hestenes (1999), states that most science/engineering problems are solved by developing a model, from which the answer to the problem is extracted through model-based reasoning. Consequently, modeling theory of instruction is centered on the idea that scientists justify their arguments using models and students use conceptual models to investigate certain phenomena.

According to Hestenes (1999) model specification is composed by a model encompassing four types of structure, considered as Modelling indicators: (1) systemic structure which specifies: composition (internal parts of the system), environment (external entities linked to the system), connections (external and internal causal associations); (2) Geometric structure which specifies: position with respect to a reference frame (external geometry), configuration (geometric positions of the different parts of the system); (3) Time evolution of the system (change in state variables of the system); and (4) Interaction structure which is governed by the fundamental laws and the interaction properties of the constituents of the system.

Mathematical models in inquiry based approach

Simulation-based Engineering and Science is considered as the cognitive area that provides the Scientific and Mathematical basis for simulating Natural Science and Engineered systems. According to (Xie et al., 2011) computational models used for problem solving are the best way to create a curriculum that is both deeper and wider. Resorting to first principles in Physics and Mathematics to build educational tools may be considered as exaggeration by some researchers in STEM Education, but it is essential in order to bring learning experience related to authentic phenomena. Building models of simulation from first principles relies on the use of Computational Thinking (CT). Using the computational model approach, students write the mathematical relations they suppose that govern the problem, they select the simulation method, they develop the algorithm and finally they are engaging in writing source code using software or programming languages.

Argument-Driven Inquiry (ADI) laboratory activities

The importance of argumentation in STEM Education is a fundamental research issue in this cognitive area (Demircioğlu, & Uçar, 2012). Although argumentation has a significant function in STEM Education, it is not common to use arguments in Science and Mathematics education courses and laboratory activities (Kim & Song, 2005). Engagement of learners in an argumentation process shares many features with inquiry based learning, since it requires from students to make hypotheses, to develop models and use them to reason their findings. This makes researchers to consider that argumentation has a lot of common elements with inquiry based approach and enhancing argumentation could be a factor to increase students’ involvement in the inquiry process. Argumentation is also closely related to scientific reasoning. Examination of students’ scientific reasoning revealed that students have difficulty in evaluating and constructing different alternatives to an issue under investigation and argumentation has been presented as a remedy to this problem (Kuhn, 1993; Osborne et al., 2004). According to (Walker et al., 2010), laboratory activities, in which steps are provided by the instructor and the lab manual, do not enhance argumentation. The remedy to overcome this problem is to leave students free to make their own hypotheses and test them experimentally following their own procedures. This methodology is known as (ADI) methodology and was initially proposed by (Sampson, 2009). In this method students design their own experiment, make their hypotheses, select the variables that they consider as important, they conjecture about the relations of these variables and finally they verify the validity of their model by comparing the data of the experiment with the data from the textbook or other sources.

Each laboratory activity in (ADI) includes six steps (Sampson, 2009): (a) The definition of the problem by the instructor; (b) Proposing an inquiry method with collaborative groups of students; (c) Development of an argument that consists of an explanation, evidence and reasoning by each group; (d)A round-robin argumentation session; (e)Production of a lab report that answers what they were trying to do and why, what they did and why, what their argument was; (f) A double-blind peer-review of written reports. In our research we did not use collaborative groups and peer-review of written reports and each student had the task to work alone.
Threshold Concepts (TCs)

In each discipline there are some concepts or certain learning experiences that have proven particularly difficult for students to understand them. These are often labeled as “troublesome” concepts. The term threshold concept denotes concepts that are troublesome, but at the same time essential to knowledge and for understanding within particular disciplines (Loch & McLoughlin, 2012). According to (Meyer & Land, 2003), “Threshold concepts (TCs) can be considered as akin to a portal, opening up a new and previously inaccessible perspective of thinking about concepts and phenomena.”

(TCs) represent a transformed way of understanding, or interpreting, or viewing concepts without which the learner cannot develop its own understanding. A starting point in discussion of threshold concepts is a set of five characteristics identified by (Meyer & Land, 2003): Threshold concepts are transformative (occasionally triggering a significant shift in the perception of a subject), probably irreversible (unlikely to be forgotten), integrative (learners perceive previously hidden relationships), troublesome (threshold concepts embody knowledge that is troublesome for learners to grasp, knowledge of these is counter-intuitive, and cannot be easily integrated with the learner’s current mental schema) and often disciplinarily “bounded” (a threshold concept will probably delineate a particular conceptual space, serving a specific and limited purpose. Troublesome knowledge goes beyond knowledge that is difficult to understand and it is joined with incorrect mental models, misconceptions and lack of ability to transfer understanding from one context to another. According to Irvine and Carmichael (2009), the transformative aspect of threshold concepts makes them to be closely related to constructivist models of learning and, consequently, inquiry based process could have an effect on the transformative component of the threshold concepts by supporting students to develop conceptual models.

(TCs) in Mathematics

Mathematical concepts are categorized into two levels of importance: Core concepts that define important stages in learning and threshold concepts that, once grasped, will change the way the student think. In Mathematics, once a threshold concept is grasped, learners will realize they are working in a different light. A review of the literature on threshold concepts in Mathematics indicates that while there are many troublesome concepts that students struggle with, identifying (TCs) among those may not be, is a non-trivial exercise. One of the first named threshold concepts is that of limit of a function (Meyer & Land, 2003). Other (TCs) are the “addition of functions,” the ratios, solutions of inequalities, proportions and percentages (Loch & McLoughlin, 2012), while (Scheja & Pettersson, 2010) justify the inclusion of the concept of function. (Worsley et al., 2008), looking at second year university mathematics, suggested –for the area of ordinary differential equations-, the method of substitution, and multiple integration as threshold concepts. In the research of Worsley (2011), the following concepts were identified as (TCs). (1) Limits; (2) Recognition of different number systems; (3) Proof by induction; (4) Describing physical problems in the language of mathematics; and (5) The algebra of complex numbers.

(TCs) in Physics

Incomplete understanding or misunderstanding of (TCs) is likely to have long-term implications for students’ learning in Physics, including their capacity to apply their knowledge in new and different contexts. Such incomplete learning could lead to passive learning (Davies, 2006) and passive-rote-learning may explain why students may be able to apply specific methods from a cognitive area well enough to handle an exam, but may not be able to apply their knowledge to a new context.

Authors also report that five Physicists from four Australian Universities suggested the following threshold concepts in Physics: field, flux, force, gravity, momentum, entropy, impulse, energy, potential, temperature, induction, acceleration, wave-particle duality, conservation laws, space-time, gravity, relativity, equilibrium, diagrams, modelling, vectors, frames of reference, significance, approximation, orders of magnitude, uncertainty, and measurement. Physicists were asked to identify the key discipline-core concepts without reference to their possible
threshold nature, and subsequently to identify candidates for threshold concept status using criteria as they are described in (Meyer & Land, 2003).

Meyer and Land (2003) also refer to the concept of gravity – the idea that any two bodies attract one another with a force that is proportional to the product of their masses and inversely proportional to the distance between them – and state that this represents a threshold concept, whereas the concept of a centre of gravity does not, although the latter is a core concept in many of the applied sciences.

**Methodology**

**Participants**

Seventy-nine (79) pre-service engineering school prospective teachers (at the third year of their studies) participated in this research as experimental group. Students were enrolled in the course “Applications of Information and Communication Technologies in Education” in a Tertiary Institution in Greece and they had already studied the concepts under discussion from relevant courses in Physics and Mathematics.

**Materials and procedure**

The research lasted four months, during the Academic Year 2013-14. A three-hour course was carried out on a weekly basis, for thirteen weeks, and one and a half hour was devoted to the use of the computational experiment and the development of computational models as a cognitive acceleration procedure. The rest of the time was used to discuss how to create an inquiry based scenario that includes the essential features of inquiry approach and the essential features of Computational Thinking. In some cases students faced problems with the concepts from Physics and Mathematics, and the Instructor (Author of the article) helped students to clarify them providing the relevant theory and definitions to students.

Our strategy was to develop specific examples from Physics and Mathematics as cognitive acceleration lessons in the form of a pedagogical scenario, which integrates the essential features of inquiry with the methodology and the spaces of the Computational Experiment-- Mathematical Modelling (CEMM) and the features of Computational experiment.

One aim of this research was the design of inquiry based-computational scenarios relevant to specific threshold concepts and the investigation of the impact of the scenarios to the acquisition of these thresholds. Another aim was to study the impact of our approach to students’ argumentation level.

According to Davies (2006), (TCs) may be identified with many ways. One of them is by comparing the reflections of experts and comparing them with the reflections of students. Another approach is by using semi-structured interviews with students (Park & Light, 2009). However this approach is considered as resource intensive. In our approach we used the less resource intensive approach based on the views of experienced teachers. During a training seminar, which lasted 360 hours, on the “exploitation of ICT in Science Education” we discussed with experienced teachers the nature of (TCs) they experienced as teachers or they had experienced at their studies.

Combining the results from the interviews and the literature review, we decided to work with the following three threshold concepts: (1) The gravitational force; (2) The addition and multiplication of complex numbers; and (3) The calculation of integrals and in particular the calculation of an integral that gives the value of number π. For the last case we consider that it covers two threshold concepts: The calculation of integrals and the concept of measurement, i.e., students should understand that calculation of integrals is an approximate method connected to the calculation not only of abstract entities but of entities we use in everyday life like the number π. For the case of integrals we used the Monte Carlo method to calculate them because we wanted to involve students in the concept of the stochastic processes. Initially, students were exposed –for nine (9) hours-to models in Physics and Mathematics from the repository (www.opensourcephysics.org). Students were asked to identify –under the guidance of the Instructor- the essential structures of models (modelling indicators) as they are described by Hestenes (1999). Most of the
models in this repository were developed by the software Easy Java Simulations (EJS) and they were relevant to the discipline of Engineering but they also contained concepts form Physics and Mathematics, i.e., they could be described as STEM examples. EJS software does not demand knowledge of the programming language Java and its interface is not difficult for use since it resembles the interface they have met in traditional lectures embedded with specific tools to write down mathematical expressions. Students discussed with the author the use of modelling indicators and they were asked to recognize their existence in certain models presented by the Instructor from the repository www.opensourcephysics.org. After three sessions of Instruction (nine hours), a questionnaire was provided to students in order to measure students’ recognition of Modelling indicators (systemic structure, geometric structure, temporal structure and interaction structure) in the models exhibited in the class by the Instructor. The score of students for the recognition of the modelling indicators was based on the following scale: For the Modelling indicators (systemic structure, geometric structure, temporal structure and interaction structure) 4 stands for the fully favourable part, i.e., students recognized all the indicators in the model, 3 for the recognition of the existence three of the indicators of three indicators(less favourable part), 2 for the recognition of two indicators (partially favourable part) and 1 for the recognition of one indicator (unfavourable part).

Before the intervention students participated in another test related to their conception for the threshold concepts, namely: (1) the motion under the gravity force; (2) the addition and multiplication of complex numbers: and (3) the calculation of integrals. In the Table 2, 3, and 4, we present the questionnaires given to students relevant to the three thresholds concepts we wanted to examine. Each assignment was marked with minimum score 1 and maximum score of 5 (for the five levels of students’ conception), so the minimum score for a student is 3 and the maximum score 15. Table 2 presents the questionnaire for investigation of students’ conceptions for the gravitational force.

**Table 2. Questionnaire/rubric for investigation of students’ conception of gravitational force**

<table>
<thead>
<tr>
<th>Level 1</th>
<th>No conception of gravitational force; no thought of it in relation to experimental outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2</td>
<td>Gravitational force is an absurd concept useful only to scientists</td>
</tr>
<tr>
<td>Level 3</td>
<td>Gravitational force is seen as not a universal force</td>
</tr>
<tr>
<td>Level 4</td>
<td>Gravitational force is a universal force, constant, and applies at a distance</td>
</tr>
<tr>
<td>Level 5</td>
<td>Gravitational force is not constant, is universal, and we have to find the resultant force on a body which interacts with other masses</td>
</tr>
</tbody>
</table>

Table 3 presents the questionnaire for investigation of students’ conceptions for complex numbers.

**Table 3. Questionnaire/rubric for investigation of students’ conception for complex numbers**

<table>
<thead>
<tr>
<th>Level 1</th>
<th>No conception of complex numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2</td>
<td>Complex numbers are an absurd concept useful only to scientists</td>
</tr>
<tr>
<td>Level 3</td>
<td>Complex’s numbers addition is like the addition of real numbers but their multiplications is something absurd</td>
</tr>
<tr>
<td>Level 4</td>
<td>Complex numbers can be used only in mathematics, operations with them are easy and well defined, is a nice construct but without applications</td>
</tr>
<tr>
<td>Level 5</td>
<td>Complex numbers can be used in transport phenomena; they can be applied in heat transfer and wave’s propagation, while operation with complex numbers is something which can be handled if we follow the rules and the axioms</td>
</tr>
</tbody>
</table>

Table 4 presents the questionnaire for investigation of students’ calculation of integrals and the number $\pi$.

**Table 4. Questionnaire/rubric for investigation of students’ conception of calculation of integrals and the number $\pi$**

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Number $\pi$ is equal to 3.14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2</td>
<td>Integrals have always a specific value</td>
</tr>
<tr>
<td>Level 3</td>
<td>There are certain techniques to calculate integrals and random processes do no lead to accurate values</td>
</tr>
<tr>
<td>Level 4</td>
<td>Random process can produce results in the interval of integration, but results are not valid</td>
</tr>
<tr>
<td>Level 5</td>
<td>Integrals can be calculated with any method that considers the points that function under integration has the greater values</td>
</tr>
</tbody>
</table>

During the intervention the Instructor developed from scratch three examples in the form of scenarios, one for each of the threshold concepts. This procedure lasted for 9 hours following the three sessions where models were
exhibited to students. For each example presented, Instructor developed a scenario including the essential features of inquiry based learning and the Computational Thinking (CT). He also created from scratch the source code explaining the algorithm included. Next, for six sessions of instruction, students had the task to either select specific models from the repository (www.opensourcephysics.org) and modify them or to develop from scratch their own models. In any case they had to create one scenario using their model and the features of inquiry based teaching and learning approach and the (CT) as well as the level of argumentation as it is described below.

Computational Thinking abstraction was used by students in defining patterns (e.g., for the case of aurora phenomenon, students discovered the helix motion of particles in Earth’s magnetic field), parameterization (definition of parameters and relation between variables and parameters), and development of algorithms for processes that take data and execute a sequence of steps in the model part of the EJS software.

Students were asked to be involved in the Driven Inquiry approach for Laboratory experiments and to design their own lab activities without provision of steps by the lab manual. These lab activities should be included in the scenarios that students had to create. Analysis of ADI reports was based on the methodology developed by (Osborne et al., 2004). Argumentation levels for each report were determined through the levels suggested by (Erduran et al., 2004; Demircioğlu, & Uçar, 2012). The argumentation levels are shown in Table 5 and each was scored from 1 to 5 according to the quality of argumentation.

Table 5. Analytical framework used for assessing the quality of argumentation

| Level 1 | Level 1 argumentation consists of arguments that are a simple claim versus a counter-claim or a claim versus a claim. |
| Level 2 | Level 2 argumentation has arguments consisting of a claim versus a claim with either data, warrants, or backings but do not contain any rebuttals. |
| Level 3 | Level 3 argumentation has arguments with a series of claims or counter-claims with either data, warrants, or backings with the occasional weak rebuttal. |
| Level 4 | Level 4 argumentation has arguments with a claim with a clearly identifiable rebuttal. Such an argument may have several claims and counter-claims. |
| Level 5 | Level 5 argumentation displays a next ended argument with more than one rebuttal |

During the course students had to develop their own scenario that should include the seven features of Inquiry, the spaces of the computational experiment, the modelling indicators and the levels of argumentation. Their score was based on the following scale. For the seven features of the inquiry based learning, the minimum score is 1 if only one feature was included and the maximum score is 7 if all the seven features of inquiry were included in a scientific way. For the Modelling indicators (systemic structure, geometric structure, temporal structure and interaction structure), marking was described above. For the quality of argumentation, the minimum score is 1 (for level 1) and the maximum score is 5 (for level 5).

After the intervention students participated in another test related to their understanding of the (TCs) namely: (1) the motion under the gravity force; (2) the addition and multiplication of complex numbers; and (3) the calculation of integrals. Each assignment was marked with minimum score 1 and maximum score of 5 (for the five levels of students’ conception), so the minimum score for a student is 3 and the maximum score 15.

The variable “Learning Performance,” was based on a test given to students which included exercises from Physics and Mathematics, spanning from classical mechanics to electromagnetism. The test contained twenty (20) questions, the maximum score for the learning performance was 100 and the minimum zero (0). The questions were based to the concepts they had been taught in the class during the intervention with the computational experiment.

Learning performance was also used in regression-stepwise- analysis in order to study the effect of the computational experiment, the ADI level and the use of Inquiry features in the dependent variable “Learning Performance.”

We also examined the differences in the mean values-before and after the intervention- for the variables “use of Modelling Indicators” and “acquisition of (TCs).”
Results

At the beginning of the course, the Instructor presented applications which included the model development, the simulation method and the selection of the numerical method in order to construct an algorithm. Most of the applications were developed using the Java language and the (EJS) with focus on models and the algorithm used to create the model. Most of the models were included in the repository www.opensourcephysics.org.

After the exploration of the models, the Instructor started to develop his own models based on the tree threshold concepts referred above.

During the development of the models, students discussed with the Instructor the fundamental concepts from Physics and Mathematics related to the threshold concepts under discussion and in some cases they had to recall knowledge from these cognitive areas.

Finally, for six sessions, students had the task to create their own pedagogical scenario based on the computational experiment approach, which should include the features of inquiry based learning combined with the modelling indicators and the ADI approach. During the six sessions, the Instructor guided the students on how to use the models from the repository www.opensourcephysics.org, while he also intervened to teach concepts from Physics and Mathematics relevant to the models developed by the students.

For example, a student developed a model for the Aurora, and the discussion focused in the motion of a charged particle in a non-uniform magnetic field. All the scenarios developed by the students were based on concepts related to the three threshold concepts we mentioned before. For the (TCs) the scores before and after the intervention are presented in Table 6.

| Table 6. Mean scores for the variable “Acquisition of Threshold Concepts” |
|-----------------------------|-----------------|-----------------|-----------------|
| Acquisition of (TCs) Before the Intervention | 7.531 | 79 | 2.246 | .252 |
| Acquisition of (TCs) After the Intervention | 10.468 | 79 | 2.800 | .315 |

A paired-samples t-test was conducted to determine the effect of the intervention on students’ acquisition of (TCs). There was a significant difference before the intervention (M = 7.53, SD = 2.246) and after the intervention (M = 10.468; SD = 2.800), where the difference in the mean values is statistically significant (t = -20.813; df = 78, p < .001).

In students’ modelling recognition (i.e., recognition of the modelling Indicators of Hestenes), before the intervention, students scored (M = 2.227, SD = 0.750) and after the intervention the score for the use of Modelling Indicators in their scenarios, the score is (M = 2.924; SD = 0.944), where the difference in the mean values is statistically significant (t = -10.120; df = 78, p < .001).

Using stepwise linear regression analysis, we considered the learning performance as the dependent variable and we investigated the impact of the other variables in the dependent variable.

The results indicate (Table 7) that the variable with the greater impact is the “use of Inquiry Features.”

| Table 7. Regression analysis results for the variable “Learning Performance” |
|-----------------------------|-----------------|-----------------|-----------------|
| Variables | R | B | F | t |
| 1 | .953<sup>a</sup> | 6.523 | 770.375 | 7.406 |
| 2 | .978<sup>b</sup> | 10.971 | 84.836 | 6.936 |
| 3 | .981<sup>c</sup> | 5.546 | 10.093 | 3.177 |

Note. Dependent Variable: Learning performance. <sup>a</sup> = Predictors: (Constant), Inquiry Features; <sup>b</sup> = Predictors: (Constant), Inquiry Features, Modelling Indicators After; <sup>c</sup> = Predictors: (Constant), Inquiry Features, Modelling Indicators After, ADI.
Using stepwise regression analysis, we observe that the variable “use of Inquiry Features” is responsible for the 92% of the Learning performance \( (F_{1,77} = 475.047, p < .001) \).

The “use of Modelling Indicators” is responsible for the 2% \( (F_{1,76} = 25.546, p < .001) \) and another 1% is from the variable “ADI” \( (F_{1,75} = 8.388, p < .001) \).

The results for the variable “Learning Performance” motivated us to go for another regression-stepwise-analysis, with dependent variable the “Acquisition of Threshold Concepts” after the intervention and the results are presented in Table 8.

### Table 8. Regression analysis results for the variable “Acquisition of Threshold Concepts”

<table>
<thead>
<tr>
<th>Variables</th>
<th>( R )</th>
<th>( B )</th>
<th>( F )</th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.928(^a)</td>
<td>0.613</td>
<td>475.047</td>
<td>4.467</td>
</tr>
<tr>
<td>2</td>
<td>.946(^b)</td>
<td>0.787</td>
<td>25.546</td>
<td>3.194</td>
</tr>
<tr>
<td>3</td>
<td>.952(^c)</td>
<td>0.788</td>
<td>8.388</td>
<td>2.896</td>
</tr>
</tbody>
</table>

*Note. Dependent Variable: Acquisition of Threshold Concepts. \(^a\) = Predictors: (Constant), Inquiry Features; \(^b\) = Predictors: (Constant), Inquiry Features, Modelling Indicators After; \(^c\) = Predictors: (Constant), Inquiry Features, Modelling Indicators After, ADI.*

Using stepwise regression analysis, we observe that the variable “Use of Inquiry Features” is responsible for the 92% of the Learning performance \( (F_{1,77} = 475.047, p < .001) \).

The “Use of Modelling Indicators” is responsible for the 2% \( (F_{1,76} = 25.546, p < .001) \) and another 1% is from the variable “ADI” \( (F_{1,75} = 8.388, p < .001) \).

**Discussion**

The objective of this study was to implement the computational-inquiry based teaching and learning approach, and through the use of ADI levels and modelling indicators, measure student’s acquisition of specific (TCs).

Results show that there was a statistically significant difference between the use of modelling indicators before and after the intervention in the class.

The main predictor for the variable “Learning Performance” is the “Use of Inquiry Features,” while the other main prediction is coming from the “use of Modelling Indicators.” The main predictor for the transformative character of (TCs) is the again the “Use of Inquiry Features,” while the other main prediction is coming again from the “Use of Modelling Indicators,” as was also the case for the dependent variable “Learning Performance.”

Our results indicate that students’ engagement in models of simulation could enhance their conceptual understanding of the (TCs) and this is probably an answer to the question about the identification of threshold concepts and how would they alter what we teach, how we teach, and how we assess them (Rountree & Rountree, 2009).

The (ADI) variable seems also crucial in students understanding of threshold concepts, since it is strongly correlated with the acquisition of threshold concepts \( (r = 0.925) \) while it also appears in the regression analysis for the dependent variables we discussed before.

Our findings are in accordance to research results by (Walker et al., 2010), who reported that undergraduate students in (ADI laboratory) sections improved significantly their ability to use evidence and reasoning to support a claim compared to the students in traditional lab sections. (Walker et al., 2010) state that assessments demonstrated that students in the ADI laboratory sections were better able to use evidence and reasoning to support a conclusion, and had a more positive attitude towards science. The proposed instruction model, including the (ADI) methodology, was designed to provide students with an opportunity to develop methods to generate data, to carry out investigations and use data to answer research questions. Results indicate that when students are involved in the (ADI) methodology, they were able to use evidence and reasoning to support a conclusion, and had a more positive attitude towards using models in EJS.
We started this research study from the theoretical point of view that theories in Science and Mathematics are developed through a process of successive elaboration and improvement in which scientific models are created and modified to account for new phenomena that are uncovered in exploring a domain. According to (White & Frederiksen, 1998) this scientific process resembles to the inquiry process where students are involved in learning scientific concepts and develop conceptual models.

In our research the experiments were set in the context of computational models and models used not only as representations of concepts but they contained mathematical algorithms in order to be in alignment with the Computational Thinking (CT) and the scientific problem paradigm.

Students’ choices spanned many thematic areas from (e.g., investigating resistance, system of automata, motion in an electromagnetic field, use of complex numbers to study laminar flow). Students developed abilities such as “logical reasoning” and “critical thinking,” as it was apparent from student’s reports and the scenario they developed. In the reports written by the students for their involvement in simulation based activities, we noticed that some students proposed explanations very close to scientific ones and they collected data to support their research questions and justify their claims. This process was in alignment with certain features of inquiry, namely the Evidence the Analysis, the Explanation, the Connection, the argumentation process and the ADI methodology.

After the intervention, discussions with the students revealed that this type of instruction could help them to act as a scientist. They expressed the view that EJS is a very useful tool to declare variables and write down the mathematical relations.

Students found the examples useful and necessary for becoming more knowledgeable about instructional design, as it was clearly articulated by a student who wrote: “When I first read the requirements for the scenario, I felt so uncertain about it but I learned so many things about designing instruction with computational methods and inquiry based method.” Other students also wrote that “the inquiry based teaching and learning approach cautiously made me aware of the fact that in order to do calculations I first had to reflect on model development through the EJS tool and this sometimes confronted me with their incorrect conceptions about the variables used and the equations of state.”

**Conclusion**

The present paper initially raised and discussed several epistemological issues regarding the construct of computational models for the purpose of clarifying (TCs) and proposed models which included integrated scenario of computational thinking with inquiry based teaching and learning approach.

During the intervention Teaching Observation Protocol was used based on Lesson design and implementation (i.e., for example if the instructional strategies and activities respected students’ prior knowledge and the preconceptions inherent there in, if this lesson encouraged students to seek and value alternative modes of investigation or of problem solving).

In the research one instruction model for developing Inquiry based Computational pedagogical scenario was described. This instruction approach was proposed as an approach for understanding and promoting a methodology toward the development of (CT) in inquiry based models in Physics and Mathematics.

At the heart of the instruction model is the development of models of simulation based on the EJS tool and this could be of help to teachers as they could develop more sophisticated ways of thinking about how technology can transform their teaching practices. Finally, despite the optimistic results of the results of this research, we consider that the development of Computational Models-Inquiry based scenario, is not an easy task. Consequently, intensive, coordinated and dedicated systematic efforts need to be planned and implemented in pre-service and in-service education programs in order to develop curriculum activities based on this instruction model and future research should be undertaken in order to validate, modify, or improve the framework proposed in this research paper.

Finally, research shows teachers and many staff members are reluctant to suggest (TCs) concepts, or delineate them from the core concepts (Wilson et al., 2010). This research study could provide the framework to educators to deal
with threshold concepts in a more systematic way by developing learning process that integrate the inquiry features and the computational models.

References


ABSTRACT

This paper used a novel literature review approach—co-citation network analysis—to illuminate the latent structure of 87 empirical papers in the field of young children's learning with technology (YCLT). Based on the document co-citation analysis, a total of 206 co-citation relationships among the 87 papers were identified and then graphically visualized as a network diagram. Accordingly, the five most highly co-cited pairs in the co-citation network were identified. Moreover, seven research streams in the network (three located in the main component, four scattered in the peripheral area) were characterized based on a general assessment of the document research domains. Of these, “technology-assisted language learning” was denoted as the main stream, and “digital literacy” (active from 2008 to 2013) was identified as the emerging stream of the YCLT research network. Looking inside, the six most central papers within the above two research streams were recognized in terms of centrality measures. Furthermore, on the basis of the analysis of follow-up research, the current and future extension research topics are discussed.

Keywords
Young children's learning with technology (YCLT), Document co-citation analysis, Social network analysis, Literature review

Introduction

With the fast-developing technologies and prevalent technology use in everyday life, the issue of young children's learning with technology (YCLT) has become increasingly important (Plowman, Stephen, & McPake, 2010). The growing trend of YCLT has attracted researchers’ attention to the need to review the contents of the YCLT literature. They have confirmed the effects of technology on various domains of young children’s learning. For example, Blok, Oostdam, Otter, and Overmaat (2002) performed a meta-analysis to review 42 research articles regarding computer-assisted instruction in support of beginning reading instruction. These researchers found that computer-assisted instruction was generally effective in terms of facilitating children’s beginning reading abilities. Lankshear and Knobel (2003) developed a four-quadrant framework to categorize 31 studies that adopted new technologies to support early childhood oral language and literacy development. In their work, most of the studies were assigned to the quadrant that used stand-alone technological devices to facilitate children’s encoding and decoding skills. Yelland (2005) synthesized the literature on the use of computers in early childhood education from 1994 to 2004. Based on a qualitative review, she indicated that the use of computers is helpful for promoting young children’s language, cognitive, and social development. Slavin and his colleagues (2009) presented an alternative meta-analytic review, called best-evidence synthesis, to investigate the effectiveness of reading programs for elementary students. In their review, technology-enhanced reading was found to be correlated with reading achievement.

Recently, Hsin, Li, and Tsai (2014) conducted a content analysis to investigate 87 empirical studies of young children’s use of technology for learning. They concluded three factors influencing children’s learning through technology, and identified “children’s development of digital literacy” as the emerging trend. Despite the abundant information obtained from Hsin et al. (2014) and the other abovementioned reviews, how YCLT studies relate to each other is still unknown. Moreover, some important issues need to be explored: What are the most often cross-referenced research streams of YCLT? To what extent have these streams in the YCLT research network been extended by the follow-up papers? In addition, from the perspective of a large number of citers, what is the latent structure in the field of the main YCLT literature?

To answer the above research questions, this study presents a structural overview of the current YCLT research based on the analysis of large citation data. Two complementary methods were used to achieve the above purposes:
document co-citation analysis (DCA) and social network analysis (SNA). First, co-citation analysis, known as a powerful computational analysis from the field of bibliometrics, is used to detect the most frequently referenced topics underlying the literature structure (Small, 1973). A highly co-cited pair indicates that the nature of these two research papers is highly correlated; thus, follow-up studies are most likely to cite them together. All these citing and cited papers together form an invisible research network around the same issue (Small, 1973). Therefore, with the use of DCA, researchers can identify the core research issues within YCLT research based on the bibliographic evidence. Next, embedded in graph layout, SNA is advantageous for profiling the complex co-citation matrix which is obtained from co-citation ties (or research pairs) of the literature network (Borgatti, Everette, & Freeman, 2002). Moreover, based on the measures of centrality, SNA can identify the most central papers within the focal network. For example, some critical bridging papers that connect different research streams can be identified with the measure of betweenness centrality (Freeman, 1979).

In sum, the current study goes beyond Hsin et al.’s (2014) descriptive overview to determine the closeness and relationships of the 87 articles based on two novel methods. Therefore, this study contributes a bibliographic insight into the review study and provides a visualization overview of the contemporary YCLT literature.

**Methods**

**Document co-citation analysis**

The most important step for co-citation analysis is the collection of research data. Based on Hsin et al.’s (2014) research review on the influence of young children’s use of technology for learning, a total of 87 empirical articles on YCLT were identified as candidate papers for this study. Accordingly, the first step of the analysis was to retrieve the bibliographic citation data of these 87 core documents from the Web of Science (WoS).

In line with the quality of these 87 journal papers, only journal citations were counted. For example, Clements and Sarama’s research (2007) is the most highly cited paper among the 87 core papers, as it has been cited by 47 other journal articles (11 citations from book chapters and conference proceedings were excluded). Then, all of these 47 citing papers were retrieved from the WoS and the 47 citation counts were accredited to the work of Clements and Sarama (2007). By adopting the selection criteria of the citation data, a pool of 87 core papers together with all 560 citing articles were collected from the WoS. Overall, there are over 200 journals in this dataset, indicating the diverse nature of YCLT research. The data retrieval was performed on October 30, 2014.

Next, a series of computation analyses for constructing the co-citation matrix were processed. Specifically, all of the above 87 core papers were paired document by document, and then the co-citation counts for each research pair were calculated by matching the reference records within the 560 articles. According to Small (1973), when a pair of core studies (i.e., two of the 87 core papers) is referenced in the same citing article (i.e., 560 follow-ups), the frequency of co-citation for that research pair is then reckoned as one. For example, two highly cited YCLT papers, Clements and Sarama (2007; 2008), have been cited 47 and 28 times, among which, 10 follow-up studies have co-cited the two works. Endorsed by these 10 citing papers, the value of co-citation for this research pair (Clements & Sarama, 2007, 2008) in the co-citation matrix is input as 10. Likewise, the co-citation relationships among all the research pairs of the 87 core YCLT papers were calculated based on the large citation data from the 560 follow-up references. Among the 87 papers, however, 34 which have not yet been co-cited by any other YCLT studies were excluded because papers without co-citation were not able to create a co-citation linkage on the co-citation map. Note that most of these excluded papers (24 out of 34) were published after 2010, and thus have less chance of having co-citation relationships with other core YCLT papers. The computation results for the remaining 53 papers were then compiled into a symmetric co-citation matrix of 53 rows and 53 columns. The full list of 53 core papers is available at: http://goo.gl/XLJuqq.
Social network analysis

In recent educational review studies, social network analysis (SNA) is a relatively novel technique for visualizing results; using this technique, one can more easily understand the research network of the most prominent studies (Tang, Tsai, & Lin, 2014; Tang & Tsai, 2016). The visible network displays the academic interactions in the bibliometric contexts like a scholarly communication of invisible colleagues (i.e., the co-citation network) that previous researchers may be unaware of themselves. Therefore, the use of SNA in this review permits the exploration of existing linkages among 53 YCLT articles.

Furthermore, three measures of centrality (i.e., the degree, closeness, and betweenness centralities) were also adopted to profile the centrality features of the most central works in the YCLT co-citation network. The degree centrality is the most straightforward measure of the connectedness of each node (which is the label for a core paper presented on the graph). The closeness measures nodes’ independence (i.e., less connectedness); therefore, the more central role a node plays, the less closeness centrality it has. The betweenness centrality is the index reflecting the bridging role of a node in the network; the more critical bridging role a node plays, the more strongly it controls the communication between sub-networks (called a “broker”) (Freeman, 1979). In this study, the measure of degree centrality is presented as the amount of connectedness which a node has. The closeness and betweenness centralities are presented as normalized measures based on Freeman’s (1979) approach.

Results

Results of the document co-citation analysis

Through matching the reference data within 560 citing articles, a total of 206 co-citation ties within 53 core papers were found in the co-citation matrix. The higher the correlation between research pairs, the more bibliographical similarity they will have. In the co-citation matrix, the highest co-cited research pair is composed of two studies by Shimizu and McDonough (2006) and Donker and Reitsma (2007). The two studies are related to the use of a computer mouse by young children and have been co-cited 14 times. In Shimizu and McDonough (2006), they taught three four-year-olds with a developmental delay the skill of pointing with a computer mouse. The children learned pointing after receiving instructional prompts. In Donker and Reitsma (2007), they investigated the efficiency of two ways of moving objects over the screen, drag-and-drop and click-move-click. They found drag-and-drop more efficient for young children.

Another highly co-cited pair consists of two studies by Clements and Sarama in 2007 and 2008 (co-cited 10 times), which focused on the evaluation of a research-based mathematics curriculum, Building Blocks, for preschoolers of low-income families. The curriculum integrated computer and off-computer activities and included three types of materials: computer software, manipulatives, and printed materials. In 2007, they used a quasi-experimental design and implemented the curriculum in four classrooms; in 2008, they then adopted an experimental design and expanded the participants to 36 classrooms. Both studies showed that the children improved their mathematics performances after attending the intervention program.

The remaining three highly co-cited pairs are closely tied with each other as a triangle as shown in Figure 1. Of these, the first pair is composed of the studies by De Jong and Bus (2004) and Chera and Wood (2003), which received the most co-citations, numbering 11. Moreover, these two studies are both tied with Segers, Takke, and Verhoeven’s (2004) article, forming the remaining two highly co-cited pairs. The topic of the three cross-linked studies is the effectiveness of electronic books for young children’s literacy development. Chera and Wood (2003) investigated the effect of e-books on the development of phonological awareness of 3-6-year-olds in comparison with children who did not read an e-book. They found that children who read e-books improved their phonological awareness. Sharing the same research question, the research of De Jong and Bus (2004) and Segers et al. (2004) compared the effects of children’s reading of e-books independently with the effects of children being read to by an adult. The results revealed that children’s comprehension of the stories and vocabulary learning were similar in the two reading conditions—being read to by a computer or by an adult.
Overall, the co-citation analysis has identified five highly referenced YCLT research pairs, which have a tight connection and a high similarity with each other. From the large cited’s perspective, these five research pairs provide the research foci of the current YCLT literature, including the use of a computer mouse by children, a children’s technology-assisted mathematics curriculum, and the effect of e-books on children’s literacy development. It is noteworthy that although these three research topics with high co-citation numbers may indicate that they are the most influential core works of YCLT, the analysis of their follow-up studies revealed that they are immature and have not yet been fully explored.

Results of the social network analysis

To further illuminate the configuration of the research network of the 53 YCLT papers, social network analysis was conducted to visualize all of the co-citation ties in the co-citation network. By using the NetDraw module of UCINET (Borgatti et al., 2002), the layout of the 53 research works was denoted as 53 nodes, and every co-citation tie was a link between each pair of nodes. The thickness of the co-citation links stands for the weights of the different co-citation ties.

As a result, the three most highly referenced ties, which have been co-cited over 10 times, are highlighted as strong ties and are presented as three red solid edges in the diagram. The other two blue lines represent moderate ties, which have been co-cited 5 times. While those edges which received the strength of ties ranging from 2 to 4 are presented as black solid lines, the rest of the ties with minimal co-citation linkages are marked as gray lines. As shown in Figure 1, the whole structure of the YCLT co-citation network can be divided into two sectors: (1) the main component (in the center), and (2) the isolated components (on the upper-right side). The structural relationships among the nodes are further explored according to the research purposes and learning domains of the 53 studies.

Overall, seven research streams are identified. Three major research streams are acknowledged in the main component of the network (i.e., technology-assisted language learning, digital literacy, and technology competence) and the other four (relatively marginal) streams (including mathematical learning with technology, social interaction through technology, video-enhanced learning for autistic children, and technology as learning companion) are mapped into the isolated components. A summary is presented in Table 1.
<table>
<thead>
<tr>
<th>Research streams</th>
<th>Papers</th>
<th>Main journals</th>
<th>Active year (average year)</th>
<th>Total co-citations</th>
<th>Average co-citation (per paper)</th>
<th>Total citations</th>
<th>Average citation (per paper)</th>
<th>Average citation (per year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology-assisted language learning</td>
<td>29</td>
<td><em>Computers &amp; Education</em> <em>(6)</em>&lt;br&gt;<em>Reading and Writing</em> <em>(3)</em></td>
<td>2003-2012 (2008)</td>
<td>254</td>
<td>8.8</td>
<td>292</td>
<td>10.1</td>
<td>45</td>
</tr>
<tr>
<td>Digital literacy</td>
<td>10</td>
<td><em>English Teaching-Practice and Critique</em> <em>(2)</em>&lt;br&gt;<em>Computers &amp; Education</em> <em>(2)</em></td>
<td>2008-2013 (2011)</td>
<td>46</td>
<td>4.6</td>
<td>37</td>
<td>3.7</td>
<td>12</td>
</tr>
<tr>
<td>Technology competence</td>
<td>3</td>
<td><em>British Journal of Educational Technology</em> <em>(1)</em>&lt;br&gt;<em>Computers &amp; Education</em> <em>(1)</em>&lt;br&gt;<em>Research in Developmental Disabilities</em> <em>(1)</em></td>
<td>2003-2007 (2005)</td>
<td>34</td>
<td>11.3</td>
<td>47</td>
<td>15.7</td>
<td>5.4</td>
</tr>
<tr>
<td>Social interaction through technology</td>
<td>3</td>
<td><em>Cambridge Journal of Education</em> <em>(1)</em>&lt;br&gt;<em>Computers &amp; Education</em> <em>(1)</em>&lt;br&gt;<em>Research Papers in Education</em> <em>(1)</em></td>
<td>2010-2012 (2010)</td>
<td>9</td>
<td>3</td>
<td>13</td>
<td>4.3</td>
<td>3.3</td>
</tr>
</tbody>
</table>
Main component of the co-citation network

The main component of the YCLT network consists of 42 YCLT papers, representing the major YCLT research community. Studies included in this component are well connected and are covered by three research streams: (1) technology-assisted language learning, (2) digital literacy, and (3) technology competence.

Graphically, “technology-assisted language learning” constitutes the largest area of the main component and includes 29 papers, indicating that it is the most researched stream in the currently reviewed YCLT studies. In addition, this research stream has received the highest numbers of citations and co-citations, signifying the emphasis on language learning in YCLT research. Moreover, the papers of “technology-assisted language learning” are published between 2003 and 2012, that is, right across the time period of this review. This indicates that the use of technology to support language learning is a longstanding and continuing research stream. While most papers in this stream were to compare the effects of technology-assisted learning and the traditional method, some studies focused on the effects of different pedagogical or mechanical designs, and the mediator-roles of adults.

Electronic books (e-books) and computer programs were generally implemented to facilitate young children’s language learning. For example, Helland, Tjus, Hovden, Ofte, and Heimann (2011) examined the effectiveness of different computer-based training programs on children with developing dyslexia. They found that bottom-up training (i.e., from sound to meaning) is appropriate for young children’s pre-literacy skills, such as phonological awareness, and top-down training (i.e., from meaning to sound) has a significant effect on promoting literacy skills, such as word reading and spelling. On the other hand, the use of e-books (e.g., animated multimedia talking books) was found to have significant effects on promoting young children’s phonological awareness (Chera & Wood, 2003). Therefore, there are two sub-networks distinguished by the types of technology adopted to support learning, that is, “e-books” and “computer programs.”

Most research nodes of “e-books” are connected in the right area of the co-citation network. This research sub-stream includes 4 highly cited papers (i.e., cited more than 10 times) and 7 more recently published works. The former three of four highly cited papers (Chera & Wood, 2003; De Jong & Bus, 2004; Segers et al., 2004), which link as a triangle-shaped tie in this sub-stream, provide a key foundation for the research network of e-book related studies. Overall, the high average number of co-citations (11.7) also implies frequent interactions in this community.

The other sub-stream of language research consists of 15 studies that implemented computer programs to facilitate children’s learning of literacy. The network connection in computer programs research (7.9) is less tight than it is in e-books (11.7). Although five papers are cited more than 10 times, no strong or moderate co-citation ties are found in this sub-stream. Two computer programs were highlighted: while some researchers investigated the effectiveness of a technology-rich curriculum, named PictoPal, in terms of promoting emergent literacy, the others evaluated a web-based tool, Abracadabra, designed for supporting reading.

The second research stream derived from the main component is “digital literacy.” This stream comprises 10 studies emphasizing young children’s digital literacy by (1) exploring the associations between children’s technology experience in daily life and their development; and (2) investigating children’s access and use of computers outside schools and informing educators of the importance of taking children’s technological abilities and experience into account when arranging appropriate learning environments and experiences for young children. Most of the papers in this research stream were published after 2011. Compared to studies of language learning, “digital literacy” is a relatively new and emerging research stream that has started to draw attention from researchers during recent years.

The third research stream is labeled as “technology competence” and includes only three papers. Donker and Reitsma (2007) and Shimizu and McDonough (2006) are the most highly co-cited YCLT works in the current review and represent this research stream. These two works, which are cross-referenced as many as 14 times, investigated...
children’s use of the computer mouse. The other study was conducted by Romeo and his colleagues (2003) to explore children’s interactions with touchscreens.

**Isolated components of the co-citation network**

Compared with the main component of the network, the co-citation relationships in the isolated components are relatively loose, with only 4.7 co-citation counts on average. Four isolated components are located in the upper right corner of Figure 1, each representing a small research stream: (1) mathematical learning with technology, (2) social interaction through technology, (3) video-enhanced learning for autistic children, and (4) technology as a learning companion.

Although shown as marginal research streams, they include the top four highly cited papers of all the reviewed YCLT studies: Clements and Sarama (2007) (47 citations), Cihak, Fahrenkrog, Ayres, and Smith (2010) (41 citations), Ryokai, Vaucelle, and Cassell (2003) (28 citations), and Clements and Sarama (2008) (28 citations). The importance of these isolated research streams represented by the four leading papers cannot be overlooked.

Two highly cited works of Clements and Sarama (2007, 2008) represent the research stream of “mathematical learning with technology.” In these two works, a computer-supported math curriculum (Building Blocks) was adopted to promote the concepts of numbers and geometry. The average co-citation of this research stream is the highest among the four isolated papers. This indicates that the papers in this stream are often cross-referenced and are highly recognized, even though this research community is very small and not part of the main research stream.

“Social interaction through technology” is the most up-to-date research stream in the YCLT network that is under development. The studies were published between 2010 and 2012 and started to discuss the roles of social interaction (including adult-child or child-child interactions) in supporting children’s learning during technology use. Two frequently co-cited works (Plowman et al., 2010; Wolfe & Flewitt, 2010) both employed the sociocultural approach to discuss the importance of adults’ roles in children’s use of technology and how adult-child interaction facilitates children’s literacy learning, computer competence skills, understanding of the cultural role of technologies, and communication with family members.

The other two smallest research streams both consist of two studies. In “video-enhanced learning for autistic children,” Dauphin, Kinney, and Stromer (2004) employed video-enhanced activity training to teach play and communication skills to a child with autism. Cihak et al. (2010) adopted video modeling to improve transitional behaviors for young students with autism. The last research stream identified in the isolated components is “technology as learning companion.” Ryokai et al. (2003) created a virtual peer to interact with children in a storytelling activity to facilitate their literacy learning. Luckin, Connolly, Plowman, and Airey (2003) employed digital toys and the associated software and investigated children’s interactions with the technologies. These two streams are the least active research community of the YCLT studies.

**Central papers in the co-citation network**

In addition to the angle of classification for the research network of YCLT, three measures of centrality were used to identify the most central papers in the network. Accordingly, the most central papers in terms of the role of centrality are the works of Chera and Wood (2003), Levy (2009), Lonigan et al. (2003), Magnan and Ecalle (2006), Segers et al. (2004), and Segers and Verhoeven (2005).

Among the six central papers, the most critical paper in the overall YCLT research network is Levy (2009). This paper has the highest degree of centrality and betweenness centrality showing that it not only connects to the largest number of YCLT studies, but also plays an important role of broker in the network. Besides, its closeness centrality is also listed as the second highest, indicating its close relationship with other studies of the YCLT research. In Figure 1, Levy’s work is connected to 12 studies including 7 in “digital literacy” and 5 in “technology-assisted language learning.” This suggests the close connection between Levy’s paper and the two main research streams.
Levy’s (2009) role of broker in two research streams can be further revealed from the research emphasis of the study. In Levy’s work, she investigated the influences of how schools teach children to read print texts on their ability to read multimodal texts. She found that children had developed the ability to read multimodal texts (digital literacy) at home by holistic understanding rather than by decoding the texts. However, when they learned decoding texts in school, their confidence in reading multimodal texts diminished. While researchers have emphasized the effectiveness of using computer programs or e-books to promote children’s particular language abilities, Levy’s study started a new direction for exploring the association between children’s technology use and their development of literacy and digital literacy.

Other than Levy’s (2009) work, the other five papers are core papers in the stream of “technology-assisted language learning.” These studies are targeted mainly on the facilitation of phonological skills and vocabulary learning. Two central works on “e-book” studies were conducted by Chera and Wood (2003) and Segers et al. (2004). These two studies compared the effects of using e-books to promote language learning with general instruction activities or teacher-mediated storybook reading. The other three studies implemented computer programs (including computer games) to enhance the phonological skills of children with reading disadvantages (e.g., Lonigan et al., 2003). Moreover, the nodes of Lonigan et al. (2003) and Chera and Wood (2003) adopted technology to improve children’s phonological awareness, and seem to represent the beginning of two sub-streams of research in language learning.

Discussion

In this section, we begin with a discussion of the two major findings in terms of the five major co-citation pairs and of the overall YCLT co-citation network. In addition, a comparison between two different methods of content and co-citation network analysis is also provided. Some methodological remarks regarding the use of co-citation network analysis are also suggested for future review studies.

Five major co-citation pairs in the YCLT research network: exploration and extension

As seen in Figure 1, five major pairs of research in the YCLT research network were identified. Among them, the highest co-cited studies (i.e., Shimizu & McDonough, 2006; Donker & Reitsma, 2007) are related to the use of the computer mouse by young children. To understand how this topic has developed over time, we further examined 14 studies which co-cited this research pair. Most studies (13 out of 14) were from Shih and his colleagues (e.g., Shih, Shih, & Peng, 2011). Their series of studies focused on evaluating disabled persons’ pointing performances (including children and adults with disabilities) after using various evolving programs which were developed to improve their pointing skill. Their studies also found that young children’s use of the computer mouse is related to the development of digital literacy (some gray links tied with the areas of technology competence and digital literacy, Figure 1). For future study, researchers can conduct studies connecting digital literacy with other developmental domains for young children. For example, how the design of touch screens influences young children’s scientific learning.

The second highly cross-referenced research pair (Clements & Sarama, 2007; 2008) is focused on the evaluation of a mathematics curriculum, Building Blocks. After examining the pairs of follow-up references, we found that 5 out of 10 follow-up papers modified the original curriculum and explored the relationships between the performances of mathematics and literacy as well as oral language (Clements & Sarama, 2011a; Clements et al., 2011b; Clements et al., 2013; Sarama, Clements, Wolfe, & Spitler, 2012; Sarama, Lange, Clements, & Wolfe, 2012). However, only Clements, Sarama, and their colleagues have worked on the important topic regarding the role of technology-enhanced activities in children’s learning of mathematics. Therefore, we suggest that future research can compare the effectiveness of (1) mathematics computer programs and off-computer activities, (2) mathematics computer programs developed on the basis of different pedagogical approaches, and (3) different types of technology (e.g., augmented reality, robots, computer games, electronic books, and online learning) in young children’s learning of mathematics. Last, the topic of the relationship between young children’s competencies of mathematics and other developmental domains or subjects (e.g., social development) needs further investigation.

The remaining three highly co-cited pairs in the main component of the network are tied with each other as a triangle (i.e., De Jong & Bus, 2004; Chera & Wood, 2003; Segers et al., 2004. see Figure 1). Their research topic is the
effectiveness of electronic books in terms of young children’s literacy development. Our co-citation analysis shows that 13 follow-up studies of the three pairs focused on the same issue among which eight studies, conducted by the research team of Korat, Shamir, and their colleagues, thoroughly explored the effects of e-books on emergent literacy. The remaining co-citing references include four empirical studies and one review paper. The topics of these articles all focus on how technology helps young children’s literacy development. A close analysis of the empirical studies revealed that young children’s e-book reading has been extended in three directions. First, Krčmar and Cingel (2014) turned their attention to adults’ reading of storybooks, and examined the difference in young children’s literacy growth between the reading of e-books and printed books. Second, Gong and Levy (2009) designed and implemented two kinds of e-books: a ball bouncing above a word when the word was read to increase children’s awareness of the sound of each word, and unreadable words being inserted into sentences to draw children’s attention to the components of readable words. Third, other types of technology rather than e-books were used to promote young children’s emergent literacy, such as computer programs (Van der Kooy-Hofland, Bus, & Roskos, 2012) and videos (Strouse, O’Doherty, & Troseth, 2013). In conclusion, five research themes have emerged as future research directions, including: (1) the role of adults in children’s reading of e-books, (2) the role of peers in co-reading e-books, (3) the effective designs of e-books, (4) the effects of other types of technology on emergent literacy, and (5) the literacy gains of reading e-books for children of various developmental levels.

The co-citation network of 53 YCLT empirical studies: overview

Early childhood education is multidisciplinary and interdisciplinary. However, from the co-citation network, it is apparent that the YCLT research is dominated by studies of language learning and digital literacy. Other learning domains (e.g., math and social interaction) have, however, received far less attention from researchers and been isolated from the main streams (i.e., language learning and digital literacy). A gap seems to exist between the main research streams and the other smaller ones. In addition, it is apparent that the current YCLT research is domain specific rather than domain integrated. Technology has opened a way to integrate learning from different aspects. To bridge the gaps in the YCLT research, we suggest that more studies can adopt various technologies (e.g., augmented reality, robots, and interactive whiteboards) and utilize them to integrate different learning domains.

In addition, the co-citation network of YCLT suggests that digital literacy appears as an important emerging stream. Our results also reveal a close connection between “digital literacy” and “technology-assisted language learning” in the YCLT network. This implies that researchers might consider digital literacy as an extension or an alternative form of literacy. “Technology competence” is naturally considered part of digital literacy; however, it is almost isolated from the network of “digital literacy.” This might be due to the fact that the advances in technology (e.g., touch screen devices) have made the computer mouse less important. This notion is also confirmed by the above analysis of the follow-up citations.

Overall, digital literacy plays an important role in technology-assisted learning. Based on the results of the co-citation network, digital literacy is expected to have a close relationship with all of the other learning domains. Therefore, future research may explore how digital literacy is related to other learning domains for young children, and whether digital literacy can be assessed as general abilities required for learning or whether it can be demonstrated only through content or domain-dependent tasks. Compared with Hsin et al.’s (2014) conceptual analysis, ours provides further bibliographic evidence by co-citation network analysis to exhibit the current developments of digital literacy in the YCLT literature.

The use of co-citation network analysis for literature reviews

Content analysis has long been used as a typical method of literature review in education research. Most coding schemes for categorization used in content analysis are based on researchers’ deductive observations or on a priori framework of past studies. For example, Hsin et al. (2014) proposed a typology based on the content analysis of five categories, and indicated some key factors influencing children’s learning through technology. In this present study, we take a distinct viewpoint by integrating co-citation analysis and social network analysis to examine the YCLT literature using their cited references.
First, the co-citation analysis provides abundant citation-based evidence to determine the closeness and relationships among the core papers. Based on the view of a large number of citers, therefore, the results can identify some bibliographically-related documents and group them into individual research streams. Second, the technique of networking visualization presents a clear structural overview of the literature, delivering a fresher review methodology. Some central cores and most referenced research sub-fields can be directly observed in the research network. Overall, the integrative approach of co-citation network analysis in this study offers researchers a relational overview of the core papers, and a structural understanding of the YCLT research network which reflects a field’s view of itself. This study is especially significant because, for instance, based on the results, researchers who are very new to the YCLT area can start by reading some focal research papers. On the other hand, senior scholars may conduct some interdisciplinary studies to connect different research topics.

For future studies, researchers can use co-citation network analysis to explore the relationships among other research topics in education or in any scientific research areas. Some statistical methods (e.g., factor analysis or cluster analysis) are suggested to integrate with co-citation analysis to quantitatively classify documents into different research groups. Furthermore, other co-citation analyses (e.g., author co-citation or journal co-citation analyses) can be adopted to provide different views of a discipline.

Conclusion and limitations

Overall, this paper contributes to the literature in a number of ways: it provides a fresh approach to research review and presents a critical content-based discussion based on the results of network analysis. The application of co-citation network analysis also provides a new research model for exploring research domains in other scientific disciplines.

Two limitations are, however, noted. First, the research network presented in this study is a citation-based result. Although highly co-cited pairs are endorsed by large citation data, some false citations in the publication may not be detected by co-citation network analysis. A further extensive content analysis is suggested to examine in more depth the relationships of citation and co-citation among the core papers. In addition, only journal papers were included in this study. Future research may consider other sources, such as conference proceeding papers, to quickly track the contemporary research trends. Last, the co-citation network is dynamic as the citation data changes daily. Researchers may include new citation data and newly published papers to analyze the YCLT research network. Bearing these limitations in mind, this study presents the overall intellectual structure of YCLT for further scholarly discussion.

Acknowledgments

This research is partially supported by the Ministry of Science and Technology, Taiwan, under Grant number MOST 101-2511-S-011-003-MY3, 104-2517-S-011-001-MY3, and MOST 104-2511-S-011-004-MY3.

References

The core papers are preceded by an asterisk (*).


A Decision Tool for Selecting a Sustainable Learning Technology Intervention

Maryam Raji1* and Imran Zualkernan2

1Engineering Systems and Management, American University of Sharjah, Sharjah, United Arab Emirates //
2Computer Engineering, American University of Sharjah, Sharjah, United Arab Emirates //
maryam.raji@alumni.aus.edu // izualkernan@aus.edu

*Corresponding author

(Submitted May 20, 2015; Revised August 24, 2015; Accepted October 15, 2015)

ABSTRACT

Education is a basic human right. In pursuit of this right, governments in developing countries and their donors often invest scarce resources in educational initiatives that are sometimes not sustainable. This paper addresses the problem of selecting a sustainable learning technology intervention (LTI) for a typical developing country. By solving this problem, more sustainable LTIs can be selected for implementation; this will in turn improve the efficiency with which limited resources are utilised, thereby allowing greater access to education for all. The paper introduces a unique decision framework for project selection by combining “soft” and “hard” decision analysis techniques. The selection problem is modelled as a multi-criteria decision-making (MCDM) problem and a combination of the Future Search Conference technique and the Analytic Network Process (ANP) is used to develop a decision tool for selecting the most sustainable LTI for a developing country. Our analysis revealed that of the nine LTIs considered, school on wheels was the most sustainable LTI in the collective opinion of all the experts involved in this study.

Keywords

Education, Learning technology interventions, Sustainability, Multi-criteria decision-making, Analytic network process

Introduction

Any nation’s development is related to the literate ability of its people (UNESCO, 2014). Nations recognise that an educated citizenry is better equipped to develop and advance its society. Therefore, nations invest heavily in education and learning projects. Despite these investments, global illiteracy remains high. The United Nations Educational Scientific and Cultural Organisation (UNESCO, 2014) has linked poor literacy to bad educational initiatives, reporting that one in four young people living in developing countries lack basic literacy skills.

In this paper, we address the issue of how a typical developing country can approach the problem of selecting a sustainable learning technology intervention (LTI). Solving this problem will allow more sustainable LTIs to be selected for implementation, increase the efficiency with which scarce resources are utilised and eventually improve access to education for all.

Learning interventions are usually assessed independently based on their individual benefits and drawbacks in specific circumstances. Attewell (2005) for instance, reported on a mobile learning (m-learning) project with disadvantaged European youth while other studies by Arbaugh and Duray (2002); Begičević, Divjak, and Hunjak (2007); Packham, Jones, Miller, and Brychan (2004); and Wu and Chen (2013) all assessed electronic learning (e-learning) implementation in higher education. Other studies reported include Norbrook and Scott’s (2003) study into using mobile phones for foreign language learning and Kim’s (2009) research into m-learning for underserved populations in Latin America. While these earlier studies all explored individual learning interventions in particular contexts, our study is the first of its kind to develop a comprehensive framework for assessing a collection of LTIs and selecting the most sustainable.

We begin the paper by defining key terms, then review the literature on nine LTIs and their applications. We further explain the methodology used, describing in detail the future search conference and the analytical network process and their bases in the literature. We then build the decision tool and apply it to select a sustainable LTI. We conclude the paper by mentioning limitations of the study and suggesting directions for future research.
Background

Sustainable development

“Sustainable” traditionally refers to something that can continue long into the future. Where educational interventions are concerned, a sustainable LTI is one that works well enough in practice to justify its continued implementation (Sterling, 2001). This implies that the LTI not only enhances learning immediately, but also well into the future. Assessing LTIs using the sustainability standard therefore requires that both present and future benefits are fully considered.

Learning technology interventions

The Association for Learning Technology (http://www.alt.ac.uk/) defines learning technology as the “broad range of communication, information and related technologies that can be used to support learning, teaching, and assessment.” This comprehensive definition encompasses a wide variety of technologies from e-learning and m-learning to educational radio and television programmes.

We conducted a literature search to reveal several LTIs that have facilitated learning in different contexts. We then selected nine LTIs from this list. The nine LTIs are: educational radio, educational television, mobile phones and applications, ubiquitous learning, school on wheels, one-laptop or tablet personal computer (PC) per child, massive open online courses (MOOCs), Internet cafés and remote online tutoring. What follows is a précis of the literature on these LTIs.

Educational radio

Radio is an effective medium for communicating with large audiences, even recognised by UNESCO (2006) as having significant potential for literacy development. The United States Aid Agency (USAID) and UNESCO have both taken advantage of the expansiveness and relative low-cost of the radio to broadcast educational programmes for children in developing countries across the globe including Malawi, Honduras, Haiti, Pakistan and Nigeria (UNESCO, 2009; VOA, 2009). Of the nine LTIs considered in this study, radio is unique in its ability to reach displaced people, communities living in remote areas and nomads with no permanent abode (Perkins, 2011; UNESCO, 2006; UNESCO, 2009; USAID, 2011).

Educational television

Television, like radio has the potential to reach large audiences but has the further advantage of being able to convey additional (visual) forms of information. Cognitive research conducted for Cisco Systems by the Metiri Group (Fadel & Lemke, 2006) indicates that when used as a learning aid, television not only has the power of increased visual, cognitive and emotional engagement, but can also result in improved academic performance when coupled with the right pedagogy. Earlier research by Ellis and Mathis (1985) and later comprehensive studies of educational television as an instructional tool by Wetzel, Radtke and Stern (1994) conclude that despite the instructor-learner distance, educational television is at least as effective as face-to-face instruction. Consequently, educational television has been successfully implemented for literacy education in India (Kothari, 2008), language instruction in Mexico, and also for teaching subjects like literacy, mathematics, science and life skills in other parts of the world (Anderson, 1998; Mares, Cantor, & Steinbach, 1999; Power, 2002; Willoughby, 1986).

Mobile phones

Mobile learning (m-learning for short) involves the use of mobile phones and applications to disseminate educational material for learning purposes. The educational material runs either on a standard mobile phone, or on a dedicated mobile learning device (Kim, Miranda, & Olaciregui, 2008). More recent surveys by Rushby (2010; 2011) signal that m-learning is highly favoured by learning technology educators. M-learning’s growing popularity among


educators can be attributed to several factors, such as the preponderance of mobile devices, and their unique characteristics of portability, versatility, collaboration, relative affordability and broad reach, which all combine to give m-learning a distinctive edge in the 21st century (Attewell, 2005; Papanikolaou & Mavromoustakos, 2006; Roschelle, 2003). Mobile phones are also increasingly popular for language learning (Kennedy & Levy, 2008; Norbrook & Scott, 2003; Stockwell, 2007).

**Ubiquitous learning**

Ubiquitous learning (u-learning) is a form of learning that employs ubiquitous computing technologies to facilitate learning anywhere and at any time (Sakamura & Koshizuka, 2005). It uses groups of computing devices (usually microchips) that sense from the environment and exchange the information with other sensors in the network (Sakamura, 2003). This arrangement allows the device to sense and supply information to learners instantaneously. Working in this non-obtrusive fashion, u-learning allows students to become fully immersed in, and active participants of the learning process (Jones & Jo, 2004). It also has advantages for teachers by allowing them greater understanding of the unique characteristics of each learner (Kinshuk & Huang, 2015). U-learning displays high levels of mobility and embeddedness (storage memory) (Lyytinen & Yoo, 2002) and is uniquely adaptable to the specific learning needs of students (Boyinbode & Akintola, 2008; Jones & Jo, 2004). For these and other reasons, students display a preference for this “real-world” form of learning (Hwang, Yang, Tsai, & Yang, 2009). U-learning has been applied to facilitate experimental learning in laboratories (Hwang et al., 2009), for language learning inside and outside classrooms (Huang, Yang, & Hwang, 2010; Ogata & Yano, 2004), and even for game-based learning (Chang, Wang, Lin, & Yang, 2009; Zualkernan, 2011).

**School on wheels**

The mobile school (or school on wheels) transports learning to students at their locations, instead of requiring them to congregate at conventional schools. It is especially useful where populations are sparsely distributed over wide areas, thus making the task of providing conventional schools to serve everyone an economically unfeasible option. The mobile school could also complement traditional schools by providing lessons and materials in subjects such as information and communications technology (ICT) or laboratory science where expenses prevent individual schools from having their own fully equipped facilities. Instances of school on wheels are the *ilm-on-wheels* project in Pakistan (Tele Taleem, 2015) and the *science-on-wheels* project in the United States (Carl, Don, & Constance, 2001).

The mobile school has the advantage of being more accessible to greater numbers of students over wider geographical regions than stationary conventional schools (Carl et al., 2001; Wood, 1965). Further, mobile schools are also a useful starting point for communities with no previous educational history or system, such as those emerging from the throes of conflict (Wood, 1965).

**One laptop or tablet PC per child**

One laptop-per-child (OLPC) is a nascent educational initiative where underprivileged children receive a laptop each to allow them educate themselves. These laptops may be used by the children working alone or, alternatively, assist as part of instructor-led teaching in schools. The laptops are designed to be low-cost, low-power, rugged portable computers which can run a variety of educational programmes. The OLPC initiative has been implemented in several countries across the globe, from African countries such as Rwanda and Nigeria (Fajebe, Best, & Smyth, 2013; Roberts & Zamora, 2012) to South American countries like Portugal, Peru and Uruguay (Kraemer, Dedrick, & Sharma, 2009; Warschauer & Ames, 2010).

**Massive Open Online Courses (MOOCs)**

The acronym “MOOC” was first coined to describe “an online course with the option of free and open registration, a publicly shared curriculum, and open-ended outcomes” (McAuley, Cormier, Siemens, & Stewart, 2010, p. 10). MOOCs are typically free, online, non-credit educational courses which are accessible anywhere by anyone with an
Internet connection (Daniel, 2012; Vardi, 2012). Examples of MOOC platforms on the Internet are EdX (http://www.edx.org), Udacity (http://www.udacity.com) and Coursera (http://www.coursera.org). Despite being most widely used in higher education, MOOCs also have potential for use by primary schoolchildren as proved by a recent online introductory engineering course by Brown University (Canvas Network, 2015; MOOCs Directory, 2015). Likewise, organisations like Next Vista for Learning (http://www.nextvista.org/), Khan Academy (https://www.khanacademy.org/) and MIT K-12 (http://k12videos.mit.edu/) all provide videos, lessons and quizzes covering an extensive range of subjects from mathematics and science to economics and history; all accessible by pupils with an Internet connection.

Internet cafés

Internet cafés (or cybercafés) are public places where people pay to access the Internet and/or perform other IT related activities on computers (Haseloff, 2005). Cybercafés are popular in developing countries as they allow of using a computer and connecting to the web to people who would otherwise be unable to bear those costs individually (Furuholt & Kristiansen, 2007; Haseloff, 2005). The youth typically make up the highest proportion of cybercafé clientele (Haseloff, 2005). A study of high school students in Turkey showed that young people use cybercafés to learn a variety of subjects and gain new skills in languages, ICT and even in social and health issues (Cilesiz, 2009).

Remote online tutoring

Online tutoring typically involves a teacher giving lessons to one or more students over the Internet or other network connection. Both parties typically use an Internet enabled device (for example, a computer, tablet PC or smartphone) and the lessons can occur either synchronously or asynchronously (Balakrishnan & Duraiswamy, 2011). Online tutoring offers the unique advantages of novelty, convenience and time flexibility. It has also been shown to improve reading skills based on a study of poor performing fourth-grade students in the United States (Vasquez & Slocum, 2012). Examples of online tutoring companies are NetTutor (http://www.nettutor.com/) and Tutorvista (http://www.tutorvista.com/) that offer tutoring in subjects like Mathematics, English and Science.

The future search conference

The future search conference is a group and community consensus building technique used for collective brainstorming (Schindler-Rainman & Lippitt, 1980). The brainstorming results in a collective mind map for a future that can be ranked by the community members by assigning priorities to various nodes of the mind map. Sellnow (2006, p. 4) documents five main steps in conducting a future search:

- Focusing on the past to highlight the main problems currently being faced.
- Shifting to an external-internal focus on the present.
- Seeking out future direction.
- Identifying common ground shared by all conference participants.
- Seeking out consensus on schemes for development.

<table>
<thead>
<tr>
<th>Icon</th>
<th>Colour</th>
<th>Points</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>-1</td>
<td></td>
<td>Lowest priority/Do not implement</td>
</tr>
<tr>
<td>Blue</td>
<td>0 (tiebreaker)</td>
<td></td>
<td>No rating/Not practical</td>
</tr>
<tr>
<td>Yellow</td>
<td>0.5</td>
<td></td>
<td>Medium priority/Maybe implement</td>
</tr>
<tr>
<td>Green</td>
<td>+1</td>
<td></td>
<td>Highest priority/Definitely implement</td>
</tr>
</tbody>
</table>

The future search conference method enjoys wide applicability across fields, from urban city planning (Sellnow, 2006) to manufacturing product redesign (Weisbord & Janoff, 2005). Several adaptations of the future search conference method exist, and this study used one such adaptation in the form of mind maps. Here, each decision
maker received four coloured dots to assign to each branch of a mind map, with the dots colour-coded as in Table 1. The points from each dot were summed and the branch with the highest points emerged as the most important.

The Analytic Network Process

The Analytic Network Process (ANP) is a multi-criteria decision-making (MCDM) method proposed by Thomas Saaty (Saaty, 1987; Saaty & Takizawa, 1986) for solving complex decision-making problems. The ANP is a generalisation of the Analytic Hierarchy Process (AHP) that structures a decision-making problem as a hierarchy; placing the goal at the top level, the criteria (and sub-criteria) below, and the decision alternatives at the bottom level. The AHP requires experts to compare alternatives pairwise (two at a time) with respect to each criterion and then rank those alternatives according to which of them best satisfies the goal.

When a decision problem involves interactions and dependencies between different-level elements of the hierarchy, it cannot be organised with the AHP; rather, the ANP must be used. The ANP structures the problem as a network without assuming independence between nodes (Saaty, 2008; Saaty, 2009). To determine the best alternative using the ANP, two sets of questions are posed:

- For each criterion, which one of a pair of alternatives is more dominant?
- For each alternative, which one of a pair of criteria is more dominant?

The AHP/ANP selection technique has several advantages (Saaty & Vargas, 2006): it can consider non-financial factors, put a figure on intangible human judgements and eventually decode the resulting mathematical model into easily comprehensible, non-technical results. Its major limitations according to Chan et al. (2008) are: it limits the size of the model (since a larger model requires additional comparisons which may lead to panel fatigue), it requires decision makers to have expert knowledge and experience of the topic, and it also requires them to be able to reason logically and make rational judgements.

Methodology

The problem of how to select a sustainable learning technology intervention (LTI) from many alternatives can be represented as a multi-criteria decision-making (MCDM) problem. This means the final decision of which LTI to select is influenced by several criteria such as benefits, opportunities, costs (required resources) and risks (BOCR) which, depending on their relative importance with respect to the goal, favour selecting a particular LTI.

We conducted this study in two stages: Stage 1 was the model identification stage, while Stage 2 was the building of the model. Stage 1 involved identifying and selecting alternatives and criteria to be used as inputs to the ANP model. This was achieved through conducting a literature search that identified nine LTI alternatives and several BOCR criteria for the model. We presented these data to a panel of 12 experts comprising government officials, CEOs of NGOs, professors, teachers, a school principal, and a representative of The World Bank. The panel used the future search conference to reduce the model elements to five LTI alternatives and three parameters per BOCR criterion. Figure 1 shows the priorities assigned to each LTI alternative by the panel.

After assigning priorities to each of the nine LTI alternatives, we selected the five alternatives with the highest priorities to advance to the next stage of the study. School on wheels was the most popular LTI among the experts; Educational Television came fifth place while Internet cafés received the lowest rating possibly because of the recently increasing prevalence of Internet in homes. The top five LTIs are arranged in order of priority, from highest to lowest in Table 2.

After ranking the LTIs, the panel ordered each of the BOCR parameters, lessening them to a more practical number for the ANP stage. The mind map branches shown in Figures 2 through Figure 5 show the ranking process for the BOCR parameters, while Table 3 displays the final selection of three parameters per BOCR criterion.
Starting with the benefits node of the network, the panel of experts prioritised the parameters as in Figure 2. They voted access to better teachers as the most important benefit criterion by assigning it seven green dots (corresponding to a score of +7).

Figure 2. Ranking of benefits parameters

Figure 3 pictures the rankings of the parameters on the opportunities node of the network. Access to teachers/learning resources was the highest rated opportunities parameter with six green dots (a score of +6).
Figure 3. Ranking of opportunities parameters

Figure 4 displays the prioritisation of the parameters on the resources node of the network. The experts considered *affordability* the most important resources parameter and assigned it five green dots and one yellow dot resulting in an effective ranking of +5.5.

Figure 4. Ranking of resources parameters

Figure 5. Ranking of risks parameters
The panel ranked parameters on the risks node of the network as in Figure 5. They considered Teacher non-adoption the most important risk factor and assigned it seven green dots (a score of +7).

After exploring each of the four branches of the mind map (corresponding to benefits, opportunities, costs/resources, and risks respectively), we collated the three highest rated parameters per branch in Table 3. The table shows the total 12 selected parameters used to assess the LTI alternatives in the second stage of the study.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Sub-criteria/Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefits</td>
<td>Access to better teachers</td>
</tr>
<tr>
<td></td>
<td>Improved learning outcomes</td>
</tr>
<tr>
<td></td>
<td>Increased productivity</td>
</tr>
<tr>
<td>Opportunities</td>
<td>Drive economic development for companies that make LTIs</td>
</tr>
<tr>
<td></td>
<td>Eradicate poverty</td>
</tr>
<tr>
<td>Costs/Resources</td>
<td>Affordability</td>
</tr>
<tr>
<td></td>
<td>Teacher training/competence</td>
</tr>
<tr>
<td></td>
<td>Access to power/electricity</td>
</tr>
<tr>
<td>Risks</td>
<td>Teacher non-adoption</td>
</tr>
<tr>
<td></td>
<td>Not addressing end-user requirements</td>
</tr>
<tr>
<td></td>
<td>Religion and taboo</td>
</tr>
</tbody>
</table>

After identifying the most important BOCR parameters and LTI alternatives through the future search conference in Stage 1 of the study; the study moved into Stage 2, which entailed combining these elements into an ANP model to identify the most sustainable LTI.

Stage 2 of the study was the ANP stage which used the outputs from the future search conference to construct a network. A new panel of eight experts then made pairwise comparisons between the nodes of the network. The panel comprised two teachers, two CEOs of NGOs, three learning technology experts and one government administrator. Each expert made three sets of pairwise comparisons. They rated successive pairs of:

- Alternatives with respect to each BOCR parameter. For example, school on wheels was rated against mobile phones as regards which provides greater access to better teachers.
- Parameters with respect to each alternative. For instance, access to better teachers was compared against improved learning outcomes to see which was a more pronounced benefit of school on wheels.
- Criteria with respect to the overall goal of selecting a sustainable LTI. For example, benefits was compared against opportunities and whichever the expert considered to be more important in selecting a sustainable LTI received a higher weighting.

We used the geometric mean to combine the judgements of all eight experts to produce a single collective group rating. The geometric mean is a widely used technique for synthesising group decisions into an ANP model. In its simplest form, the geometric mean of a group of \( n \) numbers is found by taking the \( n \)th root of the product of \( n \) numbers (Mitchell, 2004).

\[
G(n) = \sqrt[n]{x_1 x_2 x_3 x_4 \cdots x_n} 
\]

The geometric mean is relatively easy to implement in software, making it a popular method for combining multiple judgements (Dyer & Forman, 1992; Jharkharia & Shankar, 2007; Wu, Lin, & Lee, 2010; Wu, 2008). After combining the multiple judgements into one collective judgement, we combined this single judgement using the additive formula in the Superdecisions ANP software (Saaty, 2003). The additive formula is best for long-term results and takes the form:

\[
(benefits + opportunities) - (costs + risks)
\]

Next, we conducted sensitivity analysis by disturbing the weights of each BOCR criterion in turn and recording the effect of this disturbance on the rankings of the alternatives.
Model implementation and results

Figure 6 illustrates the weightings of each alternative according to the collective opinion of the panel. The results suggest that school on wheels is the most sustainable LTI overall, in the collective opinion of all experts involved in this study. Educational TV received the lowest ranking and is therefore the least sustainable LTI.

![Figure 6. Experts' rankings of LTI alternatives](image)

We also recorded the consistency ratios (CRs) for each set of pairwise comparisons. A CR below 10% is preferred and a higher CR may indicate the experts need to revise their judgements (Saaty & Vargas, 2006). The CRs for the criteria comparisons with respect to the goal was 0.01, which is consistent. Figure 7 shows the CRs for alternatives.
compared with respect to each parameter while Figure 8 shows the CRs for parameters compared with respect to each alternative.

Analysing the consistency ratios reveals that 5 out of 12 of the CRs by parameter were inconsistent while precisely 7 out of 20 of the CRs by alternative were inconsistent. An improvement of this inconsistency is desirable, particularly for those comparisons with CRs above 0.2.

The original model had equal weightings for each of the BOCR criteria; i.e., benefits, opportunities, resources and risks were each weighted 0.25. However, it is also desirable to explore the relative impact of each BOCR criterion on the results of the study. To do this, we sequentially disturbed one criterion (one of benefits, opportunities, resources or risks) to a high, medium or low value while preserving the three remaining criteria at a constant value. We recorded the impact of this disturbance on the rankings of the LTI alternatives; then used the additive formula to synthesise the results from the sensitivity analysis as before. Table 4 displays the sensitivity analysis on the benefits node and reveals that rankings of all alternatives are stable to perturbations of the benefits node.

![Figure 8. Consistency ratios by alternative](image)

<table>
<thead>
<tr>
<th>Alternative</th>
<th>Reference Ranking</th>
<th>Benefits (High: 0.6)</th>
<th>Opportunities (Medium: 0.25)</th>
<th>Resources (Low: 0.1)</th>
<th>Risks</th>
</tr>
</thead>
<tbody>
<tr>
<td>School on wheels</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mobile phones</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Online tutoring</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Educational TV</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>MOOCs</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4. Benefits node sensitivity analysis
Table 5 shows the sensitivity analysis on the opportunities node and reveals that the weighting of the benefits node does not affect the rankings of the alternatives.

Table 5. Opportunities node sensitivity analysis

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>High (0.6)</th>
<th>Medium (0.25)</th>
<th>Low (0.1)</th>
<th>Reference (original ranking)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative</td>
<td>Ranking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School on wheels</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mobile phones</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Online tutoring</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Educational TV</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>MOOCs</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 6 displays the sensitivity analysis on the resources node and reveals that a high weighting on the resources node significantly alters the rankings of the alternatives. With a high weighting, school on wheels drops from first to third place because it has the second highest resource requirement of all the alternatives. Similarly, mobile phones drops from third to fifth place as it has the greatest resource requirement of all the alternatives. On the other hand, MOOCs moves from fourth to second place when the resource node carries a high weighting as it has the lowest resource requirement of all the alternatives.

Table 6. Resources node sensitivity analysis

<table>
<thead>
<tr>
<th>Resources</th>
<th>High (0.6)</th>
<th>Medium (0.25)</th>
<th>Low (0.1)</th>
<th>Reference (original ranking)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative</td>
<td>Ranking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School on wheels</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mobile phones</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Online tutoring</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Educational TV</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>MOOCs</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 7 shows the sensitivity analysis on the risks node and reveals that only a high weighting on the risks node has any impact on the rankings of the alternatives. In this case, the positions of educational TV and MOOCs are interchanged, with the latter becoming the least desirable alternative as it has the highest risks of all the alternatives.

Table 7. Risks node sensitivity analysis

<table>
<thead>
<tr>
<th>Risks</th>
<th>High (0.6)</th>
<th>Medium (0.25)</th>
<th>Low (0.1)</th>
<th>Reference (original ranking)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative</td>
<td>Ranking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School on wheels</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mobile phones</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Online tutoring</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Educational TV</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>MOOCs</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

To summarise, this study reveals school on wheels is the most sustainable LTI, it keeps this position except when the resources node is assigned a high importance. Also, Educational TV is the least sustainable LTI and remains so even during the sensitivity perturbations, except when either the resources or risks node is assigned a high priority.

Sensitivity analysis also reveals that the priorities of the benefits and opportunities nodes have no impact on the rank ordering of the alternatives. On the other hand, a high weighting on either resources or risks significantly alters the rankings of alternatives. The greatest change to ranking occurs when the resources node is assigned a high weighting as this alters the positions of all five LTI alternatives. A lesser impact occurs with a high weighting on the risks node as this merely swaps the positions of educational TV and MOOCs.
Summary

The goal of this study was to develop a unique decision-making instrument for selecting a sustainable learning technology intervention (LTI). This tool should aid stakeholders in improving access to education in the developing world. A literature survey conducted at the start of the study revealed the LTIs currently implemented worldwide; alongside the benefits, opportunities, costs and risks associated with applying them. The study presented a novel decision-making tool in the form of a two-stage model combining qualitative and quantitative techniques (the future search conference mind mapping and the analytic network process). By applying the instrument to address the problem of selecting a sustainable learning technology intervention (LTI) for a developing country, we find that school on wheels is the most sustainable LTI in the joint opinion of the panel of experts.

In this study, we developed the future search mind mapping and ANP tool and applied it to address one decision problem (selecting a sustainable LTI) in a specific developing country (Pakistan). Nonetheless, this tool is applicable to any community decision-making problem where expert judgement is required and accessible.

This study is limited by the single geographic location and the number of participating experts. We expect that engaging a wider pool of experts from several countries should improve the generalisability of the findings. It would also be desirable to conduct extra rounds of pairwise comparisons to ensure all consistency ratios fall below 0.1. Finally, the techniques used in this study are subjective by nature. Thus, we infer that the final decision represents the personal opinions of the experts as shaped by their knowledge, skills and experiences.

References


Online Metacognitive Strategies, Hypermedia Annotations, and Motivation on Hypertext Comprehension

Hui-Fang Shang
Department of Applied English, I-Shou University, Kaohsiung, Taiwan // hshang@isu.edu.tw

(Submitted May 24, 2015; Revised September 29, 2015; Accepted November 19, 2015)

ABSTRACT
This study examined the effect of online metacognitive strategies, hypermedia annotations, and motivation on reading comprehension in a Taiwanese hypertext environment. A path analysis model was proposed based on the assumption that if English as a foreign language learners frequently use online metacognitive strategies and hypermedia annotations, then they would increase their learning motivation, which, in turn, would enhance their hypertext comprehension performance. Participants consisted of 37 English majors enrolled at a private university in southern Taiwan. The results revealed that online metacognitive strategies most significantly predicted learning motivation and motivation directly affected hypertext comprehension. However, metacognitive strategies and hypermedia annotations did not directly affect hypertext comprehension. Consistent with expectations, motivation was found to mediate the relation between online metacognitive strategies and hypertext comprehension. On the other hand, the relation between hypermedia annotations and hypertext comprehension mediated by motivation was not obvious. Finally, the implications of these findings for future research are discussed and presented.

Keywords
Online metacognitive strategies, Hypermedia annotations, Motivation, EFL hypertext comprehension

Introduction
With the rapid growth of computer technology, computer-based instruction has become an instructional tool to motivate English as a foreign language (EFL) learners to read autonomously. Thus, some printed texts are designed as hypertexts for helping EFL learners to search for and comprehend materials in a timely manner using computer-assisted functions (Wang, Kinzie, McGuire & Pan, 2010). Hypertext often incorporates multimedia features into linked texts and presents information in various forms of media such as pictures, audio, video, animation, or graphic annotations. Thus, the terms of hypertext and hypermedia are often used interchangeably (Akyel & Erçetin, 2009).

Previous studies have demonstrated that the effective use of metacognitive strategies has been recognized as an important means to increase hypertext comprehension because learners can actively monitor their progress, determine where problems exist, adjust their learning strategies, and make decisions regarding which annotations they should access during reading (Akyel & Erçetin, 2009; Huang, Chern, & Lin, 2009; Lan, Lo, & Hsu, 2014). However, not all EFL learners have the essential knowledge and skills to master effective metacognitive reading strategies to relate to a text (Chen et al., 2011).

Individual differences regarding how often learners access various built-in annotation tools and whether these tools assist in the lower-level process of understanding individual words (and ultimately, the higher-level process of text comprehension) have been noted (Chun, 2001). In addition, some studies have shown that learners can control the use of annotation by choosing what to view as well as selecting the number of viewing times and the duration (Chang & Ho, 2009). According to Moos and Marroquin (2010), control gives learners the opportunity to make decisions and to affect reading outcomes, resulting in more confidence and greater intrinsic motivation. Tsai and Tsai (2003) found that students with high motivation had better information searching strategies and learned more than those with low motivation. Although it has been reported that disorientation is a motivational problem that learners tend to have while navigating within a hypertext environment and that metacognitive processing skills can determine a successful search on the Internet, the evidence of previous research is inconclusive with regard to the effects of metacognitive strategies and annotation applications on the development of EFL hypertext comprehension (Ariew & Erçetin, 2004). Moreover, few empirical studies have explored the possibility that motivation acts as a mediator variable with online metacognitive strategies, hypermedia annotations, and hypertext learning outcomes.
Given its importance in a learner-controlled hypertext environment, it is essential to understand the factors that influence the extent to which EFL learners engage in metacognitive strategy use, hypermedia annotations, and learning motivation on hypertext comprehension. Previous studies have not revealed whether metacognitive strategies and hypermedia annotations can influence learning motivation and whether these three factors can facilitate hypertext comprehension (Sakar & Ercetin, 2005). It is assumed that these three antecedents are likely to influence EFL students’ hypertext comprehension performance, and motivation is a mediator between metacognitive strategies, hypermedia annotations, and hypertext comprehension.

Literature review

Hypertext comprehension

Hypertext is defined as an analog to traditional reading environments in which documents are presented on a computer screen, and learners can decide which hyperlink to access during the hypertext learning process (Moos & Marroquin, 2010). Within a document, elements of the apparatus, such as definitions of terms, glossaries, annotations, and references can all be hyperlinked to provide readers with additional information for text comprehension (Chou, 2012). Numerous studies have investigated the ways in which hypertexts can be employed to enhance students’ reading comprehension. The results showed that students attain significantly higher comprehension scores by reading the hypertext version, and that they prefer this mode over the printed version (Chen et al., 2011; Huang, Chen, & Lin, 2009; Ray & Belden, 2007). However, few studies have reported the effect of hypertext reading on comprehension in the EFL context (Huang et al., 2009). Although hypertexts can offer many benefits, students may become lost in a “sea of information” since reading hypertexts often requires a more cognitively demanding mode of learning compared to reading printed texts (Nowak, 2008). Troffer’s (2000) research showed that reading from the screen is approximately 30% slower than reading from the printed text, and most hypertext readers tend to scan a document rather than read the text word-for-word. Liou (2004) also found that, even though students in Taiwan have good computer literacy, they still have difficulty coping with electronic literacy due to the lack of using appropriate metacognitive reading strategies to comprehend online English texts. Therefore, it is essential to investigate the impact of metacognitive strategy use on hypertext comprehension.

Online metacognitive strategies

In order to explore the relation of metacognitive strategy use and hypertext comprehension, considerable attention has been paid to understanding what proficient EFL readers typically do while reading hypertexts, including identifying the types of strategies and determining how they use them as well as under what conditions they are used (Anderson, 2003; Gene, 2011). A number of studies have shown that metacognitive strategy use can increase the comprehension of low-proficiency EFL readers’ to the level of skilled readers as well as improve reading comprehension and motivation towards hypertext reading (Lan et al., 2014).

Anderson (2003) investigated EFL and English as a second language (ESL) learners’ online metacognitive reading strategies through the Online Survey of Reading Strategies (OSORS), which focused on problem-solving and support strategies as well as local and global strategies. It was found that problem-solving strategies, such as adjusting reading rates, re-reading difficult texts, and pausing to think about what one is reading, were reported most frequently, whereas support strategies were identified the least. Konishi (2003) analyzed online reading strategies employed by Japanese ESL students while reading online texts and found that both local strategies (i.e., commenting on the meaning of the work, pronunciation or grammatical interpretation) and global strategies (i.e., making inferences and utilizing background information) were used in reading tasks. Huang et al. (2009) examined the effects of strategies used by 30 EFL students on comprehension and found that students used support strategies i.e., translating, using dictionaries or highlighting) much more frequently than any other strategy. In sum, the use of global strategies significantly contributed to better comprehension for low-proficiency learners.

Chang (2005) investigated whether strategy use facilitated web-based learning and found that metacognitive strategies made EFL students more responsible for their learning, thus enhancing their overall academic achievement. In the study, less-proficient learners showed the greatest improvements. Therefore, it is important for EFL instructors to consider how to structure a learning atmosphere in which thinking about what occurs during
hypertext reading can lead to better learning outcomes (Anderson, 2003). The majority of previous studies, however, solely relied on self-reported questionnaire data and did not explore the actual use of annotations for hypertext comprehension (Akyel & Erçetin, 2009).

**Hypermedia annotations**

An annotation is defined as a list of words and phrases of which their meanings are outside the learner’s current competence level. The purpose of an annotation is to provide lexical, syntactic or additional information about a topic in order to clarify text understanding (Ariew & Erçetin, 2004). As revealed by Chun (2001), hypertext readers have the freedom to progress through a text and choose their own order of annotation based on their preference. Previous studies have demonstrated that EFL readers have a strong preference for annotations that provide word definitions and first-language translations (Chun & Payne, 2004; Erçetin, 2003) or annotations that offer topic-level information (Sakar & Erçetin, 2005). Regarding the effectiveness of hypermedia annotations on reading comprehension, some studies have illustrated that annotated texts were more useful for comprehension than those without annotations (Davis, 1989), while others did not find a significant relation between annotation use and reading comprehension (Ariew & Erçetin, 2004; Li et al., 2014).

Chun (2001) conducted a study with 23 students divided into high- and low-ability groups. The results indicated that the learners looked up the definitions of more words when both an internal and an external dictionary were available and performed better on the recall task. However, no relation was found between the number of words that the learners looked up and the number of propositions recalled. Liou’s (2000) study investigated EFL learners’ strategies of word consultation on reading comprehension and found that advanced learners had more background information, they read faster, looked up fewer words, possessed more word guessing and inference strategies, and often chose to ignore some unknown words. Meanwhile, weak learners had to double-check some of the unknown words in order to comprehend the text. Erçetin (2003) investigated the annotation use of intermediate and advanced level students in terms of the frequency that they accessed the annotations. The results showed that intermediate students accessed all types of annotations more frequently. Moreover, both groups preferred accessing annotations that frequently provide extra topic-level information and spent more time on these annotation types than the annotations that offer word-level information.

Ariew and Erçetin (2004) conducted a study in order to determine whether different types of hypermedia annotations at the word and topic levels facilitated reading comprehension for intermediate and advanced ESL learners. The results indicated that annotations did not enhance reading comprehension for either group, while a negative relation was observed between the time spent on video and graphics annotations and reading comprehension for the intermediate group. Li et al. (2014) argued that learners’ motivation plays an important role during the process of online reading since it affects the success or failure of using a computer as an instructional tool. Thus, it is essential to explore learners’ motivation in the hypertext learning process.

**Motivation**

Motivation is defined as physiological processes involved in the vigor and persistence of behavior, and it has been suggested that motivational theories can include beliefs, values, and goals (Eccles & Wigfield, 2002). Alvermann (2002) addressed that it is essential for instructors to create opportunities for readers to manipulate information in the hypertext environment and engage learners in their own hypertext learning in order to enhance their motivation and academic performance. Hoffman (1998) investigated the effect and participants’ motivation of a hypertext learning environment on second-language reading comprehension. Most of the participants indicated that, although reading hypertexts was not an easy task, it was faster. More specifically, some of the participants positively perceived its richness, ease of access to information, and quickness of locating the target information. However, they also expressed some concerns regarding the issue of disorientation and the complexity of the hypertext environment, both of which decreased their learning motivation. DeStefano and LeFevre (2007) also confirmed that complex hypertext structural features did not reliably enable learning motivation and navigation.

Becker and Dwyer (1994) examined the impact of students’ intrinsic motivation for a learning task and found that learners who used the hypertext program were more self-determined and their motivation was higher than those who
used paper-based texts. Conversely, Cho (1995) revealed no effect of students’ motivation towards hypertext learning, whereas Gray (1988) found a negative effect between learners’ motivation and hypertext reading. Based on the findings of these aforementioned studies, the evidence is apparently inconclusive.

The present study, guided by the work of Murphy and Alexander (2000) and Hasan (2003), considered motivation under the following three constructs: (1) computer self-efficacy; (2) goal orientation; and (3) situational interest. Computer self-efficacy is the self-perception of one’s capabilities of using a computer to successfully perform a task (Compeau & Higgins, 1995). Goal orientation focuses on the desire to gain competence (or master a new set of skills) and perform better through minimal effort (Archer, 1994). Finally, situational interest represents a short-lived interest that results from interactions with specific aspects of the learning context (Alexander et al., 1997).

**Relationships among online metacognitive strategies, hypermedia annotations, motivation, and hypertext comprehension**

Numerous researchers have proposed that reading strategies in hypertext can indirectly affect comprehension by leading readers to process a particular text in terms of amount of information accessed (Salmerón et al., 2005). Research has also showed that learners with greater levels of metacognitive strategy use performed better on a performance measure and had higher levels of motivation (Schmidt & Ford, 2003). Azevedo et al. (2005) found that students who monitored their understanding by linking nodes of information generally develop a deeper understanding of challenging topics, compared to those who did not use this approach. Moos and Azevedo (2008) found that, as students made decisions about which annotation to access in order to solve their reading problems, they would actively participate in the learning process, and the extent to which they actively participated was related to motivational constructs. In particular, students’ participation in the learning process is associated with the perceptions of their capabilities, which can facilitate higher motivation and learning outcomes (Debowski, Wood, & Bandura, 2001). Scott and Schwartz (2007) further suggested that hypermedia requires the need to balance effective annotation application and content comprehension during hypertext reading. This need is best met through the use of metacognitive processes such as monitoring overall understanding, the feeling of knowing, and the progress towards the learning goal (Azevedo et al., 2005). However, research has also demonstrated that many learners have difficulty using such learning processes (Moos & Azevedo, 2008). For example, Moos and Azevedo (2009) suggested that learners need to have a requisite amount of prior knowledge in order to make decisions regarding which hyperlinks to access.

Although the inherent nature of hypertext reading is often thought to engage learners, research examining the relation between motivation and hypertext reading has resulted in mixed findings, especially since various factors can influence the effect of learner motivation with hypertext learning (Moos & Marroquin, 2010). Moreover, it is unclear whether the results can be generalized to EFL students in a Taiwanese learning context. Hence, it is important to understand the factors that may influence the extent to which EFL learners engage in hypertext reading activities.

As noted in the preceding section, the relations among online metacognitive strategy use, annotation application, and motivation on hypertext learning have resulted in inconclusive findings. Therefore, the purpose of this study was to explore the impact of the antecedents on hypertext comprehension. This study also presented a model of the hypothesized relations among online metacognitive strategies (i.e., local, global, problem-solving, and support strategies) adopted by Anderson (2003), hypermedia annotations (i.e., word glossaries, sentence translations, and background information) developed by Shang (2015), and motivation (i.e., computer self-efficacy, goal orientation, and situational interest) adapted from Murphy and Alexander (2000) and Hasan (2003). The following hypotheses (see Figure 1) were posited:

H1: There is a significant correlation between online metacognitive strategies and hypermedia annotations.
H2: Online metacognitive strategies positively affect motivation.
H3: Hypermedia annotations positively affect motivation.
H4: Motivation positively affects hypertext comprehension.
H5: (a) Online metacognitive strategies positively affect hypertext comprehension.
(b) Motivation mediates the relation between online metacognitive strategies and hypertext comprehension.
H6: (a) Hypermedia annotations positively affect hypertext comprehension.
(b) Motivation mediates the relation between hypermedia annotations and hypertext comprehension.
Figure 1. The hypothesized path model of online metacognitive strategies, hypermedia annotations, self-efficacy, and hypertext comprehension

Research methodology

Subjects

Thirty-seven English majors (10 males and 27 females), who were approximately intermediate learners of English, from a junior reading class in fall 2014 at a private university in southern Taiwan participated in this study. A demographic questionnaire including age, years of using a computer, hours of accessing a computer in a day, and self-evaluating level of computer knowledge was administered to gather information about the subjects’ backgrounds. Results from the questionnaire survey showed that the subjects of this study ranged in ages from 20 to 23 years old, with an average of 20.5 years old (SD = .804). More detailed demographic characteristics of the subjects are provided in Table 1.

Table 1. Sample characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>Female</td>
<td>27</td>
<td>73</td>
</tr>
<tr>
<td>Years of using a computer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-4 years</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5-10 years</td>
<td>20</td>
<td>56.8</td>
</tr>
<tr>
<td>11-15 years</td>
<td>15</td>
<td>40.5</td>
</tr>
<tr>
<td>16-20 years</td>
<td>1</td>
<td>2.7</td>
</tr>
<tr>
<td>Average</td>
<td>11.1 (SD = 2.61)</td>
<td></td>
</tr>
<tr>
<td>Hours of accessing a computer per day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-3 hours</td>
<td>23</td>
<td>63.9</td>
</tr>
<tr>
<td>3.5-4.5 hours</td>
<td>6</td>
<td>16.7</td>
</tr>
<tr>
<td>5-6 hours</td>
<td>7</td>
<td>19.4</td>
</tr>
<tr>
<td>Average</td>
<td>3.1 (SD = 1.48)</td>
<td></td>
</tr>
<tr>
<td>Self-evaluate level of computer knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beginner</td>
<td>3</td>
<td>8.1</td>
</tr>
<tr>
<td>Intermediate</td>
<td>27</td>
<td>73.0</td>
</tr>
<tr>
<td>Advanced</td>
<td>7</td>
<td>18.9</td>
</tr>
</tbody>
</table>
Instruments

A hypertext system

Seven printed texts, relating to daily-life topics, selected from the magazine “English Digest” were developed in a hypertext format, and annotations were added by the researcher with the aid of multiple forms of digital media. The developed hypermedia annotations provided information about the text with the learner’s native language at three levels: (a) word-level information, i.e., definition of words, (b) sentence-level information, i.e., translation of sentences, and (c) topic-level information, i.e., background information about the topic. The developed hypertext system (Shang, 2015) is composed of the following three main sections:

- The introductory section is to hyperlink encyclopedias for background information or knowledge of the world. Readers are free to select various topic-level information which is available through buttons placed outside the text and information can be viewed in the form of text, audio, and video.
- The text-based section is to direct students to link within and among texts. Within a document, elements of the annotations, such as vocabulary glossary with definitions and examples as well as sentence translations, are all hyperlinked to provide readers with word- and sentence-level information at the point of need to enhance their text understanding.
- The monitoring section is to assess learners’ understanding of the text by asking text-based questions. As the information in hypermedia environments is presented in multiple segments, connected via hyperlinks, readers have to recycle these same processes many times until they accomplish the reading goal.

The screenshots for the developed hypertext reading system are illustrated in the following figures (Figure 2, 3 and 4).

![Figure 2. A sample of magazine article (adopted from English Digest, Unit 4, September 2013)](image-url)

326
Reading comprehension tests

A 15-item multiple-choice test to assess understanding of the content knowledge was administered to the subjects after reading each hypertext. As one of the purposes of the present study was to examine whether annotations facilitated comprehension of the text, the questions on the test assessed the specific factual information such as main ideas and details provided in the text, not in the annotations. One score for each subject was calculated. The maximum achievable score for each test was 15 to estimate subjects’ reading comprehension. Seven test scores were collected after completing seven hypertext reading tasks. Subjects’ mean scores of the seven tests were calculated and served as the hypertext comprehension scores. To evaluate content validity, two English instructors from the
Department of English at the research site were invited to modify the test items to be more appropriate and content-based.

Questionnaires

An OSORS (Online Survey of Reading Strategies) adopted by Anderson (2003) was adapted and administered to measure participants’ use of metacognitive strategies while reading the hypertexts. The OSORS questionnaire survey involved 27 items measuring three broad categories of reading strategies: global strategies (10 items), problem-solving strategies (8 items), and support strategies (9 items). Global reading strategies are the carefully planned techniques by which learners monitor or manage their reading. Problem-solving strategies are the localized techniques used when problems develop in understanding textual information. Support strategies are the basic support mechanisms to aid the reader in comprehending the text (Mokhtari & Sheorey, 2002). A five-point Likert scale, indicating the frequency of hypertext strategy use, ranging from 1 (never do) and 5 (always do) was maintained in the OSORS. The Cronbach’s alpha for the overall OSORS was .901. The reported reliabilities for each subsection were global strategies, .773, problem-solving strategies, .748, and support strategies, .785.

The second instrument used for data collection was the motivation questionnaire adapted from Murphy and Alexander (2000) and Moos and Marroquin (2010). The motivation questionnaire, with a total of 17 items, measuring participants’ computer self-efficacy (4 items, α = .827), goal orientation (4 items, α = .746), and situational interest (9 items, α = .814), was administered to assess participants’ motivation for doing the hypertext activity. The Cronbach’s alpha for the overall motivation scale was .869. Individuals responded on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). These data help to establish that the OSORS and motivation questionnaires were reliable instruments for assessing EFL learners’ online metacognitive strategies and learning motivation while reading hypertexts. A pilot study was conducted by five students randomly chosen to ensure that each item was fully understandable.

Data collection procedures

This study was conducted within the context of a reading course that met once per week for 100-minute class periods throughout an 18-week semester in fall 2014. The reading course, Guided Reading (requirement/2 units) aimed at enhancing junior students’ general English reading comprehension. To ensure students’ full participation, they were informed that doing this hypertext reading activity was part of the course requirements. Participants were encouraged to choose and view as many annotations as provided for text understanding, and they were allowed to return to previously studied pages as often as they chose. The software tracked every interaction of the reader with the text, including which annotations the reader chose to view and the frequency of accessing a particular annotation. The data were all saved as a log file. Because the tracking tool was hidden, the collection of the data regarding the participants’ interaction with the text did not hinder the flow of reading (Sakar & Ercetin, 2005).

More specifically, participants received an orientation on metacognition by the instructor who was also the researcher, introducing the concept of metacognition and students were informed of the importance of monitoring their own knowledge as they proceeded reading. After a brief orientation on metacognition, participants were then instructed on how to use the hypertext system and asked to read the hypertext individually in front of a computer in the computer lab to display the document. To encourage students’ participation in doing this hypertext activity, they were informed that all their consultation with the annotations, such as linking vocabulary glossary for word-level information, translation for sentence-level information, and background for topic-level information, would all be stored in the computer database for further assessment.

The hypertext reading task was provided in two 50-min sessions, lasting for seven weeks. No attempts to teach reading comprehension were implemented during class periods, yet students were free to take notes and encouraged to ask questions at any time during the computer operation. Appropriate levels of assistance were provided by the instructor in a timely manner, so as to help solve any possible technical or operational problem. To prevent students from using the other unrelated resources without concentrating on the hypertext task, the instructor monitored students’ learning process all the time and modified the intervention accordingly. Participants would read the
expository hypertext until they felt satisfied that they could answer text-based questions to assess their reading comprehension. They were not allowed to refer to the text during the reading comprehension test.

After finishing the hypertext reading task and comprehension test, participants were requested to complete the OSORS and motivation questionnaires immediately within 30 minutes. The participants were instructed to read the cover letter to understand the purpose of the study, and they were informed that their answers on the questionnaires would not affect their scores in the course. The researcher would stay in the classroom to answer any questions from the participants. Then, the questionnaires were collected and returned to the researcher for further analyses.

Data analysis

The purpose of this study was to explore the possible causal relationships among online metacognitive strategies, hypermedia annotations, motivation, and hypertext comprehension. Since the literature (Hasan, 2003; Murphy & Alexander, 2000; Salmerón et al., 2005) indicates that motivation is an important mediator towards hypertext comprehension development, it was employed as a mediator in the hypothesized model. To be more specific, a path analysis model was proposed based on the assumptions that if EFL learners perceived high frequent use of online metacognitive strategies and hypermedia annotations, they would increase their motivation, and in turn would enhance hypertext comprehension performance. The exogenous variables included online metacognitive strategies (i.e., global, problem-solving, and support strategies) and hypermedia annotations (i.e., word glossary, sentence translation, and background information); motivation (i.e., computer self-efficacy, goal orientation, and situational interest) acted as a mediator served as an endogenous variable. Three main research questions are composed based on the research purpose:

- To what extent is online metacognitive strategies related to hypermedia annotations?
- To what extent is online metacognitive strategies and hypermedia annotations related to motivation during learning with hypertext?
- To what extent is the relationship among online metacognitive strategies, hypermedia annotations, and motivation related to hypertext comprehension?

Results

Relationship of online metacognitive strategies and hypermedia annotations

To answer the first research question, which sought to identify the nature of the relationship between online metacognitive strategies and hypermedia annotations, a bivariate correlation was conducted. The frequency that participants accessed a particular annotation was computed for analysis.

![Figure 5. The path results of the research model](image-url)
As shown in Figure 5, no statistically significant relationship was found between the self-reported metacognitive strategy use and the actual use of annotations ($r = .150, p = .191$). Thus, this result does not confirm with hypothesis one, which indicates that there is a significant correlation between these two exogenous variables. To further seek the relationship of specific type of strategy use and annotation application, it is found that global strategy was significantly related to sentence translation ($r = .342, p = .041$), suggesting evidence of a positive correlation between these two individual variables.

**Relationships among online metacognitive strategies, hypermedia annotations, and motivation**

To investigate the relationships among online metacognitive strategies, hypermedia annotations, and learning motivation, a standard multiple regression analysis was performed; online metacognitive strategies and hypermedia annotations served as independent variables and motivation was used as the dependent variable. Results (see Figure 5) indicated that the correlation coefficients ($r$) between online metacognitive strategies and hypermedia annotations on motivation were .646 ($p = .000$) and .320 ($p = .028$), respectively. Regarding the regression results, only the variable of metacognitive strategy use entered the regression model ($\beta = .646, p = .000$) and it explained 41.7% of the variability ($R^2 = .417$) in learning motivation. However, the path coefficient between hypermedia annotations and motivation was .228, revealing an insignificant impact between these two variables ($p = .084$). It is, therefore, suggested that the usage of online metacognitive strategies most significantly facilitated learning motivation; as the frequency of metacognitive strategy use increased, so did the learning motivation. This result showed that only hypothesis two is confirmed, indicating that online metacognitive strategies most significantly and positively affected learning motivation, but the significant impact of the overall hypermedia annotations on motivation (H3) was not observed.

To further explore which type of specific metacognitive strategies and annotations might affect learning motivation, individual regression results showed that global strategy ($\beta = .703, p = .000, R^2 = .495$) and sentence translation ($\beta = .469, p = .004, R^2 = .220$) were observed to reveal significant and positive relationships in enhancing motivation while reading hypertexts.

**Relationships among online metacognitive strategies, hypermedia annotations, motivation, and hypertext comprehension**

A series of standard multiple regression analyses were further performed on hypertext comprehension scores ($M = 12.27, SD = 2.04$) as the dependent variable, with online metacognitive strategies, hypermedia annotations, and motivation as the independent variables. The mean scores and standard deviations for each variable are illustrated in Table 2. Based on the regression analysis results, only the variable of motivation entered the regression model; that is, the variable of learning motivation significantly predicted the hypertext comprehension score ($\beta = .372, p = .025$), and it explained 13.8% of the variance ($R^2 = .138$) for hypertext comprehension. As hypothesized, the variable of motivation had a significant unique relationship with hypertext comprehension outcomes. In particular, computer self-efficacy was found to positively and significantly influence hypertext comprehension performance ($\beta = .363, p = .027, R^2 = .132$).

As presented in Table 3, the exogenous variable of online metacognitive strategies was the most significant indicator that positively affected learning motivation ($\beta = .646, p = .000$), and motivation was the most significant factor to directly affect hypertext comprehension ($\beta = .358, p = .015$). In addition, through learning motivation, metacognitive strategy use had an indirect effect on hypertext comprehension ($\beta = .231, p < .05$). Thus, the total effect of online metacognitive strategies on hypertext comprehension was .309 (.078+.231). Another exogenous variable of hypermedia annotations, however, did not have a significantly direct effect on motivation ($\beta = .228, p = .084$), nor had an indirect effect mediated by motivation on hypertext comprehension ($\beta = .082, p > .05$). Based on the research results, the hypotheses of H4 and H5 are supported, revealing that motivation positively affected hypertext comprehension, and motivation significantly mediated the relationship between online metacognitive strategies and hypertext comprehension. On the other hand, the two direct effects between online metacognitive strategy ($\beta = .078, p = .322$) and hypermedia annotation ($\beta = .250, p = .071$) on hypertext comprehension were not observed, and thus hypothesis six is not confirmed.
Table 2. Descriptive statistics of metacognitive strategies, annotations, and motivation

<table>
<thead>
<tr>
<th>Metacognitive strategies</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>3.73</td>
<td>.98</td>
</tr>
<tr>
<td>Problem-solving</td>
<td>3.96</td>
<td>.94</td>
</tr>
<tr>
<td>Support</td>
<td>3.47</td>
<td>1.05</td>
</tr>
</tbody>
</table>

Annotations
- Word glossary: .28, .45
- Sentence translation: .64, 1.13
- Background information: 1.19, 1.09

Motivation
- Computer self-efficacy: 3.86, .79
- Goal orientation: 3.44, .89
- Situational interest: 3.68, .89

Reading comprehension: 12.27, 2.04

Note. N = 37.

Table 3. Summary of regression analysis

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Dependent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exogenous variable</td>
<td>Motivation</td>
</tr>
<tr>
<td>Online metacognitive strategies</td>
<td>Direct effect</td>
</tr>
<tr>
<td></td>
<td>Indirect effect</td>
</tr>
<tr>
<td></td>
<td>Total effect</td>
</tr>
<tr>
<td>Hypermedia annotations</td>
<td>Direct effect</td>
</tr>
<tr>
<td></td>
<td>Indirect effect</td>
</tr>
<tr>
<td></td>
<td>Total effect</td>
</tr>
<tr>
<td>Endogenous variable</td>
<td>Motivation</td>
</tr>
<tr>
<td></td>
<td>Direct effect</td>
</tr>
<tr>
<td></td>
<td>Indirect effect</td>
</tr>
<tr>
<td></td>
<td>Total effect</td>
</tr>
</tbody>
</table>

Note. *p < .05; **p < .01.

Discussion and conclusion

The purpose of this study was to explore how online metacognitive strategies, hypermedia annotations, and motivation affected reading comprehension in a Taiwanese hypertext learning environment. A path analysis model was proposed based on the assumptions that if EFL learners perceive high frequent use of online metacognitive strategies and hypermedia annotations, they will increase their learning motivation, and in turn will enhance hypertext comprehension performance. Several major investigation findings are revealed based on the current research results.

First, no significant relationship was observed between the overall online metacognitive strategy use and actual use of hypermedia annotations. Nevertheless, there was a significant relationship between the specific type of global strategy and sentence translation, and these two individual variables can significantly and directly enhance learners' motivation and in return performed better on hypertext comprehension. Consistent with previous findings which indicate that with amount of translation accessed and with greater levels of global strategy use, learners had higher levels of motivation so as to affect hypertext comprehension positively and indirectly (Ercetin, 2003; Huang et al., 2009; Salmerón et al., 2005). It is, therefore, suggested that the use of first language translations and a series of global strategies are correlated to monitor text understanding due to possessing higher levels of motivation.

Counter to expectations, the use of overall metacognitive strategies and hypermedia annotations did not directly and significantly contribute to better hypertext comprehension. This finding aligns with the previous studies (Ariew &
Ercetin, 2004; Ercetin, 2003; Li et al., 2014), showing that metacognitive strategy application and annotated text may not be useful for reading comprehension. Part of the reason may contribute to students’ easy loss in a sea of information because reading hypertexts often requires a more cognitively demanding mode of learning (Nowak, 2008). As suggested by Chen et al. (2011), it is not easy to be an active, constructive, and self-regulated reader; the reading process, therefore, should be scaffolded by an appropriate strategy instruction and pre-, during-, and post-reading guidance with multiple representations in order to enhance EFL students’ reading comprehension.

Another interesting finding indicates that the variable of online metacognitive strategies was the most significant factor that directly affected learning motivation. This finding is consistent with Schmidt and Ford’s (2003) research result, revealing that learners with greater levels of metacognitive strategy use had higher levels of motivation. Further, a significant and positive direct effect of motivation on hypertext comprehension was observed; particularly, computer self-efficacy had the most significant effect on hypertext reading performance. This lends to Becker and Dwyer’s (1994) finding, demonstrating that learners who used the hypertext program were more self-determined and they perceived to have high capabilities to perform the hypertext task successfully. Thus, to increase students’ computer self-efficacy and engagement, it is important for instructors to create sufficient opportunities in doing meaningful learning tasks that can enhance EFL students’ achievement and interest. Activities involving effective computer applications that students can use in their own lives may be useful in increasing their learning intentions (Hsu & Huang, 2006).

Limitations and future research

This study represents a preliminary attempt to explore the possible causal relationships of online metacognitive strategies, hypermedia annotations, and motivation on hypertext comprehension. Due to the research finding of no significant relationships between metacognitive strategy use and annotation application, and between these two variables on hypertext comprehension, the present study suffers from several limitations.

First, there is no qualitative evidence to figure out how EFL students build sequences of information within a hypertext, nor how they navigate hypertexts to influence their learning motivation. Therefore, the results of the present study should be interpreted with caution. As scholars suggest that structural features and order of selecting annotation links are factors that influence the effect of hypertext reading (Akyel & Erçetin, 2009; DeStefano & LeFevre, 2007), future research should focus more on qualitative investigations regarding the sequence of selecting annotations, and how hypertext structural features affect the choice of reading strategies, as well as learners’ motivation and reading comprehension in hypertext learning environment.

The second limitation is that the research results might not be reliable because of conducting only seven hypertext tasks as well as a small sample size, and these might be judged to be too few. More samples and reading tasks are obviously needed to be done in the future to clarify the effect of metacognitive strategy use, hypermedia annotations, and motivation on EFL reading comprehension.

Acknowledgements

This work was supported by the National Science Council (NSC 102-2410-H-214-016).

References


Mobile Seamless Technology Enhanced CSL Oral Communication

Yu-Ju Lan* and Yen-Ting Lin

Department of Applied Chinese Language and Culture, National Taiwan Normal University, Taiwan // yujulan@gmail.com // roylin1003@gmail.com

Corresponding author

(Submitted June 8, 2015 Revised September 28, 2015 Accepted October 29, 2015)

ABSTRACT

The current study aimed at investigating how mobile seamless technology can be used to enhance the pragmatic competence of learners of Chinese as a second language (CSL). 34 overseas CSL learners participated in this study. They were randomly assigned into two groups: the classroom group, executing language tasks in fake contexts in a traditional classroom; and the real-world group, executing identical tasks in real world with the support provided by a mobile seamless learning platform (MOSE). All the CSL learners were asked to make a plan of receiving a friend abroad after collecting the information about the shops or stores in the neighborhood of the campus. Both quantitative (Mandarin communication performance test) and qualitative data (videos recorded during the learning process) were collected and analyzed in this 4-week study. The analytical results show that both groups made significant improvements in the test-based Mandarin communication performance. However, according to the qualitative data, the CSL learners in the real-world group made significantly fewer errors when executing language tasks than did those in the classroom group. Furthermore, they did not depend on their first language to communicate with the people they visited in the real world and they had more peer cooperation with the support provided by the MOSE platform compared with those in the classroom group.

Keywords

Contextual language learning, Immersive learning, Mandarin, Chinese as a second language, Context awareness, mobile seamless learning

Introduction

Being able to use a foreign/second language (FL/L2) appropriately is an essential component in evaluating the success of FL/L2 education. Pragmatic competence referring to the ability to use language appropriately in different social situations, thus, should be considered in FL/L2 teaching, as described in The Common European Framework of Reference for Languages (Council of Europe, 2001) and the proficiency guidelines developed by the American Council on the Teaching of Foreign Languages (Swender et al., 2012).

In order to develop the pragmatic competence in the target language, some approaches have been suggested and adopted in FL/L2 education. Among them, context-based language learning is heavily highlighted by FL/L2 researchers and educators (Serrano, Llanes, & Tragant, 2011). Context-based instruction has a foundation on the sociocultural theory of second language acquisition (SLA) which emphasizes the integrated nature of learner. Moreover, social context elements in the learning process (Eun & Lim, 2009) include the contexts and the interaction mediate language learning, and thus they play an important role in the SLA process (Ellis, 2008). According to the perspective of sociocultural SLA, immersing in an authentic context is important for L2 learning (Lan, 2014) because an L2 cannot be acquired merely via context-reduced practicing by rote. L2 learning which emphasizes the importance of learners using the target language in an authentically immersive environment befits L2 learners’ oral performance and forms accuracy (Lan, Kan, Hsiao, Yang, & Chang, 2013). The evidence obtained from brain-related research also supports context-immersive learning for L2 acquisition (Zinser & Li, 2012).

As Mandarin Chinese learning has become popular globally over recent years (Ramzy, 2006), many people have traveled to such countries as China, Taiwan, and Singapore, in which Mandarin Chinese is the dominant or primary language, in order to acquire the language. Taiwan has been one of the most popular countries for learners of Chinese as a second language (CSL), and especially for overseas Chinese students from around the world because the Chinese tradition and culture has been preserved on this Asia Pacific Island (Lan, 2014). How we could meet the learning needs of those overseas Chinese students in appropriately using Mandarin in real-life occasions, consequently, becomes a challenge to Mandarin training institutes (Lan, Lin, & Tsai, 2014). To take up the challenge and to consider the importance of contexts for SLA, context-based real-life language tasks are usually included in the
course program of Mandarin Chinese in addition to in-class Mandarin language skill instructions to provide CSL learners with diverse experiences in exploring Mandarin Chinese (Lan, Lin, & Tsai, 2014).

A language task is something that people do in their everyday lives (Long, 1985), in which the settings and the conditions under which the task takes place are two essential elements (Nunan, 1989). The two elements referring the authentic contexts and the social interaction, as described above, both are essential in the SLA process (Ellis, 2008). Obviously, real life contexts should be first constructed for CSL learners, and then learners should practice using the learned language in social interaction. However, the two issues described below should be dealt with if successful real-life context-based language learning is anticipated. (1) It is uneasy for CSL teachers in traditional classrooms to create authentic contexts for learners to immerse themselves into the situations and carry out language tasks (Yue, 2009). The lack of similar real life contexts does not only lower CSL learners’ performance but also their motivation (Lan, Lin, Kao, Chang, Sung, & Liu, 2015). (2) While carrying out language tasks, especially in real world, some obstacles are encountered by CSL learners, including the insufficient pragmatic competence for having appropriate social communication (Lan, Lin, & Tsai, 2014) and the low motivation in using Mandarin Chinese in daily interaction rather than using their first language (Edge, Searle, Chiu, Zhao, & Landay, 2011).

The above-mentioned obstacles encountered by both the CSL learners and teachers inspire the applications of mobile seamless technology into CSL learning (e.g., Chen, Seow, So, Toh, & Looi, 2010; Edge et al., 2011; Liu, Sunaoka, & Urano, 2005; Wong, Chin, Tan, & Liu, 2010). Recently, mobile seamless technology has attracted FL/L2 researchers and educators because of its potential in integrating the contexts in real world into language learning and teaching (e.g., Wong, 2012). This paper, therefore, focuses on bridging the in- and out of classroom via mobile seamless technology to provide CSL learners with authentic and real contexts in their language learning. Relevant literature about mobile seamless technology used for FL/L2 learning is briefly described below.

**Review of mobile seamless technology for FL/L2 learning**

Mobile seamless learning, compared with general mobile learning, emphasizes more on the transparency of the barriers among different learning contexts than the use of personal mobile devices as a mediator to do individual learning (Milrad, Wong, Sharples, Hwang, Looi, & Ogata, 2013). Through mobile seamless learning, learners can easily and quickly switch from one scenario to another, such as between in and out of the school or between formal and informal learning contexts without any interruptions caused by context/scenario transformation (Wong, 2012). In addition to applying seamless technology to expand learning opportunities (Huang, Kuo, Lin, & Cheng, 2008; Wei, 2012), seamless learning has also developed a good reputation for addressing some of the problems that have long existed in traditional FL/L2 language learning. One example is that teachers often use second-hand experience that places a limit to the classroom environment (Uosaki, Ogata, Sugimoto, Li, & Hou, 2012). The successful application of seamless technology in learning English as a foreign language (EFL) (e.g., Uosaki et al., 2012) also attracts the attentions of researchers of CFL/CSL, even though the volume of the related literature is still small. For example, Edge and colleagues (Edge et al., 2011) developed a mobile learning platform to enable CSL learners to access location-based Mandarin learning materials. Step by step, CSL learners used this platform to learn the most frequently used Mandarin sentences pertaining to their current situation. Liu and colleagues (Liu et al., 2005) provided Japanese students with corpus-based drill and testing systems to learn Chinese by using a phone or a PDA. According to preliminary evaluation conducted by Liu et al. (2005), the corpus-based Chinese learning system not only successfully promoted Japanese students’ motivation of learning Chinese but also benefited their pre- and review results. Comparing with the studies conducted by Edge et al. (2011) and Liu et al. (2005), Wong and colleagues (2010) asked elementary-school CSL students in Singapore to create their own learning materials of Chinese idioms via mobile phones rather than using the ready-to-use embedded materials, as those mentioned in the two studies above. The elementary-school students in Wong et al.’s (2010) study took pictures from the real world after school to illustrate the idioms they learned in their Chinese classes. Next, they uploaded them to the class-shared wiki page and some sentences to explain the meanings of the pictures they had uploaded. Through analyzing students’ meaning making process, Wong et al. (2010) concluded that seamless language learning has the potential of transforming language learning in the class into an authentic learning experience.

Although the above-mentioned studies either assisted CSL/CFL learners learning Chinese on the move (e.g., Edge et al., 2011) or bridged the gap between the in-classroom and real-world contexts (Wong et al., 2010), none of them investigated how seamless learning enhances CSL learners’ pragmatically oral communication skills which are
definitely important to their daily lives. Since the real communicative occasions are often different from the pre-embedded learning materials, the promising characteristics of general seamless learning systems that simply promise an easy and quick switch among scenarios/contexts are not enough for CSL learners to carry out socially real-world interaction. Obviously, providing them with timely and context-fitted support is also important as well. Although Chen and her colleagues (Chen et al., 2010) mentioned a number of approaches to and the potential of connecting in- and out-of-classroom via using mobile seamless technology in Chinese learning, only suggestions, rather than any practical findings or evidence were proposed in their reports.

To add to the knowledge pool of the potential of seamless learning for enhancing CSL learners’ pragmatic competences, as well as to solve the heritage problems caused by de-contextual teaching approaches existing in traditional CSL classrooms, the purposes of the study are (1) to bridge the learning in- and out of the classroom, and (2) to confirm the effects of learning Mandarin Chinese in real life contexts on CSL learners’ pragmatic competences through mobile seamless learning.

Three research questions are dealt with to reach the purposes of this study:
- What are the differences in the effects on CSL learners’ test-based oral communication performances between different learning contexts (conventional classroom learning vs. mobile seamless learning)?
- What are the differences in the effects on CSL learners’ social-interaction skills between different learning contexts (conventional classroom learning vs. mobile seamless learning)?
- What are the differences in the effects on CSL learners’ usages of communication strategies in social interaction between different learning contexts (conventional classroom learning vs. mobile seamless learning)?

Methodology

Participants

34 overseas CSL beginners participated in this study (28 were from Thailand; 3 were from Canada; 2 were from USA; and 1 was from Brazil). They were randomly assigned into two groups, the conventional classroom group and the mobile seamless group. Consequently, 18 were assigned to the mobile seamless group (mean age = 20.00 years, SD = 1.25 years, range = 18 – 23 years) and 16 to the conventional classroom group (mean age = 19.44 years, SD = 1.38 years, range = 17 – 22 years). Although the numbers of the participants in each group were not the same, there was no difference in age ($F = 1.554, p > .05$) between the two groups. Additionally, because all the participants were beginners and were randomly assigned into the two groups, the participants’ Mandarin Chinese abilities between the two groups were considered the same.

Research design

A mixed research design (Creswell, 2003) was adopted in this study in which both qualitative and quantitative data were collected and analyzed because CSL learners’ pragmatic competences cannot be confirmed solely by the traditional test items without observing how the CSL learners use the target language in social interaction. For the quantitative data, participants’ test-based Chinese oral communicative performances were evaluated via both pre- and posttests. On the other hand, the processes of carrying out tasks by the two groups were recorded and analyzed according to a video coding scheme which focuses on identifying CSL learners’ communication strategy as the qualitative data. The coding scheme will be described in the section of “Video coding scheme.” Additionally, the quantitative data were analyzed via a two-way mixed design ANOVA (group: classroom vs. mobile seamless; test: pre- vs. posttest), while the content analysis was adopted to analyze the qualitative data.

Instruments

MObile SEeamless (MOSE) learning platform

The mobile seamless platform used in this study, “MOSE,” was developed by the authors. According to the sociocultural SLA theory, the social context and interaction are essential to the SLA process (Ellis, 2008; Lantolf,
The main functions of the mobile seamless learning platform (MOSE) are (1) to bridge the learning that occurs inside and outside the classroom by providing CSL learners with opportunities for what have been learned in the classroom to be used at some occasions out of the classrooms; (2) to provide timely support needed in socially real-world communication; and (3) to record and share learners’ exploring experiences with their peers. The above-described functions of the MOSE are implemented by the technologies of location-based-service and quick response (QR) code. Simply put, the MOSE allows CSL learners to access the learning materials fitting the contexts where they visit. In addition, they can further access any ready-to-learn materials via a QR code scanner installed in their mobile devices (e.g., a smart phone or an iPad). While exploring a real-world context, CSL users can also upload any materials, such as pictures and texts to the MOSE to share them with their peers. Figures 1 and 2 show the examples of context-fitted materials provided by the MOSE when the attending CSL learners were carrying out the assigned tasks; while Figure 3 shows what CSL learners uploaded to the MOSE according to the mission requests when they visited a nearby shop.

Figure 1. Left: locations visited (“learning spots”) with learning materials; Right: learning materials for learners exploring one representative building (the Commons) in the campus

Figure 2. The context-fitted materials needed for CSL learners to carry out the assigned mission when they visit an attar shop close to the campus
Figure 3. Screenshots of information collection and sharing, showing the mission assigned by the teacher (left), photographs of fabric-made objects taken by student #130622 (center and right), and the information shared by student #130622 (right)

Language task

One language task, “Receiving a friend from abroad,” was developed and used in this study. The goal of this task is to collect needed information for making a plan to receive a friend coming from other countries. All the participants were asked to carry out this task by (1) exploring the neighborhood of the campus, (2) collecting the needed information for arranging a 7-day tour, including daily and unexpected activities, and (3) scheduling a visit plan by deciding the must-visit spots (such as a bakery, shops, or restaurants). To help the CSL learners successfully accomplish this mission, the needed key words and sentences were taught before their mission started. Appendix A shows the identical learning materials for the two groups, while Appendix B shows the paper-based materials used by the CSL learners who role-played as the owners and the customers in the classroom context. Additionally, the information listed in the paper-based materials is real, existing in the real world. Therefore, although the CSL learners in the conventional classroom group did not visit the real spots, they could collect the needed information for accomplishing the assigned tasks. In contrast, the CSL learners in the MOSE group would use the learned words and sentences listed in Appendix A to interact with the real owners of the stores or restaurants visited in the real world. When they encountered problems or needed help in the real-world social interaction, the CSL learners in the MOSE group could reach the support by accessing the MOSE.

Mandarin communication performance test

The mandarin communication performance test (hereafter referred to as the Mandarin test) focuses on communication skills. Thus, both the test items and answer options were delivered aurally; no printed texts were shown to the participants. All the participants were first asked to listen to the questions and 3 options carefully and then choose the correct answer to each question. The following is one example of the items in the performance test.
Question: (Aural) 哪一種商品最受歡迎? (Which product sells the best?)

Option 1: (Aural) 你當然受歡迎了。 (You are definitely popular.)

Option 2: (Aural) 每一種都賣得很好。 (All the products sell well.)

Option 3: (Aural) 那個穿紅衣服的女生。 (It is the girl in a red T-shirt.)

**Video coding scheme**

The video coding scheme used in this study includes 4 dimensions: completion of the conversation goals, unexpected conversation, appropriateness between the contexts and the used words, and the strategies used (see Appendix C). Additionally, the communication strategy coding scheme follows the structure proposed by Oxford (1990). While coding a video, the encoder counted and recorded the frequency of each target behavior or conversation on the scheme.

**Procedure**

This study lasted 4 weeks, 3 hours a week, from May 9th to 27th, 2014. The pretest of the Mandarin test was administered followed by the traditionally classroom-based instructions of the basic Chinese words and sentences needed in carrying out the language tasks in the first week. Then, in the second and third weeks, students had to accomplish the language task, “receiving a friend from abroad.”

The task mission assigned to the two groups was identical as described in the section of instruments: to explore the campus neighborhood and collect information of certain spots that might be visited while receiving their friends. The difference in carrying out the language task is the contexts they explored. The participants in the conventional classroom group (hereafter referred to as the classroom group) role-played as customers and the owners or the waiters of the spots visited. As an owner or a waiter, the student would use the poster (see Appendix B-1) to show the customers what type of business s/he was running. S/he also needed to answer the customers’ questions according to the information listed on the information cards (see Appendix B-1). On the other hand, as a customer, the student had to ask the waiters or the owners of the visited spots proper questions chosen from what are listed on the worksheet (see Appendix B-2) to collect the information needed. In contrast, the participants in the mobile seamless group (hereafter referred to as the MOSE group) explored the neighborhood of the campus and visited the target spots to collect the information needed by asking the real waiters or the owners appropriate questions. With their smart phones or iPads, they could access the MOSE anytime to obtain timely support, if encountering any problems about not knowing how or what to ask or about the lack of the knowledge about the visiting spots. They could also take pictures or record the needed information and upload them to the MOSE to share them with their peers.

Having explored and collected the information, they had to schedule a receiving plan with their peers. Finally, in the fourth week, the identical Mandarin test was administered as the posttest. Furthermore, in order to investigate how CSL learners interacted with one another, the processes of carrying out the language tasks in the two groups were video-recorded.

**Results**

To evaluate how the learning affected overseas CSL learners’ both oral communication performance and pragmatic interaction between different groups, both quantitative (the scores of the performance test) and qualitative data (the video data) were collected and analyzed. The results are described below.

**Comparison of learning gains on test-based mandarin communication performance**
Table 1 lists the descriptive statistics of the scores of the two groups at pre- and posttest, while Table 2 is the summary of the analysis results of two-way mixed design ANOVA. According to the data listed in Table 2, there was neither significant interaction between group and test, $F(1, 1) = .115, p > .05$; nor the group effect, $F(1, 1) = .014, p > .05$. However, the test effect reached a significant level [$F(1, 1) = 376.927, p < .001$], indicating that both groups made significant improvements at posttest of the Mandarin communication performance. The results of the main effect analysis are as follows: classroom: $F(1, 1) = 67.052, p < .001$; MOSE: $F(1, 1) = 76.281, p < .001$. These results indicate that the improvement made by the MOSE group is larger than that made by the conventional classroom group, although the difference in the improvement between the two groups was not significant.

### Table 1. Descriptive statistics of the Mandarin test

<table>
<thead>
<tr>
<th>Test</th>
<th>Classroom ($N = 16$)</th>
<th>MOSE ($N = 18$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Pretest</td>
<td>5.31</td>
<td>2.80</td>
</tr>
<tr>
<td>Posttest</td>
<td>15.44</td>
<td>4.08</td>
</tr>
</tbody>
</table>

### Table 2. The analysis results of two-way mixed design ANOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>.265</td>
<td>1</td>
<td>.265</td>
<td>.014</td>
<td>.907</td>
</tr>
<tr>
<td>Test</td>
<td>1677.687</td>
<td>1</td>
<td>1677.687</td>
<td>376.927</td>
<td>.000***</td>
</tr>
<tr>
<td>Group*Test</td>
<td>.511</td>
<td>1</td>
<td>.511</td>
<td>.115</td>
<td>.737</td>
</tr>
<tr>
<td>Errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS$_{BS}$</td>
<td>609.500</td>
<td>32</td>
<td>19.047</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS$_{WS}$</td>
<td>142.431</td>
<td>32</td>
<td>4.451</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. ***$p < .001$.

**Comparison of oral communication process via video data analysis**

The task executing processes of the two groups were recorded via digital cameras and were analyzed. The collected video data were coded according to 4 dimensions as shown in Appendix C by two encoders who received training on coding the videos. The Kappa coefficient of agreement of the results from the two encoders was .674 ($p < .001$), indicating that the inter-encoder agreement reaches a substantial level.

Regarding the completion rate of the communication goals, both groups successfully completed all the assigned tasks, showing no differences. However, the big differences between the two groups are in the dimensions of the unexpected conversations, the context fitness, and the communication strategy usage. For the unexpected conversations, it was found that this kind of conversation was caused by different occasions, including error questioning or answering, task executing, and friendship expressing. This first occasion, error questioning or answering, was observed when the CSL learners made errors and sequentially the dialogue went beyond what was taught in classes.

Table 3 lists the frequencies of the unexpected conversations found in the two groups. The Chi-square analysis result reveals that the difference between the two groups is significant [$X^2(1,2) = 27.239, p < .001$]. It was also found that in the MOSE group, most of the unexpected conversations were found when CSL learners executed the assigned missions, while most of what happened in the classroom group was caused by errors made by the CSL learners. As a result, with the support provided by the MOSE, the CSL learners made significantly fewer errors when executing the assigned tasks than did those in the other group.

### Table 3. The frequency and chi-square test of the unexpected conversations in the two groups

<table>
<thead>
<tr>
<th>Contexts</th>
<th>Error questioning</th>
<th>Task executing</th>
<th>Social interacting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom</td>
<td>10</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MOSE</td>
<td>1</td>
<td>19</td>
<td>5</td>
</tr>
</tbody>
</table>
It was also found that most of the unexpected conversations identified in the MOSE group happened when they were executing the assigned tasks. Additionally, in order to begin a conversation or show appreciation to the interlocutors in real world, some dialogue expressing friendship or kindness was also identified in the MOSE group, the type of unexpected conversations not found in the classroom group.

With regard to the context fitness, the video data also depicted that most of the CSL learners in the classroom group usually asked all the questions listed on the learning material (Appendix A) without taking the appropriateness between the questions and the visited spots into consideration. In contrast, fewer the same type of errors were found in the MOSE group, although sometimes they were confused due to the lack of the knowledge of a specific store in Taiwan. Actually, only one was identified, as listed below in Table 4.

Table 4 lists the frequency of both appropriate and inappropriate conversations found in the two groups. Although the CSL learners in the MOSE group also asked inappropriate questions as their peers in the other group did, it happened only once. Furthermore, the Chi-square analysis also reveals the significant difference in having inappropriate conversations between the two groups \( \chi^2(1,1) = 15.834, p < .001 \).

Table 4. The frequency of the context-fitness conversation happened in the two groups

<table>
<thead>
<tr>
<th>Contexts</th>
<th>Fitness</th>
<th>Non-fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>MOSE</td>
<td>30</td>
<td>1</td>
</tr>
</tbody>
</table>

The current paper now moves on to discussing the communication strategies used by the CSL learners. Interestingly, only two main categories of the communication strategies are identified, compensation and social strategies as listed in Table 5. Additionally, 4 strategies belong to the former category and 2 belong to the latter one. The Chi-square analysis shows that the strategies used by the CSL learners between the two groups are significantly different \( \chi^2(5,1) = 28.136, p < .001 \).

Table 5. The used frequencies of conversation strategies in the two groups

<table>
<thead>
<tr>
<th>Communication strategies</th>
<th>Compensation strategies</th>
<th>Social strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Language switching</td>
<td>Using mother language</td>
</tr>
<tr>
<td>Classroom</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>MOSE</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

As listed in Table 5, the strategy most commonly used by the CSL learners in the classroom group was confirmation, a social strategy done by repeating what they heard. It is also found that it was easy for them to use their first language (L1) or non-verbal strategy (body language in this study) to communicate with each other. Besides, few of them in the classroom group tried to cooperate with their peers to execute the assigned tasks. In contrast, in the MOSE group, although the CSL learners also used the strategy of confirmation by repeating what they heard from the real shop owners, they tended to cooperate with peers to complete the assigned tasks.

Discussion

Three issues about the effects of using mobile seamless learning on the learning gains of CSL learners are investigated: test-based oral communication performance, social-interaction skills, and the usages of communication strategies. A brief discussion about the three issues based on the analytical results of this study is described below.

Regarding the learning gains in terms of the test-based results, all the participants with or without the MOSE when executing the assigned tasks made significant improvements (see Tables 1 and 2). The positive results confirm the effects of context-based learning on FL/L2 oral performances as argued by numerous pieces of research on contextual influence on L2 learning (e.g., Jung, 2002; Kasper & Schmidt, 1996; Serrano, Llanes, & Tragant, 2011). However, a deeper exploration of the effects of mobile seamless learning on CSL learners’ pragmatic competency
should not be hesitated because of the similar improvements made by the CSL learners in the two groups according to the test-based results. As Solano-Flores and Trumbull (2003) suggested, examining a language should take the context factor into account. We further analyzed the video data to see the possible differences in learning gains of CSL pragmatic competency between the two groups. Prospectively, the processes of the task execution of the MOSE group significantly differed from those of the classroom group by three aspects: the unexpected conversations, the context fitness, and the usage of communication strategies.

Regarding the unexpected conversations, the participants in the real world context encountered significantly more of this kind of conversations than those in the classroom (Table 3) did. It is worthy of notice that more unexpected conversations do not mean more errors made. As seen in Table 3, more errors were made by the participants of the classroom group than those done by the other group. It depicts that the MOSE system successfully lowered the error rate happened in social conversation of the CSL learners. Furthermore, although the MOSE group had to deal with more unexpected conversations, they accomplished the assigned goals via social interaction as successfully as the classroom group did. This finding is in line with the results of related studies of mobile assisted language learning that mobile learning significantly improved FL/L2 learners’ FL performances (e.g., Lan, Lin, & Tsai, 2014; Liu & Chu, 2010).

Additionally, the data in Table 4 also support the information conveyed in Table 3: the context fitness ration of social communication is significantly higher in the MOSE group than that in the classroom group. Almost all the CSL learners in the classroom group simply read all the sentences listed on the worksheet and the information cards without carefully considering the appropriateness of the visited spots. In contrast, the MOSE group was able to distinguish the differences among the visited spots and consequently used the appropriate questions to ask the native Chinese speakers (the shop owners). The differences in choosing dialogue sentences which were taught before carrying out the assigned tasks between the two groups are similar to the findings of Lafford’s (1995) study. However, the aim of Lafford’s study was to compare the differences between a study abroad context and a study at home. She found that the students studying Spanish-abroad outperformed those studying at home. Interestingly, all the participants of the current study were studying Chinese-abroad. As Lan (2014) mentioned, the low oral communication competency is a general problem encountered by most of the long-staying overseas CSL students in Taiwan. The study-abroad period of the participants of this study was much shorter than that of the participants of Lan’s (2014) study, 6 weeks versus 1 year. Thus, it is possible that all the participants would have the same problems in social conversations. The significantly appropriate usage of Chinese in social conversations by the MOSE group, therefore, approves the effects of the MOSE system on enhancing CSL learners’ pragmatic competency in social conversation.

The arguments provided above are re-approved when the discussion switches to the communication strategies used by the CSL learners. Table 5 shows that the communication strategies used by the CSL learners in the two groups are significantly different. Results show that the CSL learners in the classroom tended to use their L1 to continue conversations when they were stuck, an act not found in the MOSE group. Although the real world context could provide the CSL learners with opportunities for immediate contact with members of the target language community which is proved to be one of the main factors facilitating the SLA (e.g., Jung, 2002), the degree of activity involvement of the L2 learners is also an important factor that might influence L2 learners’ SLA. Additionally, according to the study of Yashima and Zenuk-Nishide (2008), an imagined context could also benefit the SLA of FL/L2 learners if they are fully involved in the learning activities. Thus, with the MOSE system, CSL learners’ motivation of involving in social conversation in Mandarin Chinese was obviously promoted. This method solves the problems pointed out by Lan (2014) that due to the lack of motivation of speaking Mandarin Chinese, most of the overseas Chinese students tend to associate with other students from the same country of origin outside the classrooms and speak their common L1 outside the classrooms.

Interestingly, twice of circumlocution usage was identified in the MOSE group (see Table 5). The richer usage of communication strategies might be caused by the more unexpected conversations encountered in the MOSE group. However, the results show that the willingness of and the efforts made on continuing social conversations in the MOSE group are more obvious than those of the classroom group. Besides, although both groups used the repeat strategy in the social conversation, the purposes were different. It was found that in most of the occasions, the repetitions of words or sentences by the classroom group were just “repeats,” while the repetitions by the MOSE group were used to confirm the meaning and clarify the ambiguous situations. In addition, the inter-peer cooperation happened much more often in the MOSE group than that in the classroom group did, although all the CSL learners in
the latter group were in the same location. By checking the information uploaded to the MOSE system by the CSL learners (see Figure 3), it was found that the CSL learners in the MOSE group were highly involved in the activities guided by the MOSE system and they fervently collected and shared the needed information to accomplish the assigned mission. The findings are in line with the study of Liu and Chu (2010). The high involvement is the key factor, suggested by Yashima and Zenuk-Nishide (2008), that benefits FL/L2 learners’ SLA, regardless of the learning contexts being abroad or at home. Thus, the MOSE system appears to have successfully promoted the involvement and practice efficiency of the CSL learners in the social conversations, and consequently benefited their learning gains after the treatment.

By synthesizing the findings obtained from this study, we argue that being immersed in survival-oriented contexts inspires FL/L2 learners to actively involve social interaction in the target language, while providing timely support assists them in successfully conquering the obstacles they usually encounter and sequentially reaching the goals of having a conversation and promoting their motivation of having social interaction with native speakers.

Conclusion and limitation

Contextual influence on FL/L2 learning has been an important issue in SLA. It affects the SLA of learners of cross-aged (Kasper & Schmidt, 1996) and of different target languages (Jung, 2002; Lafford, 1995; Lan, 2014; Lan, Fang, Legault, & Li, 2015). The learning contexts also influence learners’ transferability of (Takahashi, 1993) and the amount of transfer (Takahashi & Beebe, 1987) of FL/L2 learning. As Jung (2002) argued, pragmatic competency can be only evaluated in meaningful and authentic interactions. Thus, it is uneasy to confirm CSL learners’ pragmatic competency only via traditional performance tests. Consequently, the development of a new paradigm of FL/L2 assessment should be emphasized in FL/L2 research by researchers and educators. Additionally, the context factor should be also considered when designing a learning program for various contexts, including studying-abroad versus at home country or a long-term versus short-term program.

Furthermore, as the evidence obtained from the current study approves that CSL learners’ active involvement plays an important role in the success of social communication in Mandarin Chinese, both a careful design of learning tasks and timely support should be included in any Mandarin Chinese learning programs. Based on the findings of this study, mobile seamless learning (like the proposed MOSE system in this study) depicts to be a good solution to successfully providing the CSL learners with timely support and effectively inspiring them to actively involve themselves in executing learning tasks.

Despite the positive findings obtained from the current study, the short period of treatment of this study hinders its absolute inference to possible effects on CSL learning in a longitudinal study. A longer period of treatment is anticipated in the future to further confirm the effects of the MOSE system on CSL learners’ pragmatic competency. Besides, only the communication process and the test-based performance of Mandarin Chinese, but not the tone of Mandarin Chinese, the complexity of word and sentences usages, and the fluency of social communication, are evaluated in this study. A deeper investigation on using mobile seamless learning for the SLA process of Mandarin Chinese, from phonetic system, words, sentence, social communication, to culture aspects, is needed to bridge the gap of the knowledge pool between English and Chinese as foreign/second languages.

In sum, the current study not only approves the value of mobile seamless language learning, but also provides a deeper understanding of the effects of technology-supported learning in real world on CSL learners’ pragmatic competency.

Acknowledgements

Funding for this research work is provided by the Ministry of Science and Technology, Taiwan, under Grant Nos. MOST 104-2911-I-003-301 and MOST 103-2628-S-003-002-MY3. This research is also partially supported by the Aim for the Top University Project and Center of Learning Technology for Chinese of NTNU, sponsored by the Ministry of Education, Taiwan, and the International Research-Intensive Center of Excellence Program of NTNU.
References


Appendix A. The learning materials: Basic words and sentences

1. Key words

中式早餐店 zhōngshì zǎocān diàn (Chinese breakfast store), 麵包店 miàn bāo diàn (bakery)
服装店 fúzhuāng diàn (apparel store), 眼科诊所 yǎnkē zhēnsuǒ (eye clinic), 郵局 yóu jú (post office), 中餐廳 zhōng cāng tiáol (Chinese restaurant), 水果店 shuǐguǒ diàn (fruit store)
皮鞋店 píxié diàn (shoe store), 理髮廳 lǐfà tīng (hair salon), 圖書館 túshūguǎn (library)
速食店 sùshí diàn (fast-food restaurant), 超級市場 chāojí shìchāng (super market)
眼鏡行 yǎnjīng háng (eyeglasses store), 咖啡廳 kāfēi tíng (coffee house), 飲料店 yǐnliào diàn
(soft drink store), 冰店 bīng diàn (ice shop), 藥妝店 yào zhuāng diàn (drug store)
毛筆店 máobǐ diàn (brush shop), 咖啡 kāfēi (coffee), 茶 chá (tea), 蛋糕 dāngāo (cake), 餅乾
bǐng (cookie), 素食 sūshí (vegetarian food), 中國菜 zhōngguó cài (Chinese food), 烤雞 kǎoji
(BBQ chicken), 果汁 guǒzhī (juice), 冰 bīng (ice), 早餐 zǎocān (breakfast)

2. Sentences

(1) 請問(你們)什麼時候開門/休息/公休？
   (What is your opening time/When is the break time/day off?)

(2) 請問(你們的)營業時間是什麼時候？
   (What are your business hours?)

(3) 請問可以訂位/刷卡/拍照嗎？
   (May I make a reservation/pay by card/take photos?)

(4) 請問有素食嗎？
   (Do you provide vegetarian dishes?)

(5) 請問哪裡有早餐/麵包店…？
   (Are there any breakfast store/bakery…?)

(6) 請問最低消費是多少？
   (What is the minimum spend?)

(7) 請問用餐時間有多久？
   (How long is the meal time?)

(8) 請問不吃豬肉/不吃牛肉/吃素的人可以買/點什麼？
   (What can those who don’t eat pork/beef/vegetarian order?)
Appendix B. Examples of the paper-based materials used by the classroom context group

1. Poster and information cards used by the waiter of the restaurant

<table>
<thead>
<tr>
<th>Restaurant</th>
<th>餐廳</th>
<th>cāntīng</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00-21:00</td>
<td>週一 zhōu yī</td>
<td>週日 zhōu rì</td>
</tr>
<tr>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
</tr>
<tr>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
</tr>
<tr>
<td>$120 ↑</td>
<td>2 HR ↓</td>
<td></td>
</tr>
</tbody>
</table>

Coffee
咖啡
kāfēi

Cake
蛋糕
dāngāo

Vegetarian Food
素食
sūshí

BBQ Chicken
烤雞
kǎojī
2. Worksheet used by the customers

<table>
<thead>
<tr>
<th></th>
<th>Store</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>開門</td>
<td>Open</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>休息</td>
<td>Close</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>公休</td>
<td>Day off</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>訂位</td>
<td>Reserve</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>刷卡</td>
<td>Credit card</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>拍照</td>
<td>Take Photo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>素食</td>
<td>Vegetarian diet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>最低消費</td>
<td>Min. charge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>用餐時間</td>
<td>Meal Time</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C. Video coding scheme

<table>
<thead>
<tr>
<th>Codes</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completion of the communication goals</td>
<td>According to what students initiated (purpose) and whether the conversation goal/purpose was reached. If yes, count 1, or else count 0.</td>
</tr>
<tr>
<td>Unexpected conversation</td>
<td>According to what was taught before taking the mission, if the words/sentences used in the conversation were not among the materials taught, count 1, or else count 0.</td>
</tr>
<tr>
<td>Context fitness</td>
<td>According the visited stores or shops, if the conversation fits the context, count 1, or else count 0.</td>
</tr>
<tr>
<td>Communication strategies</td>
<td>Observe what strategies the student adopted to continue the conversation when the conversation was stuck. The strategy coding scheme used is based on Oxford’s study (Oxford, 1990: pp. 59, 69, 91, 136). Each usage of the communication strategies counts 1</td>
</tr>
</tbody>
</table>
Effects of a Question Prompt-based Concept Mapping Approach on Students’ Learning Achievements, Attitudes and 5C Competences in Project-based Computer Course Activities

Hsiu-Ying Wang¹, Iwen Huang² and Gwo-Jen Hwang³*

¹Department of Education, National University of Tainan, Taiwan // ²Department of Information and Learning Technology, National University of Tainan, Taiwan // ³National Taiwan University of Science and Technology, Taiwan // cute@tkgsh.tn.edu.tw // huangi@mail.nutn.edu.tw // gjhwang.academic@gmail.com

*Corresponding author

(Submitted June 15, 2015; Revised September 7, 2015; Accepted December 1, 2015)

ABSTRACT

Concept mapping has been widely used in various fields to facilitate students’ organization of knowledge. Previous studies have, however, pointed out that it is difficult for students to construct concept maps from the abundant searched data without appropriate scaffolding. Thus, researchers have suggested that students could produce high quality concept maps when they construct them with appropriate guidance. Accordingly, this study proposes a question prompt-based concept mapping approach, which was conducted in a computer course with a duration of 14 weeks on the research issue of global warming. There were 176 senior high school students from four classes who participated in the experiment. The study adopted a quasi-experimental design for the investigation of the effectiveness of the proposed approach; two classes were randomly assigned to be the experimental group, learning with the question prompt-based concept mapping approach, while the other two groups were the control group, who used the conventional concept mapping approach. The results indicated that the students in the experimental group had better performance than those in the control group in terms of their learning achievement, 5C competences, and learning attitudes. In addition, the results also showed that students with a reflective style had better learning achievement than those with an active style. Thus, it is worth investigating, for such an issue, how to propose more effective learning models for diverse student learning styles in the future. This can provide useful guidance for educators, practitioners and researchers in the areas of science education and computer-assisted education.

Keywords

Concept mapping, Question prompt, 5C competences, Project-based learning, Website development

Introduction

Previous research has indicated that, while most students have information-searching ability on the Internet, they tend to lack effective strategies for finding useful information from the searched data (MacGregor & Lou, 2004). This may cause disruption to students’ learning, especially when they are facing resource-rich and complex learning situations (Yin et al., 2013). Therefore, it is necessary to provide scaffolding or effective tools to support students in summarizing or organizing the information they access on the Internet (Hwang, Kuo, Chen, & Ho, 2014).

The construction of knowledge depends on the mental tools used (Jonassen & Carr, 2000). Concept mapping is regarded as an effective tool that helps students organize important issues, facilitating their construction of knowledge (Hwang, Kuo, Chen, & Ho, 2014; Liu, 2011). However, it is somewhat difficult for students to construct concept maps directly from the huge network of data available to them. Researchers have suggested that students should be given scaffolding such as hints or clues when developing their own concept maps, through which knowledge can be effectively constructed (Gibbs & Habeshaw, 1993). In addition, integrating collaboration and concept mapping could help students in their knowledge sharing and exchange, while also developing their communication skills and strengthening their learning motivation (Güvenç, & Ün Açikgöz, 2007; Hwang, Shi, & Chu, 2011; Kwon & Cifuentes, 2009; Roth & Roychoudhury, 1992; Wood & O’Malley, 1996). That is, having interactive groups is the key component to applying the concept mapping approach (Cheng, Wang, & Mercer, 2014).

With regard to meaningful learning in the process of concept mapping, Brown, Collins, and Duguid (1989) pointed out that it is not only important but in fact necessary to engage students in dealing with learning tasks with real-world problems for knowledge construction. The best practice for students to understand an important issue is to conduct a project-based learning activity in which they are able to develop products or solve problems based on a specific
project that represents their knowledge (Hwang, Hung, Chen, & Liu, 2014). Previous research has proved that project-based learning can cultivate students’ diverse competences and enhance their learning achievement, such as their problem-solving abilities, independent thinking, critical thinking and communication ability, as well as their learning motivation (Bell, 2010; Chang & Lee, 2010; Hung, Hwang, & Huang, 2012; Koh, Herring, & Hew, 2010; Ravitz, Hixson, English, & Mergendoller, 2012). Thus, the project-based learning approach can be regarded as an innovative approach for one of the key abilities in the 21st century (Bell, 2010). In addition, Trilling and Fadel (2009) further indicated that learners in the 21st century should have diverse abilities for addressing complex problems, such as critical thinking, problem solving, communication and team work.

Project-based learning (PBL) is one of the most important teaching strategies; encouraging educators to use this approach in the current education field at this stage of education reform could help foster students’ diverse abilities and active attitudes. Several studies have demonstrated the potential of combining PBL with concept mapping (Kassab & Hussain, 2010; Mok, Whitehill, & Dodd, 2014). Cañas and Novak (2009, p. 1) pointed out that “a good way to delineate the context for a concept map is to define a Focus Question,” which is a question that “clearly specifies the problem or issue the concept map should help to resolve.” Following this notation, in this study, an integrated 5W1H (i.e., who, what, when, where, why and how) question prompt-based concept mapping approach is proposed to guide students to investigate specified issues via developing a website. Collaboratively constructing concept maps via the Internet can represent the domain knowledge of a project, in which students use “who” to stand for the important persons or organizations related to the issue, “what” to stand for the big events, “when” and “where” to represent the time and location of the events, “why” to explain the reasons for the phenomenon, and “how” to propose possible solutions to the problem. In order to validate the learning approach proposed in this study, the research questions are as follows:

- Is there significant difference between the learning achievements of the students who learned with different concept mapping strategies?
- Is there significant difference between the 5C performances of the students who learned with different concept mapping strategies?
- Is there significant difference between the learning attitudes of the students who learned with different concept mapping strategies?
- Is there any effect of the interaction between the concept mapping strategies and the students’ cognitive styles on their learning achievements?

**Literature review**

**Concept mapping**

Researchers have indicated that concept mapping could bring benefits to students in terms of their learning (Caelli, 1998; Hwang, Kuo, Chen, & Ho, 2014; Weinerth, Koenig, Brunner, & Martin, 2014). For example, Caelli (1998) employed the effect of concept mapping on students’ learning performance in a health care course and found that the use of concept maps assisted them in organizing or realizing abstract concepts, such that the learning became more meaningful. Wu, Hwang, Milrad, Ke, and Huang (2012) developed a concept mapping approach with an instant feedback mechanism to improve students’ learning performance in an ecology course. In the meantime, several studies have revealed the benefits of employing concept maps to organize the information searched on the Internet (Chiou, Huang, & Hsieh, 2004; Hwang, Kuo, Chen, & Ho, 2014); for example, Hwang, Kuo, Chen and Ho (2014) employed concept mapping in a web-based information-searching activity and found that the approach significantly enhanced the students’ learning achievements and perceptions as well as reducing their cognitive load.

On the other hand, researchers have pointed out the difficulties and problems of conventional concept mapping approaches (Cañas, Novak, & Reiska, 2012); for example, Chang, Sung and Chen (2001) indicated the need to offer students hints when engaging them in constructing concept maps individually; Weinerth, Koenig, Brunner and Martin (2014) further indicated that most studies ignored the usability of computerized concept mapping tools. Accordingly, it is necessary to assist students in developing concept maps with a systematic strategy. In this study, a question prompt-based concept mapping approach is therefore proposed to guide students to organize their knowledge in a project-based learning activity that engages students in collecting data from the web for developing a website based on a specified issue.
The 5C competences for the 21st Century

In this study, the 5C competences refer to communication, collaboration, critical thinking, complex problem solving and creativity. Those competences have been emphasized by scholars as important abilities that could determine the competitive ability of a country in the 21st century (Ravitz, Hixson, English, & Mergendoller, 2012; Smaldino, Lowther, & Russell, 2012; Trilling & Fadel, 2009).

Communication skills refer to the ability to understand others’ perspectives and convey ideas in oral or written ways, while collaboration skills refer to the ability to cooperate with others, respecting the value of team members’ existence, and being responsible for team work (Eggen & Kauchak, 2007; Kamarudin, Abdullah, Kofli, Rahman, Tasirin, Jahim, & Rahman, 2012; O’Donnell, Reeve, & Smith, 2007; Pheeraphan, 2013). Vygotsky (1978) indicated that several important competences and types of knowledge, such as spatial awareness, need to be developed by children via communicating and cooperating with people. Van Rooij (2009) further pointed out that project-based learning tasks could engage students in developing collaborative skills as well as communication competences via the provision of a working scenario in which they could make use of knowledge and practice social interactions and collaboration while dealing with complex problems.

Critical thinking refers to the cognition process with purposive judgment for obtaining a logical conclusion or solution (Felder & Brent, 2003; van Gelder, 2005). It has been recognized by educators as an extremely important competence that enables students to have an in-depth understanding of the information they receive; moreover, it enables students to make correct judgments and assertions based on their own interpretations of the evidence they find (Glassner, Weinstock, & Neuman, 2005). Several researchers have further indicated that the use of computer technologies might have great potential for helping students improve their critical thinking performances (Dwyer, Hogan, & Stewart, 2014; Lin, Liu, & Yuan, 2001).

Complex problem solving skills refer to students’ capabilities of clearly defining and interpreting complex problems as well as generating possible strategies and choosing the best one for dealing with the problems (Eggen & Kauchak, 2007). Kamarudin, Abdullah, Kofli, Rahman, Tasirin, Jahim, and Rahman (2012) found that students preferred to solve complex problems in small groups, such that they could share ideas with their peers, generate various strategies from brainstorming, and make decisions about how to deal with the problems. Therefore, they proposed that project-based learning activities, in which students have chances to confirm information related to the given project, and generate and choose possible solutions for the problems, could be a way to situate students in complex problem-solving scenarios.

Creativity represents a complex cognitive process for generating ideas based on several cognitive activities, including imagination, interpretation of the target issues, and the connection of the issue with personal experiences and prior knowledge (Finke, Ward, & Smith, 1992; Lynch, 2000; Simonton, 2000). Gu, Zhang, and Liu (2014) further indicated that prompting interactions among peers, instructors and experts could have positive impacts on students’ creative performances. This implies that situating students in collaborative learning tasks, such as project-based learning activities, could provide them with learning scenarios for facilitating creative thinking.

Learning style

Learning style has long been considered by researchers when assessing the learning performance of students. Turner and Cutrer (2012) pointed out that individuals with unique and diverse learning styles adopt different ways to organize their knowledge. This implies that students with different learning styles would have different learning outcomes. Scholars have defined various learning style classifications, for instance, Felder and Silverman (1988) classified learning into four different styles, active-reflective, sensing-intuitive, visual-verbal, and sequential-global. Active learners are typically interpersonal learners with a tendency to have an adventuresome personality type. They apply the information and instruction received to their individual lives. They prefer to actually experience the subject matter, rather than just gaining knowledge of the material. Reflective learners are typically intrapersonal, logical, mathematical learners with a thinking personality type. Sensing learners tend to be more practical and careful, and like to learn facts; intuitive learners prefer discovering possibilities and relationships, and dislike repetition. Visual learners prefer diagrams, sketches, schematics, etc., when reading course materials, while verbal learners tend to gain information by ways of writing or oral presentation. Sequential learners tend to follow logical stepwise paths while
gaining knowledge, while global learners seem to learn in large jumps, absorbing material almost randomly without seeing connections.

Several studies have further reported the impacts of different learning styles on students' learning performance, interest and motivation (Hwang, Sung, Hung, & Huang, 2013; Yang, Hwang, & Yang, 2013). For example, Jahanbakhsh (2012) found that, for the students whose major was mathematics, sensing-intuitive learning styles had significant correlations with their learning achievement; on the other hand, for speculative science, students' learning achievement had significant correlations with active-reflective learning styles. They also reported that the empirical science students' learning achievements were significantly correlated with both visual-verbal and sequential-global learning styles. Hwang, Sung, Hung and Huang (2013) considered the sequential-global learning styles in developing an educational computer game for an elementary school natural science course and found that the learning achievement and motivation of the style-fit students were significantly better than those of the non-fit students. That is, it is worth taking learning styles into account when developing or adopting learning strategies or tools.

Experiment design

The experiment was conducted using a quasi-experimental method, in which the experimental group learned using a question prompt-based concept mapping approach, while the control group used the conventional concept mapping approach. The CmapTools developed by the Institute for Human and Machine Cognition (IHMC) of the Florida University System (Novak & Cañas, 2006) was employed. The purpose of the study was to investigate the effectiveness of the proposed approach in terms of students' learning achievement, learning attitude, 5C ability, etc. Moreover, the effects of the learning approaches on the learning outcomes of the students with different learning styles were investigated as well. The research framework of this study is depicted in Figure 1.

![Research Framework](image)

Independent variable:
- learning mode
  - (1) Question prompt-based concept mapping approach
  - (2) Conventional concept mapping approach

Independent variable:
- Learning styles
  - active reflective
  - sensing/intuitive
  - visual/verbal
  - sequential/global

Dependent variables
- (1) Learning achievement
- (2) Leaning attitude
- (3) 5C ability
  - Communication
  - collaboration
  - critical thinking
  - complex problem solving
  - creativity

Participants

The participants were grade-two students from four classes of a senior high school in southern Taiwan. The sample was constituted of 176 students (61 males and 115 females), averaging 17 years old. Most of the students had knowledge of concept mapping, but none of them had experience of using the Cmaptools before the experiment. Two classes totaling 88 students were randomly assigned to the experimental group, which used the question prompt-based concept mapping approach. The other two classes were constituted of 88 students who adopted a conventional concept mapping approach as the control group. The students worked collaboratively in groups of four to complete a group webpage and concept map based on the given project.
Measuring tools

The measuring tools in the study include pre- and post-tests of learning achievement, learning attitude as well as the 5C ability questionnaire.

The pre-test and post-test of learning achievement were assessed with 25 multiple choice questions, giving a total score of 100. The pre-test aimed to evaluate the students’ knowledge of the issues before the learning activity, while the post-test aimed to examine their learning achievements in doing the project (i.e., the global warming issue). The test items were selected from the national examination item pool, and were related to the learning content. The test items of the pre- and post-tests were examined by three domain experts with over 10 years of teaching experience in scientific, technological and environmental issues.

The learning attitude questionnaire was developed by Hwang and Chang (2011). It consisted of six items with a five-point Likert rating scheme. The Cronbach’s α value of the questionnaire was 0.91.

The 5C questionnaire consisted of 61 items, including 12, 19, 18, 12, and 10 items for the communication, collaboration, critical thinking, complex problem-solving and creativity dimensions, respectively. The communication dimension originated from the “communicative adaptability” scale developed by Duran (1992) with a Cronbach’s α value of 0.74. The collaboration dimension was developed by Jeng and Tang (2004) with a Cronbach’s α value of 0.88. The critical thinking dimension was developed by Schraw and Dennison (1994) with a Cronbach’s α value of 0.92. The critical thinking dimension originated from the measure developed by Pan and Wang (2003) with a Cronbach’s α value of 0.89. The creativity dimension was modified by Lin and Wang (1994) based on the Creativity Assessment Package proposed by Williams (1980) with a Cronbach’s α value of 0.79. The scoring method of a five-point Likert scale was used, with 1, 2, 3, 4 and 5 respectively representing “strongly disagree,” “disagree,” “neutral,” “agree” and “strongly agree.”

The learning style measure was modified by Hwang, Sung, Hung and Huang (2013) based on the Index of Learning Style (ILS) Questionnaire developed by Soloman and Felder (2001). It consisted of 44 items in four dimensions: active/reflective, sensing/intuitive, visual/verbal and sequential/global. Each dimension contained 11 items. In this study, the active/reflective, sensing/intuitive and sequential/global dimensions were adopted since they might be related to the use of question prompts in helping students develop concept maps. The Cronbach’s α value of the adopted dimensions was 0.74.

Experimental procedure

The experimental procedure is shown in Figure 2. During the first two weeks, the educator provided the students with orientation of how to construct domain knowledge using the project-based learning process, and to understand the project-based learning system created by the educator, as shown in Figure 3. In addition, the pretests of learning achievement, learning attitude and 5C abilities were conducted in the second week. Subsequently, the online course on the issue of global warming was investigated by the students in the 3rd and 4th weeks. Instruction for the concept mapping tool was introduced, and in the 5th week, the students drew a personal concept map of what they had learnt.

For the following seven weeks, the concept map and website for the global warming issue were constructed collaboratively by the students in small groups. The experimental group students were guided by the question prompt-based concept mapping approach during the learning process. Figure 3 shows an example of how the students developed the concept map following the question prompts; Figure 4 shows a web page developed based on the first level concepts in Figure 3. On the other hand, the control group students developed the website based on a conventional concept mapping approach; that is, they were guided to develop the concept map for designing the website structure without following the question prompt procedure. It should be noted that a concept map may also be characterized by the presence of cross-links, although the example given in Figure 3 shows only a hierarchical structure.

Finally, the effectiveness of the proposed approach was examined by asking the students to take the post-test of learning achievement, learning attitude questionnaire, and 5C competence measures.
Experimental group (N=88)
Control group (N=88)

Orientation for the project-based learning activity
Pretest and pre-questionnaires of learning attitude, 5C competences and learning styles

Instruction for the global warming issue

Instruction for the concept map development tool

Project-based learning with question prompt-based concept mapping
Project-based learning with conventional concept mapping

Posttest of the global warming issue

Post-questionnaires of learning attitude and 5C competences

Figure 2. Experimental procedure of the study

Figure 3. Illustrative example of a global warming concept map developed based on the question prompt mechanism
Experimental results

Analysis of learning achievements

No significant difference was found from the t-test results of the pre-test scores of the two groups ($F = 1.01, p = 0.316$), meaning that the two groups’ learning achievements could be considered as equivalent before the learning activity.

To evaluate the effectiveness of the proposed approach, analysis of covariance (ANCOVA) was used to examine the post-test scores of the two groups by excluding the effect of their pre-test scores. As shown in Table 1, it was found that the experimental group had significantly better learning achievement than those in the control group ($F = 7.24, \ p = 0.008$), indicating that the adjusted mean ($M = 74.73$) of the experimental group was better than that of the control group ($M = 71.68$). In addition, the standard deviation ($SD = 6.78$) of the experimental group was smaller than that of the control group ($SD = 8.13$), implying that the degree of dispersion for the experimental group was smaller than that of the control group. This finding shows that the proposed approach could significantly enhance students’ learning achievement.

<table>
<thead>
<tr>
<th>Variance</th>
<th>Group</th>
<th>$N$</th>
<th>Mean</th>
<th>$SD$</th>
<th>Adjusted mean</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning achievement</td>
<td>Experimental</td>
<td>88</td>
<td>74.77</td>
<td>6.78</td>
<td>74.73</td>
<td>7.24**</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>88</td>
<td>71.64</td>
<td>8.13</td>
<td>71.68</td>
<td></td>
</tr>
</tbody>
</table>

*Note. $^{**}p < .01.$*
Analysis of 5C abilities

No significant difference was found in the pre-test scores of the two groups in terms of 5C abilities: communication ($F = 0.14, p = 0.713$), critical thinking ($F = 0.13, p = 0.717$), complex problem solving ($F = 2.16, p = 0.143$), and creativity ($F = 1.46, p = 0.228$), meaning that the four competences could be considered as equivalent before the learning activity. Collaboration, however, presented a violation of homogeneity of the regression coefficients, so the Johnson-Neyman technique (J-N) was employed as an alternative. As seen in Figure 5, the regression of the two groups intersects, with an X value of 4.580. For those students with a score for collaboration ability of over 3.941, there was no significant difference between the two groups. On the contrary, for those who scored below 3.941, the experimental group was significantly better than the control group.

Figure 5. Scatterplot illustrating heterogeneity of regression slopes: treatment-covariate interaction via the Johnson-Neyman technique

Table 2. ANCOVA of the 5C abilities of the two groups

<table>
<thead>
<tr>
<th>Variance</th>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>Experimental</td>
<td>88</td>
<td>3.87</td>
<td>0.51</td>
<td>3.85</td>
<td>5.07*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>88</td>
<td>3.70</td>
<td>0.46</td>
<td>3.71</td>
<td></td>
</tr>
<tr>
<td>Collaboration</td>
<td>Experimental</td>
<td>88</td>
<td>3.90</td>
<td>0.58</td>
<td>3.89</td>
<td>5.76*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>88</td>
<td>3.68</td>
<td>0.56</td>
<td>3.70</td>
<td></td>
</tr>
<tr>
<td>Critical thinking</td>
<td>Experimental</td>
<td>88</td>
<td>3.85</td>
<td>0.54</td>
<td>3.84</td>
<td>7.60**</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>88</td>
<td>3.64</td>
<td>0.51</td>
<td>3.65</td>
<td></td>
</tr>
<tr>
<td>Complex problem solving</td>
<td>Experimental</td>
<td>88</td>
<td>3.89</td>
<td>0.48</td>
<td>3.89</td>
<td>6.04*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>88</td>
<td>3.70</td>
<td>0.53</td>
<td>3.71</td>
<td></td>
</tr>
<tr>
<td>Creativity</td>
<td>Experimental</td>
<td>88</td>
<td>3.85</td>
<td>0.46</td>
<td>3.83</td>
<td>5.29*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>88</td>
<td>3.66</td>
<td>0.52</td>
<td>3.68</td>
<td></td>
</tr>
</tbody>
</table>

Note. *$p < .05$; **$p < .01$. 
Examination of the effectiveness of the approach in terms of the 5C abilities via the ANCOVA method showed (in Table 2) that there were significant differences between the two groups in the aspects of communication ($F = 5.07$, $p = 0.026$), collaboration ($F = 5.76$, $p = 0.017$), critical thinking ($F = 7.60$, $p = 0.006$), complex problem solving ($F = 6.04$, $p = 0.015$), and creativity ($F = 5.29$, $p = 0.023$), indicating that the proposed approach effectively enhanced the 5C abilities of the students in the experimental group.

**Analysis of learning attitudes**

No significant difference was found between the pre-test scores of the two groups ($F = 2.13$, $p = 0.147$), meaning that the two groups’ learning attitudes could be considered as being equivalent before the learning activity.

To evaluate the effectiveness of the proposed approach, analysis of covariance (ANCOVA) was used to examine the post-test scores of the two groups by excluding the effect of their pre-test scores. As shown in Table 3, it was found that the experimental groups had significantly better learning attitude than those in the control group ($F = 10.36$, $p = 0.002$), indicating that the adjusted mean ($M = 4.14$, $SD = 0.58$) of the experimental group is better than that of the control group ($M = 71.68$).

**Table 3. ANCOVA of the learning attitudes of the two groups**

<table>
<thead>
<tr>
<th>Variance</th>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning attitude</td>
<td>Experimental</td>
<td>88</td>
<td>4.12</td>
<td>0.58</td>
<td>4.14</td>
<td>10.36**</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>88</td>
<td>3.87</td>
<td>0.79</td>
<td>3.85</td>
<td></td>
</tr>
</tbody>
</table>

*Note. **$p < 0.01.$

**Interaction effect between concept mapping strategy and learning style**

In Table 4, factor A represents the learning strategies of the two groups, while factor B indicates the learning styles: active and reflective styles, when conducting the two-way ANCOVA method. The result shows that there is no significant interaction effect between learning strategy and learning style ($F = 0.38$, $p = 0.541$). Further verifying the main effect, there is significant difference in the learning achievement ($F = 7.35$, $p = 0.007$), indicating that the adjusted mean ($M = 74.88$) of the experimental group outperformed that ($M = 71.85$) of the control group. Moreover, there is a significant difference between active-style and reflective-style students ($F = 9.63$, $p = 0.002$), showing that the adjusted mean ($M = 75.10$) of the reflective-style students is better than that ($M = 71.63$) of the active-style students, as shown in Table 5.

**Table 4. Results of two-way ANCOVA for learning achievement**

<table>
<thead>
<tr>
<th>Variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>Post hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest of learning achievement</td>
<td>9.06</td>
<td>1</td>
<td>9.06</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Learning strategy (A)</td>
<td>394.61</td>
<td>1</td>
<td>394.61</td>
<td>7.35**</td>
<td>E &gt; C</td>
</tr>
<tr>
<td>Learning style (B)</td>
<td>517.36</td>
<td>1</td>
<td>517.36</td>
<td>9.63**</td>
<td>Reflective &gt; Active</td>
</tr>
<tr>
<td>Interaction (A) * (B)</td>
<td>20.20</td>
<td>1</td>
<td>20.20</td>
<td>0.38</td>
<td></td>
</tr>
</tbody>
</table>

*Note. **$p < 0.01.$ E: Experimental group; C: Control group.

**Table 5. Descriptive statistics of learning mode and learning style (Active/ Reflective style)**

<table>
<thead>
<tr>
<th>Learning mode (A)</th>
<th>Learning style (B)</th>
<th>Active ($n = 96$)</th>
<th>Reflective ($n = 80$)</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjusted mean</td>
<td>SD</td>
<td>Adjusted mean</td>
<td>SD</td>
</tr>
<tr>
<td>Question prompt-based concept map</td>
<td>73.49</td>
<td>1.06</td>
<td>76.27</td>
<td>1.18</td>
</tr>
<tr>
<td>Conventional concept map</td>
<td>69.77</td>
<td>1.06</td>
<td>73.93</td>
<td>1.16</td>
</tr>
<tr>
<td>Sum</td>
<td>71.63</td>
<td>0.75</td>
<td>75.10</td>
<td>0.82</td>
</tr>
</tbody>
</table>

**Students’ feedback to the open-ended questions**

To investigate in depth the factors that might affect the students’ learning performances, the students’ feedback to the open-ended questions was analyzed. Accordingly, it was found that the experimental group students experienced
three different aspects of the learning approach (i.e., the 5W1H question prompt-based concept mapping approach) in comparison with the feedback from the control group.

- **Facilitating cooperation:** Thirty students in the experimental group stated that students in the same team knew how to play their roles following the 5W1H question prompts to search for data and integrate their findings during the website development process. For example, ES17 and ES47 indicated that “the 5W1H prompts guided us to cooperate in a systematic way. While I was searching for data following one of the prompts, my team member would search for data based on other prompts. By cooperating in this way, it is easy to integrate the collected data.

- **Providing clear directions of thinking:** Many students in the experimental group addressed that the 5W1H-based concept mapping provided clear guidance for developing the website. For example, ES12 stated that “The 5W1H-based concept mapping approach enabled me to efficiently select the required information from a large amount of data. ES32 indicated that “5W1H helped me think in the right direction for completing the learning task.”

- **Facilitating thinking from various perspectives:** A portion of students addressed that the 5W1H-based concept mapping helped them view the learning task from different aspects. For example, ES18 stated that “the 5W1H-based concept mapping helped me think about the issue from at least six aspects, which made me see things deeply and clearly.”

**Discussion and conclusions**

This study proposes a question prompt-based concept mapping approach and compares it with the conventional concept mapping approach. The research findings show that the learning achievement, learning attitudes and 5C abilities of students can be enhanced. One possible reason for this achievement could be that the students could easily comprehend the issue of the core knowledge via the question prompt-based concept mapping approach used as scaffolding, and so improved their learning performance. This finding is consistent with previous research (Chiou, Huang, & Hsieh, 2004; Hwang, Kuo, Chen, & Ho, 2014), indicating that appropriate scaffolding to support students in organizing web information could enhance their learning effectiveness. That is, if educators provide prompts to students as scaffolding to construct concept maps, it can support their thinking, revise their knowledge structures, and facilitate knowledge organization (Chang, Sung & Chen, 2001; Gibbs & Habeshaw, 1993).

For the aspect of the 5C abilities, communication, critical thinking, complex problem solving, creativity, and collaboration, the students showed significant enhancement by using the proposed approach based on their effective communication and negotiation. Students are able to train their communication skills with their peers as a result of the group work. In particular, their communication skills can be improved via the proposed approach in the study (Kamarudin, Abdullah, Kofli, Rahman, Tasirin, Jahim, & Rahman, 2012).

Besides, the students in the experimental group had better collaboration skills than those in the control group, which echoes Pheerapan’s (2013) point of view, indicating that the proposed approach offers students an effective collaboration practice based on the given project. Most researchers believe that the project-based learning approach could facilitate students’ collaboration skills and help them acquire specific domain knowledge (Grabe & Grabe, 2001; Markham, Larmer, & Ravitz, 2003; Musa, Mufti, Latiff, & Amin, 2011; Ravitz, Hixson, English, & Mergendoller, 2012). Thus, the study not only offered students project-based learning, but also provided question prompts as scaffolding so that they were able to focus on the important issue, and enhance their communication and collaboration skills with peers.

For the aspect of critical thinking ability, the students in the experimental group were given a specific learning objective and appropriate prompts to solve problems. Previous research has indicated the importance of allowing students to address real-world complex problems through purposeful thinking processes. Especially when filtering a great deal of information, scaffolding with a question prompt-based approach, and critiquing the advantages or disadvantages of peers’ creations can enhance students’ critical thinking and application of knowledge (Sweet, 2009). As for their complex problem-solving ability, students in the experimental group outperformed those in the control group, implying that a systematic learning process would lead students to address the given problem step by step. For example, Eggen and Kauchak (2007) indicated that a clear problem definition and a systematic problem-solving
process would guide students to a successful solution. This study provided obvious rules with 5W1H project-based learning for students. Taking “Global warming” as an example, “What” was used to define global warming, “Why” required students to find out its causes, “Who” was employed to ask students to present the related important figures and organizations, or even policy issues, “Where and when” were used to find out the big events related to global warming which occurred abroad, and “How” asked them to propose possible solutions to the global warming issue. Thus, the question prompt-based concept mapping approach provided a robust learning strategy and scaffolding to facilitate the students’ problem-solving ability, as shown in Figure 6.

![Figure 6. Project-based learning with guided concept map facilitating 5C abilities](image)

It should be noted that the concept mapping tool employed in this study is a widely used tool in most schools in this country and has been adopted in the existing curriculum of the selected school for years. Therefore, the researchers followed the original learning design of the school teacher to reflect the teaching reality in this study. As other knowledge organization tools, such as Web2.0 tools, could provide unique functions, it is worth investigating the effects of integrating these tools with the 5W1H question prompt approach in the future. In addition, to promote the effect of the question prompt-based concept mapping approach, it is also important to analyze students’ learning behaviors for finding the key factors affecting their learning performance.

Acknowledgements

This study is supported in part by the Ministry of Science and Technology of the Republic of China under contract numbers NSC 102-2511-S-011 -007 -MY3 and MOST 104-2511-S-011-001-MY2.

References


