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 The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to Educational Technology & Society.

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

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

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

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

Editorial Assistant
Barbara Adamski, Athabasca University, Canada; Natalia Spyropoulou, University of Piraeus & Centre for Research and Technology Hellas, Greece

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

Assistant Editors
Yuan-Hsuan (Karen) Lee, National Chiao Tung University, Taiwan; Wei-Chieh Fang, National Sun Yat-sen University, Taiwan; Sie Wai (Sylvia) Chew, National Sun Yat-sen University, Taiwan.

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 Jemni, University of Tunis, Tunisia; Mike Joy, University of Warwick, United Kingdom; Athanasios Karoulis, Hellenic Open University, Greece; Paul Kirschner, Open University of the Netherlands, The Netherlands; William Klemm, Texas A&M University, USA; Rob Koper, Open University of the Netherlands, The Netherlands; Jimmy Ho Man Lee, The Chinese University of Hong Kong, Hong Kong; Ruddy Lelouche, Universite Laval, Canada; Tzu-Chien Liu, National Central University, Taiwan; Rory McGreal, Athabasca University, Canada; David Merrill, Brigham Young University - Hawaii, USA; Marcelo Milrad, Växjö University, Sweden; Ricchiro Mizoguchi, Osaka University, Japan; Permanand Mohan, The University of the West Indies, Trinidad and Tobago; Kiyoshi Nakabayashi, National Institute of Multimedia Education, Japan; Hiroaki Ogata, Tokushima University, Japan; Toshio Okamoto, The University of Electro-Communications, Japan; Jose A. Pino, University of Chile, Chile; Thomas C. Reeves, The University of Georgia, USA; Norbert M. Seel, Albert-Ludwigs-University of Freiburg, Germany; Timothy K. Shih, Tamkang University, Taiwan; Yoshiaki Shindo, Nippon Institute of Technology, Japan; Kevin Singley, IBM Research, USA; J. Michael Spector, Florida State University, USA; Slavi Stoyanov, Open University, The Netherlands; Timothy Teo, Nanyang Technological University, Singapore; Chin-Chung Tsai, National Taiwan University of Science and Technology, Taiwan; Jie Chi Yang, National Central University, Taiwan; Stephen J.H. Yang, National Central University, Taiwan; Yu-Mei Wang, University of Alabama at Birmingham, USA.

Executive peer-reviewers
http://www.ifets.info/
Supporting Organizations
Centre for Research and Technology Hellas, Greece
Athabasca University, Canada

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

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

Abstracting and Indexing

Guidelines for authors
Submissions are invited in the following categories:
- Peer reviewed publications: Full length articles (4000 - 7000 words)
- Book reviews
- Software reviews
- Website reviews

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

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

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

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

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

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

**Special Issue Articles**

A Structural Equation Modelling Approach for Massive Blended Synchronous Teacher Training  
Kalpana Kannan and Krishnan Narayanan  
1–15

ET4ET: A Large-scale Faculty Professional Development Program on Effective Integration of Educational Technology  
Sahana Murthy, Sridhar Iyer and Jayakrishnan Warriem  
16–28

The Factors and Impacts of Large-Scale Digital Content Accreditations  
Tony C.T. Kuo, Hong-Ren Chen, Wu-Yuin, Hwang and Nian-Shing Chen  
29–48

**Full Length Articles**

Assessing a Collaborative Online Environment for Music Composition  
Michele Biasutti  
49–63

Examining the Effect of Academic Procrastination on Achievement Using LMS Data in e-Learning  
Ji Won You  
64–74

Gamification in Education: A Systematic Mapping Study  
Daria Dicheva, Christo Dichev, Gennady Agre and Galia Angelova  
75–88

Perceived Effectiveness of Using the Life-Like Multimedia Materials Tool  
Hung-Hsu Tsai, Yen-Shou Lai, Shih-Che Lo and Pao-Ta Yu  
89–99

The Intelligent Robot Contents for Children with Speech-Language Disorder  
Hawon Lee and Eunja Hyun  
100–113

Exploring Application, Attitudes and Integration of Video Games: MinecraftEdu in Middle School  
José-Manuel Sáez-López, John Miller, Esteban Vázquez-Can and María-Concepción Domínguez-Garrido  
114–128

Personal Learning Environments Acceptance Model: The Role of Need for Cognition, e-Learning Satisfaction and Students’ Perceptions  
Salvador del Barrio-Garcia, José L. Arquero and Esteban Romero-Frias  
129–141

Flipping a College Calculus Course: A Case Study  
Alpaslan Sahin, Baki Cavlazoglu and Yunus E. Zeytuncu  
142–152

Effects of Cueing by a Pedagogical Agent in an Instructional Animation: A Cognitive Load Approach  
Hsin I. Yang and Fred Paas  
153–160

Collaborative Professional Development of Mentor Teachers and Pre-Service Teachers in Relation to Technology Integration  
Shih-Hsiung Liu, Hsien-Chang Tsai and Yu-Ting Huang  
161–172

Personal Learning Network Clusters: A Comparison Between Mathematics and Computer Science Students  
Annie Harding and Johann Engelbrecht  
173–184

Make E-Learning Effortless! Impact of a Redesigned User Interface on Usability through the Application of an Affordance Design Approach  
Hyungjoo Park and Hae-Deok Song  
185–196

MONTO: A Machine-Readable Ontology for Teaching Word Problems in Mathematics  
Aparna Lalingkar, Chandrashekar Ramanathan and Srinivasam Ramani  
197–213

Constructing Proxy Variables to Measure Adult Learners’ Time Management Strategies in LMS  
Il-Hyun Jo, Dongho Kim and Meehyun Yoon  
214–225
Orchestration: Providing Teachers with Scaffolding to Address Curriculum Standards and Students’ Pace of Learning  
Anita Díaz, Miguel Nussbaum, Hugo Ñopo, Carolina Maldonado-Carreño and Javier Corredor  
226–239

Towards a Social Networks Model for Online Learning & Performance  
Kon Shing Kenneth Chung and Walter Christian Paredes  
240–253

Enhancing Vocabulary Retention by Embedding L2 Target Words in L1 Stories: An Experiment with Chinese Adult e-Learners  
Zi-Gang Ge  
254–265

Engineering Students Learning Preferences in UNITEN: Comparative Study and Patterns of Learning Styles  
Chen Kang Lee and Manjit Singh Sidhu  
266–281

Development of an Inquiry-based Learning Support System Based on an Intelligent Knowledge Exploration Approach  
Ji-Wei Wu, Judy C. R. Tseng and Gwo-Jen Hwang  
282–300

A Role-play Game to Facilitate the Development of Students’ Reflective Internet Skills  
Wilfried Admiraal  
301–308

Using Mixed-Modality Learning Strategies via e-Learning for Second Language Vocabulary Acquisition  
Fang-Chuan Ou Yang and Wen-Chi Vivian Wu  
309–322

**Book Reviews**

What Connected Educators Do Differently (Authors: Todd Whitaker, Jimmy Casas, and Jeffrey Zoul)  
Reviewer: Christopher Bowen  
323–325

The Chicago Handbook of University Technology Transfer and Academic Entrepreneurship (Editors: Albert N. Link, Donald S. Siegel, and Mike Wright)  
Reviewer: Hadi Shaheen  
326–327

Music and the Making of Modern Science (Author: Peter Pesic)  
Reviewers: Rébecca Guillot and Isabelle Guillot  
328–329
A Structural Equation Modelling Approach for Massive Blended Synchronous Teacher Training

Kalpana Kannan* and Krishnan Narayanan

Department of Humanities and Social Sciences, IIT Bombay, India // kalpanak@cse.iitb.ac.in // knn@iitb.ac.in

*Corresponding author

ABSTRACT

This paper presents a structural equation modelling (SEM) approach for blended synchronous teacher training workshop. It examines the relationship among various factors that influence the Satisfaction (SAT) of participating teachers. Data were collected with the help of a questionnaire from about 500 engineering college teachers. These teachers had attended a blended synchronous distance mode workshop, conducted by IIT Bombay, on “Computational fluid dynamics” in June 2012. The research model was tested using the SEM approach and was found to have a good fit. Ten hypotheses were tested using path analysis. The variable that had the largest effect on Satisfaction (SAT) was Technology effectiveness (TE). The variables that had medium effect on Satisfaction (SAT) were, Behaviour intention to use (BIU), Perceived usefulness (PU), Instructor effectiveness (IE) and Extrinsic motivation (EM). The result on the role of TE in blended synchronous teacher training supports the existing body of literature. Moreover, we found that IE was fully mediated through PU and BIU and thereby implying that TE has a significant role if supported by IE, PU and BIU. Finally, we discuss the implications of our study and make suggestions for further improvement of blended synchronous mode teacher training model.

Keywords

Blended teacher training, ICT enabled workshops, Synchronous distance education, Technology Acceptance Model (TAM), Extrinsic motivation, Intrinsic motivation, Structural equation modelling (SEM)

Introduction

In recent years, Information and Communication Technology (ICT) enabled education is seen as a way to improve the quality of education at all levels. Although ICT enabled education is considered as a necessary step to scale up the educational activities, its implementation is complex. One has to take into consideration many factors, such as availability of technology, time, training and support, coordination and management, individual attitude, belief and motivation, characteristics and ethos of the organization (Tearle, 2004). It also costs a lot of money to build the ICT infrastructure and to maintain them.

Most developing countries have limited infrastructure and fewer experts. Therefore, in future, many programmes are likely to be conducted in the distance mode - in a synchronous or asynchronous fashion. Since 2009, in order to train a large number of engineering college teachers in the country, the Indian Institute of Technology (IIT) Bombay, under the National Mission on Education through ICT (NMEICT, www.sakshat.ac.in), uses ICT to conduct teacher training workshops. These workshops are a first of its kind in the country, conducted in a blended synchronous distance mode. Thousands of teachers are engaged simultaneously through Internet based video conferencing software called AVIEW (www.aview.in) for a period of two-weeks. These workshops are scalable and more than 10,000 teachers can be trained simultaneously. Therefore, understanding the factors associated with such a system are very important for delivering future blended synchronous mode training workshops. The methodology used for conducting blended teacher training workshop is briefly explained in the following section. More details can be found in our earlier studies: Kannan and Narayanan (2011); Kannan and Narayanan (2012).

This paper helps find some critical success factors in a blended synchronous teacher training program. In our study, a research model for a synchronous distance mode teacher training program was developed based on the theoretical framework of Technology Acceptance Model (TAM) by Davis (1989) and the Self Determination Theory (Deci & Ryan, 1985). We redefined the constructs from TAM and added constructs of motivation - intrinsic and extrinsic motivation to explain the relationship between various factors. Although TAM has been tested extensively for technology acceptance by users, to the best of our knowledge, there are hardly any studies carried out so far in India on such a large scale blended synchronous training workshop.

This study analyses the data drawn from 523 participants of a blended synchronous distance mode workshop, conducted for engineering college teachers, in June 2012. Specifically, the subject matter covered the field of
“Computational fluid dynamics” (CFD). In the workshop, there were about 1200 engineering college teachers from about 275 cities in India, representing 550 colleges. Therefore, the sample had a good representation from all parts of the country and from various engineering colleges.

In our study, the effectiveness of the distance mode blended synchronous teacher training workshop was measured with the help of a dependent variable – Participants’ satisfaction (SAT). The independent variables were: Instructor effectiveness (IE), Technology effectiveness (TE), Intrinsic motivation (IM) and Extrinsic motivation (EM). The mediating variables were: Perceived usefulness (PU) and Behaviour intention to use (BIU). The current research focuses on two specific objectives: (1) To test the proposed research model for blended synchronous teacher training workshop using SEM technique, (2) To understand the factors that contribute to effective blended synchronous distance mode training.

Literature review

Technological advances in the last two decades have led to increasing use of technology for teaching and learning. According to Bower et al. (2013), as the technology and bandwidth continue to improve, in future, we may have ubiquitous teaching-learning. Remote participants may be represented in any classroom via video interactions or any mixed mode with a similar experience of a real classroom. As we proceed further towards online education, understanding the factors that influence successful online education can provide useful insights to both teachers and learners.

Critical success factors in online learning

According to Dillon and Guawardena (1995), the three main variables that affect the effectiveness of e-learning environment are: technology, instructor characteristics and student characteristics. According to Willis (1994) and Collis (1991), the instructor plays a central role in the effectiveness and success of e-learning based courses. They believed that it was not the information technology but the instructional implementation of the ICT that determined the effectiveness of e-learning. Volery and Lord (2000) identified three critical success factors in e-learning: Technology (ease of access and navigation, interface design and level of interaction); Instructor (attitude towards students, instructor technical competence and classroom interaction) and previous use of technology from a students’ perspective. According to Papp (2000), e-learning critical success factors included intellectual property, suitability of the course for e-learning environment, building the e-learning course, e-learning course content, e-learning course maintenance, e-learning platform, and measuring the success of an e-learning course. Papp (2000) suggested studying each one of these critical success factors in isolation and also as a composite to determine which factors influence and impact e-learning success.

TAM in online learning

Technology Acceptance Model abbreviated as TAM was proposed by Davis (1989); Davis et al. (1989). TAM was developed based on the theory of reasoned behaviour (TRB) by Fishbein and Ajzen (1975). TAM is widely used in empirical research to explain and predict user acceptance of information system. Literature shows that in education, TAM has been applied by researchers in various studies such as, e-learning and online course delivery. Chen and Tseng (2012) modified TAM for a web-based e-learning system and found that motivation to use and self-efficacy were positively associated with Behaviour intention to use (BIU). Liaw (2008) studied the effectiveness of a Blackboard e-learning system in a university and found that Perceived usefulness (PU) and Perceived satisfaction contributed to BIU e-learning system. Liu et al. (2010) evaluated the web based interactive language learning community among senior high school students in Taiwan and found that Perceived usefulness, Perceived ease of use, and Perceived interaction positively influenced intention to use an online learning community. Teo (2011) tested a model to explain school teachers’ Behaviour intention to use technology. Most of the studies predicted the users’ behaviour intention to use information technology.
Motivation in online environment

As motivation plays an important role in learning, especially in an online environment, Self Determination Theory (SDT) serves as an important framework for addressing motivation in an online environment. According to SDT (Deci & Ryan, 1985), self-determined form of motivation (intrinsic and integrated regulation) may lead to positive outcomes, while the non self-determined form of motivation (Amotivation, external regulation, introjection regulation and identified regulation) may lead to negative outcomes (Deci & Ryan, 1991). SDT has been widely applied across various fields, such as, physical education (Moreno et al., 2010), politics (Losier, et al., 2001), health care (Williams et al., 2006) and general education (Chen & Jang, 2010).

Synchronous and asynchronous interactions

According to Offir et al. (2008), students prefer synchronous to asynchronous mode of interaction, as the transactional distance in asynchronous mode causes poor quality interaction which leads to decreased learning. Educational researchers propose several benefits of using blended synchronous learning approaches. It enables equity of access for learners who are geographically isolated and better completion rate than asynchronous (Norburg, 2012). It also promotes discussion and community learning (Roseth et al., 2013; Lidstone & Shield, 2010).

Satisfaction

The concept of satisfaction is derived from Marketing, where customer satisfaction refers to the state of mind when the expectations of the customer are met from a product or a service (Anisor et al., 2010). Satisfaction is a construct used by many researchers in the past to measure the effectiveness of online learning (Hermans et. al., 2009; Arbaugh, 2000; Kanthawongs, 2011; Marks et al., 2005; Eom et al., 2006). Arbaugh (2000) found that both perceived usefulness and ease of use were positively associated with students' satisfaction in online MBA courses. Kanthawongs (2011) found that learners’ satisfaction with the instructor, perceived ease of use, commitment, and perceived flexibility were positively related to the learners' satisfaction in a web-based ERP course in Thailand. Drawing from the literature in our study, we have taken participants’ satisfaction as a dependent variable.

Teacher training methodology

In order to train a large number of teachers spread across the length and breadth of India, IIT Bombay uses a hub and spoke model. IIT Bombay acts as the hub and a number of remote centres (RCs) located at various engineering colleges in the country act as spokes. RCs are equipped with the necessary infrastructure to accommodate 30 to 60 teachers each. The infrastructure consists of a minimum of 1MBps bandwidth to receive live transmission and two way interaction, open source software AVIEW, different types of cameras, an LCD projector, a classroom, and a laboratory. RCs are located at short distances from several engineering colleges. RCs form a convenient and effective ecosystem for interactions between the college teachers and IIT instructors and for hands on training. There are about 250 such RCs. The model is diagrammatically represented in Figure 1. It can accommodate 10,000 to 15,000 participants at a time.

The training workshop is conducted in two parts: coordinators’ workshop and the main workshop. The coordinators’ workshop is conducted about two months prior to the main workshop, where the workshop coordinators from every remote centre come to the hub (IIT Bombay) for a one week face-to-face workshop. During this workshop, the coordinators are given an orientation on various teaching-learning technologies, the laboratory facility required, and software and hardware required to conduct the main workshop. A common syllabus of the subject to be taught in the main workshop is also finalized. These coordinators help conduct the main workshop in each of the RCs.

The main workshop is conducted for a duration of two-weeks during the winter and summer vacation. Live transmission of lectures and two-way interaction between the hub and RCs happen through an Internet based video conferencing software AVIEW. The screenshot of lectures delivered through AVIEW is given in Figure 2. The
A VIEW screen has three windows - A, B and C. Window-A shows the video of the course instructor. In window-B, the participants’ video is displayed. Window-B has two other modes: the chat window and users’ window, which can be selected at any time. Whenever participants have a query, they can electronically raise their hand through the users’ window. This appears as a question mark against the RC in Window B. The instructor at the IIT Bombay hub can choose this RC and carry out live interactions, which are seen by all RCs. Window-C is used for displaying the presentation slides or as a whiteboard. Laboratory sessions and tutorials are conducted locally at each RC with the help of coordinators. All lectures and live interactions are recorded and the final content is released as open content to the teaching-learning community through T10KT website (www.it.iitb.ac.in/nmeict/home.html).

The blended mode workshop has:
- Synchronous interactions between the hub and RCs through A VIEW and chat
- Asynchronous interactions between instructors, participants and coordinators through discussion forum (Moodle) and email
- Face-to-face and peer group interactions between participants and coordinators in all RCs.

Thus, the model is unique, as the combination of all three interactions makes it very effective and engaging.

Research model and hypotheses

This section describes the research model that was tested using SEM technique. The research model is given in Figure 3. The main components in the research model are given below:
- Online environment factor consists of instructor and technology. Instructor is the subject experts at the hub and technology consists of both synchronous and asynchronous ICTs used for lecture delivery and interactions.
- Participants’ motivation consists of both intrinsic motivation and extrinsic motivation to attend the workshop.
- Participants’ perception consists of perceived usefulness of the workshop and behaviour intention to use/adopt the teaching-learning material in the classroom after the workshop.
- Participants’ satisfaction is a measure of effectiveness of blended synchronous training program.
Operation definitions of constructs

Operation definitions of constructs used in the study are given below:

*Instructor effectiveness (IE)* - Instructor effectiveness was measured in terms of organization of course material, time management, content knowledge, pedagogy and query handing from participants.

*Technology effectiveness (TE)* - Technology effectiveness was measured in terms of quality of audio and video transmission, clarity of live interaction, ease of use of synchronous and asynchronous mode of interaction during the teacher training workshop conducted in the distance mode.

*Participants’ motivation (IM and EM)* - Motivation of participants to attend a distance mode workshop was measured by intrinsic motivation (to learn from subject experts, to learn new teaching methodology and to become accomplished teacher) and extrinsic motivation (certification, promotion, sponsored by government).

*Perceived usefulness (PU)* - Perceived usefulness refers to whether the participating teachers believed that attending the workshop was useful to enhance his/ her subject knowledge and would benefit his/ her teaching performance in future.

*Behaviour intention to use (BIU)* - Behaviour intention to use refers to the participating teachers’ intention to use the knowledge gained from the workshop and adopting the teaching-learning material from the workshop in their teaching.

*Satisfaction (SAT)* - Satisfaction refers to the confirmation of participants’ expectations from the workshop.

**Hypotheses**

From the research model, we formulated the following hypotheses and tested the same using SEM technique.

- H1: Instructor effectiveness will have a positive effect on Perceived usefulness of workshop.
• H2: Technology effectiveness will have a positive effect on Perceived usefulness of workshop.
• H3 and H4: Participants’ motivation will have a positive effect on Perceived usefulness.
• H5: Perceived usefulness will have a positive effect on Behaviour intention to use.
• H6: Behaviour intention to use will have a positive effect on Satisfaction.
• H7: Instructor effectiveness will have a positive effect on Satisfaction.
• H8: Technology effectiveness will have a positive effect on Satisfaction.
• H9 and H10: Participants’ motivation will have a positive effect on Satisfaction.

Figure 3. The research model

Method

In the field of education, we often have many independent variables and dependent variables and need to examine the interplay between the variables. Therefore, the appropriate technique used in explaining such complex models is multivariate analysis using the structural equation modelling (SEM). All hypotheses in our study were tested using SEM.

Instrument

Online survey methodology was adopted to collect data from the participating teachers. We adopted the online survey method, as the number of participants was large and geographically scattered. In order to measure the variables in the study, an instrument was developed. The items that were used to measure the constructs are given in Appendix A. Some items were adopted from prior research work (Young & Lewis, 2008; Davis, 1989) and were changed to suit the context of the study. The questionnaire was first tested in a pilot study and some items were modified and re-worded. All items were measured using a five-point Likert scale from “Strongly disagree = 1” to “Strongly agree = 5.” Literature shows that Likert scale is a popular method used by researchers to measure people’s attitudes, perceptions, preferences, opinions etc. Therefore, we chose Likert scale to measure the items used in the study.
Sample size

In order to achieve adequate power in SEM, the recommended sample size is at least one hundred observations (Hair et al., 2006). We had a total sample size of 923. We randomly split the data into two parts. We performed reliability and exploratory factor analysis (EFA) with the sample of 400 and with the remaining data (523) we performed confirmatory factor analysis (CFA) and tested the structural model.

Data analysis

Data collected through the online questionnaire were analyzed using Statistical Package for Social Sciences (SPSS) version-18 and AMOS version-18.

Demography of the sample

The demography of the sample is tabulated in Table 1. The sample consisted of 85% male and 15% female. Majority of the respondents, about 84% were below 40 years and only about 16% were above 40 years. About 75% were postgraduates, 15% were undergraduates and 10% of the respondents were Doctorates. About 67% had a work experience of less than 10 years and about 33% had worked for more than 10 years. About 38% of the respondents belonged to colleges located in rural areas and about 62% were from colleges situated in urban areas.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>447</td>
<td>85%</td>
</tr>
<tr>
<td>Female</td>
<td>76</td>
<td>15%</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 to 30</td>
<td>219</td>
<td>41.8%</td>
</tr>
<tr>
<td>31 to 40</td>
<td>222</td>
<td>42.4%</td>
</tr>
<tr>
<td>41 to 50</td>
<td>67</td>
<td>12.8%</td>
</tr>
<tr>
<td>above 50</td>
<td>15</td>
<td>3%</td>
</tr>
<tr>
<td>Educational qualification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergraduate</td>
<td>80</td>
<td>15%</td>
</tr>
<tr>
<td>Postgraduate</td>
<td>390</td>
<td>75%</td>
</tr>
<tr>
<td>Doctorate</td>
<td>53</td>
<td>10%</td>
</tr>
<tr>
<td>Work experience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>less than 2 years</td>
<td>108</td>
<td>20.7%</td>
</tr>
<tr>
<td>2 to 5 years</td>
<td>123</td>
<td>23.5%</td>
</tr>
<tr>
<td>5 to 10 years</td>
<td>119</td>
<td>22.8%</td>
</tr>
<tr>
<td>more than 10 years</td>
<td>173</td>
<td>33%</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>201</td>
<td>38.4%</td>
</tr>
<tr>
<td>Urban</td>
<td>322</td>
<td>61.6%</td>
</tr>
</tbody>
</table>

Descriptive statistics

The descriptive statistics of the constructs are shown in Table 2. All means were above the midpoint of 2.5. The standard deviations indicate a narrow spread around the mean. The normality assumption using skewness and kurtosis indices was checked for each of the variables using the criteria given by Kline (2005). According to Kline (2005), lack of normality occurs when absolute value of skewness index is > 3 and absolute value of kurtosis index is > 10. In our sample, the skewness and kurtosis of all the variables were within the expectable range. The skewness ranged between −2.15 and 0.64 and kurtosis ranged between −0.99 and 7.81. Therefore, the normality assumption
was met for the purpose of structural equation modelling. Except for the extrinsic motivation (EM) variable, all other variables had significant correlations among the variables. The correlation matrix is given in Table 2.

**Exploratory factor analysis**

Exploratory Factor Analysis (EFA) on a sample of 400 observations was performed. Principal axis factoring (PAF) method with direct oblimin rotation was used. Four factors were extracted. The total variance explained by the factors was 56.7% and Kaiser-Meyer-Olkin (KMO) value was 0.870. The factor loadings of the factors were in the range 0.384 to 0.848. According to Hair et al. (2006) if the sample size is above 350, factor loadings of 0.3 and above are considered significant for interpreting the results. Therefore, we used the value of 0.3 as the threshold for factor loading. Items with low factor loading were removed from the data set.

### Table 2. Descriptive statistics and correlation matrix

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
<th>Mean</th>
<th>SD</th>
<th>α</th>
<th>IE</th>
<th>TE</th>
<th>IM</th>
<th>EM</th>
<th>PU</th>
<th>BIU</th>
<th>SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>IE</td>
<td>6</td>
<td>4.15</td>
<td>0.53</td>
<td>0.82</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td>5</td>
<td>4.14</td>
<td>0.56</td>
<td>0.81</td>
<td>.454***</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM</td>
<td>3</td>
<td>4.41</td>
<td>0.56</td>
<td>0.62</td>
<td>.295***</td>
<td>.248***</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EM</td>
<td>4</td>
<td>2.73</td>
<td>0.80</td>
<td>0.69</td>
<td>.038</td>
<td>-.085</td>
<td>-.061</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU</td>
<td>3</td>
<td>4.18</td>
<td>0.57</td>
<td>0.72</td>
<td>.568***</td>
<td>.417***</td>
<td>.287***</td>
<td>-.016</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIU</td>
<td>2</td>
<td>4.12</td>
<td>0.61</td>
<td>0.76</td>
<td>.445***</td>
<td>.351***</td>
<td>.242***</td>
<td>-.033</td>
<td>.614***</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td>3</td>
<td>3.94</td>
<td>0.60</td>
<td>0.69</td>
<td>.337***</td>
<td>.362***</td>
<td>.197***</td>
<td>.105**</td>
<td>.35***</td>
<td>.288***</td>
<td></td>
</tr>
</tbody>
</table>

**Note.*** Correlation is significant at the .01 level (2-tailed), ** Correlation is significant at the .05 level (2-tailed)

SD − Standard deviation, α − Cronbach's alpha, IE − Instructor effectiveness, TE − Technology effectiveness, PU − Perceived usefulness, IM − Intrinsic motivation, EM − Extrinsic motivation, BIU − Behaviour intention to use, SAT – Satisfaction.

### Reliability and validity

In our study, we established the reliability of the instrument with the help of internal consistency. Cronbach’s alpha (α) was calculated for all the constructs in the study. Cronbach’s alpha values ranged from 0.62 to 0.82. Except for the Intrinsic motivation construct (α = 0.62), all other constructs had satisfactory internal consistency (α ≥ 0.70). One of the reasons for low value of (α) for Intrinsic motivation construct could be high mean and low standard deviation.

### Table 3. Convergent and discriminant validity

<table>
<thead>
<tr>
<th>Construct</th>
<th>CR</th>
<th>AVE</th>
<th>IE</th>
<th>TE</th>
<th>IM</th>
<th>EM</th>
<th>PU</th>
<th>BIU</th>
<th>SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor effectiveness (IE)</td>
<td>0.85</td>
<td>0.501</td>
<td>0.707</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology effectiveness (TE)</td>
<td>0.83</td>
<td>0.500</td>
<td>0.206</td>
<td>0.707</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrinsic motivation (IM)</td>
<td>0.74</td>
<td>0.500</td>
<td>0.087</td>
<td>0.061</td>
<td>0.707</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extrinsic motivation (EM)</td>
<td>0.79</td>
<td>0.500</td>
<td>0.001</td>
<td>0.007</td>
<td>0.003</td>
<td>0.707</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived usefulness (PU)</td>
<td>0.74</td>
<td>0.497</td>
<td>0.322</td>
<td>0.173</td>
<td>0.087</td>
<td>0.000</td>
<td>0.704</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behaviour intention to use (BIU)</td>
<td>0.83</td>
<td>0.495</td>
<td>0.198</td>
<td>0.123</td>
<td>0.058</td>
<td>0.076</td>
<td>0.706</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction (SAT)</td>
<td>0.74</td>
<td>0.498</td>
<td>0.113</td>
<td>0.131</td>
<td>0.038</td>
<td>0.011</td>
<td>0.122</td>
<td>0.082</td>
<td>0.707</td>
</tr>
</tbody>
</table>

**Note.** Composite reliability (CR) = (Square of the summation of factor loadings) / (square of the summation of factor loadings) + (summation of error variances). Average variance extracted (AVE) = (Summation of the square of factor loadings) / (summation of the square of factor loadings) + (summation of error variances). Diagonals represent the square root of average variance extracted and other entries represent squared correlations.

The face validity or content validity of the instrument was done by three experts and through pilot testing. The construct validity was established by convergent and discriminant validity. According to Hair et al. (2006) convergent validity is established with composite reliability (CR) and average variance extracted (AVE). Convergent validity is judged to be adequate when composite reliability exceeds 0.70 and average variance extracted exceeds
0.50 (Hair et al., 2006). As shown in Table 3, CR and AVE for all the constructs were adequate. Thus, convergent validity for all the constructs was established.

Discriminant validity is established when the variance shared between a construct and any other construct in the model is less than the variance that a construct shares with its own indicators (Fornell & Larcker, 1981). To get satisfactory discriminant validity, the square root of AVE for each construct should be greater than the squared correlation between the constructs (Fornell & Larcker, 1981). The diagonals in Table 3 represent the square root of average variance extracted and other entries represent squared correlations. Table 3 shows acceptable discriminant validity between each pair of constructs. Thus, discriminant validity for all the constructs was established.

Measurement model

Confirmatory factor analysis (CFA) for the measurement model was performed using Maximum Likelihood Estimation (MLE) method using AMOS version 18. A total of 26 items of the measurement model were subjected to CFA. The initial estimation model was modified with the help of modification index with a few error terms in the model to covary. The standardized factor loadings ranged from 0.46 to 0.78 and were significant at \( p < 0.001 \) (\( t \)-values ranged from 4.9 to 12.7). The Goodness of fit (GOF) results of the measurement model was found within the acceptable range with \( \chi^2/df = 2.5 \) (\( \chi^2 = 725, df = 281, p = .000 \), \( CFI = 0.87, GFI = 0.91, AGFI = 0.88, TLI = 0.85 \), all values were ≥ 0.85 (recommended value for GOF is ≥ 0.9) and \( RMSEA = 0.05 \), recommended value is ≤ 0.10. As most of the observed GOF indices satisfied the recommended values of the indices, the results of CFA suggests that the measurement model had a good fit.

Structural model

The structural model was tested using AMOS. The results showed that the model was recursive with 351 distinct sample moments and 70 distinct parameters to be estimated. The model was identified and it had 277 degrees of freedom. The fit indices indicated that the model had a good fit. The \( \chi^2 \) statistics was 494 and it was significant (\( p = .000 \)). The \( \chi^2 \) value should be non-significant, but in our model due to the large sample size it was significant. Therefore, we checked the value of \( \chi^2/df \). The \( \chi^2/df \) was 1.7 (recommended value is ≤ 3), implying that the observed covariance matrix and implied covariance matrix were statistically identical.

The \( GFI \) and \( AGFI \) were 0.93 and 0.91 respectively (the recommended value is ≥ 0.90) (Hair et al., 2006). \( CFI \) and \( TLI \) were 0.93 and 0.92 respectively. The parsimonious fit index \( PNFI \) and \( PCFI \) were 0.74 and 0.79 respectively (recommended value is ≥ 0.6). The \( RMSEA \) value was 0.03 (recommended value is less than 0.05 for good fit and ≤ 0.10 for moderate fit). The results of the model indicated that overall the model had a good fit as all the indices were within the recommended values.

Results

Ten hypotheses were tested using path analysis. The regression weights for the parameters along with the critical ratios are given in Table 4. The research model with the standardized path coefficients and the explanatory power \( R^2 \) for each dependent variable is displayed in Figure 4. Except for four paths, all other paths were statistically significant. The results of the model supported H1, H5, and H8 with path coefficients of 0.57, 0.89 and 0.36 (\( p \leq .001 \)) respectively and H2, H6 and H10 with path coefficients of 0.19, 0.18 and 0.14 (\( p \leq 0.05 \)) respectively. However, the results indicated rejecting H3, H4, H7 and H9.

A coefficient linking one construct with another directly in the path diagram (Figure 4) represents the direct effect. An indirect effect is represented when the determinants affect the endogenous variable through one or two mediating variables. Table 5 shows standardized total effects, direct and indirect effects associated with each of the variables in the model. The total effect on a variable is the sum total of respective direct and indirect effects. In our model, the standardized total effects of predictor variables on the dependent variables ranged from 0.00 to 0.89.
Three endogenous variables were tested in the model. In terms of the explanatory power, the model adopted in the study explained 58% of variance in Perceived usefulness, 80% of variance in Behaviour intention to use and 34% of variance in Satisfaction.

Among the three endogenous variables, the highest amount of variance (80%) was explained by the determinants of Behaviour intention to use (BIU). The most dominant determinant was Perceived usefulness (PU) with a total effect of 0.891, followed by Instructor effectiveness (IE) with a total effect of 0.509 and Technology effectiveness (TE) with a total effect of 0.171. Both Intrinsic motivation (IM) and Extrinsic motivation (EM) had insignificant effect on BIU.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Path</th>
<th>Path coefficient (β)</th>
<th>Critical ratio</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>IE → PU</td>
<td>0.57</td>
<td>5.8***</td>
<td>Supported</td>
</tr>
<tr>
<td>H2</td>
<td>TE → PU</td>
<td>0.19</td>
<td>2.3*</td>
<td>Supported</td>
</tr>
<tr>
<td>H3</td>
<td>IM → PU</td>
<td>0.11</td>
<td>1.7 (n.s.)</td>
<td>Not supported</td>
</tr>
<tr>
<td>H4</td>
<td>EM → PU</td>
<td>0.001</td>
<td>0.02 (n.s.)</td>
<td>Not supported</td>
</tr>
<tr>
<td>H5</td>
<td>PU → BIU</td>
<td>0.89</td>
<td>9.5***</td>
<td>Supported</td>
</tr>
<tr>
<td>H6</td>
<td>BIU → SAT</td>
<td>0.18</td>
<td>2.1***</td>
<td>Supported</td>
</tr>
<tr>
<td>H7</td>
<td>IE → SAT</td>
<td>0.05</td>
<td>0.49 (n.s.)</td>
<td>Not supported</td>
</tr>
<tr>
<td>H8</td>
<td>TE → SAT</td>
<td>0.36</td>
<td>3.6***</td>
<td>Supported</td>
</tr>
<tr>
<td>H9</td>
<td>IM → SAT</td>
<td>0.11</td>
<td>1.5 (n.s.)</td>
<td>Not supported</td>
</tr>
<tr>
<td>H10</td>
<td>EM → SAT</td>
<td>0.14</td>
<td>2.3*</td>
<td>Supported</td>
</tr>
</tbody>
</table>

Note. *** indicates $p \leq .001$. ** indicates $p \leq .01$. * indicates $p \leq .05$. (n.s.) indicates non-significant.

Table 5. Direct, indirect and total effects of the research model

<table>
<thead>
<tr>
<th>Endogenous variable</th>
<th>Determinant</th>
<th>Standardized estimates</th>
<th>Direct</th>
<th>Indirect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived usefulness ($R^2 = 0.58$)</td>
<td>IE</td>
<td>0.572</td>
<td>—</td>
<td>0.572</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.192</td>
<td>—</td>
<td>0.192</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IM</td>
<td>0.113</td>
<td>—</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EM</td>
<td>0.001</td>
<td>—</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Behaviour intention to use ($R^2 = 0.80$)</td>
<td>IE</td>
<td>—</td>
<td>0.509</td>
<td>0.509</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>—</td>
<td>0.171</td>
<td>0.171</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IM</td>
<td>—</td>
<td>0.101</td>
<td>0.101</td>
<td></td>
</tr>
</tbody>
</table>
The model explained 58% of variance in Perceived Usefulness (PU). Among all the variables, Instructor effectiveness (IE) had the largest effect, with total effect of 0.572; Technology Effectiveness (TE) had a total effect of 0.192. However, Intrinsic Motivation (IM) with a total effect of 0.113 and Extrinsic Motivation (EM) with a total effect of 0.001 had statistically insignificant effect on Perceived Usefulness (PU). This implies that the participants would find the course useful if the instructor is very effective in making the subject matter interesting, easy to understand and interacts regularly with the distant participants.

The model explained about 34% of variance in Satisfaction (SAT). The most dominant determinant was Technology Effectiveness (TE) with a total effect of 0.392 followed by Behaviour intention to use (BIU) with a total effect of 0.185. Among other variables, Perceived Usefulness (PU) had a total effect of 0.165, Instructor Effectiveness (IE) had a total effect of 0.144, Extrinsic Motivation (EM) had a total effect of 0.143. However, Intrinsic Motivation (IM), with a total effect of 0.134, was insignificant as both direct and indirect paths were non-significant.

Discussion, limitations and future work

The results imply that Technology Effectiveness (TE) defined in terms of “Quality of audio and video transmission, clarity of live interaction and ease of use of synchronous and asynchronous mode of interaction” during the teacher training workshop emerged as the most important factor for Satisfaction. If one improves the technology component by one point, then it will result in 0.39 standard deviation increase in the satisfaction level of the participants. Instructor Effectiveness (IE) had an indirect effect (0.094) through the mediating variables PU and BIU on participants’ satisfaction, which was more than the direct effect (0.05). Therefore, we can infer that Instructor Effectiveness (IE) was fully mediated through PU and BIU. Extrinsic Motivation (EM) had a total effect of 0.143 on Satisfaction, which was entirely a direct effect. This implies that some kind of incentive, such as certification, sponsorship is required for the participating teachers for greater satisfaction. However, Intrinsic Motivation (IM) had insignificant effect on Satisfaction, as there was not much variation in the participants’ response that contributed to the variance.

The results of our study were consistent with the TAM model proposed by Davis & Venkatesh (1996), where PU is positively related BIU. In our study, PU and BIU had a statistically significant correlation ($\beta = 0.89, p \leq .001$). The findings of our study were consistent with studies done by Ramayah and Lee (2012) in Malaysia, where system quality positively affected user satisfaction in an e-learning environment. The findings were also consistent with the study done by Eom et al. (2006), where the course structure, self-motivation, instructor feedback, interaction, instructor facilitation had a positive effect on students’ satisfaction in an online course.

The importance of technology component in a blended synchronous mode is stressed by other researchers in the literature. According to Bower et al. (2013), the main issues that instructors confronted when teaching blended synchronous lessons were communication issues and issues related to cognitive overload. According to White et al. (2010), the success of blended synchronous learning activities depends on the quality of audio, use of teaching associates and support staff. Our study confirms that in a blended synchronous mode teaching-learning, the quality of audio-video, clarity of live interaction and ease of use of technology were important factors.

Although care was taken to ensure that the methodology used in the study was sound, there were some limitations. First, as data were collected online, just after the workshop through a self-rating questionnaire, some participants could have given higher ratings to some of the variables, which could have affected the results. Second, Intrinsic Motivation (IM) construct had a low Cronbach’s alpha (0.62); therefore, some more items need to be added or
modified. Third, the six variables in the model explained only 34% of variance in Satisfaction leaving 66% unexplained. The findings of the study suggest that there could be other variables in the blended synchronous training that could influence the Satisfaction. Nevertheless, such an attempt to systematically evaluate the effectiveness of a blended synchronous teacher training program has not been done so far in the Indian context. To that extent, this paper makes an important contribution.

In order to further improve the design and implementation of blended synchronous training, we looked at the qualitative feedback from the participants. Some of the feedback was:

- “This experience of distance learning is good and innovative.”
- “It is my first experience of distance education and it is good to see that understanding is as good as the lectures conducted in a lecture hall in the conventional manner.”
- “The RC facility was excellent and the centre coordinator was very helpful. He was knowledgeable and we could clarify most of the doubts.”
- “Laboratory session was very helpful for understanding the topic which was taught during the theory session.”
- “During the online quiz there was bandwidth problem in our RC.”
- “Sometimes there was problem with the audio and video.”
- “Facilities at my RC were not good.”

From the participants’ feedback, we found that in some RCs the workshop coordinators were very helpful and knowledgeable. They helped the participants in clearing doubts during laboratory sessions and encouraged discussions among the participants. While in some other RCs, the coordinators did not take much initiative. Facilitating conditions at the RCs also varied from one centre to another. Therefore, we infer that workshop coordinators, facilitating conditions at RC and peer group interactions could also influence satisfaction in the blended training model. Therefore, in future, one may consider introducing the following additional factors in the SEM: (1) Coordinator effectiveness (Educational qualification, subject knowledge, motivation, query handling, workshop management and coordination) (2) Facilitating conditions at RC (Audio-video setup, laboratory facility, Internet bandwidth and support staff) (3) Peer group interactions (formal and informal discussions with peers in RCs).

Conclusion

In this paper, a research model for a blended synchronous mode teacher training program was developed based on the theoretical framework of TAM (Davis, 1989) and Self Determination Theory (Deci & Ryan, 1985). The research model was tested with empirical data from 523 teachers, who had participated in a blended synchronous workshop in June 2012. Ten hypotheses were tested with the help of path analysis. Six hypotheses were supported and four hypotheses were not supported. The model adopted in the study explained 58% of variance in Perceived usefulness (PU), 80% of variance in Behaviour intention to use (BIU), and 34% of variance in Satisfaction (SAT). The results imply that in a blended synchronous mode teacher training the technology factor is the most significant factor that contributes to the participants’ satisfaction.

This study has many implications for educational institutions, administrators and policy makers who are involved in ICT enabled education. First, the study shows that Technology effectiveness is an important factor, the technology that is used for blended synchronous mode has to be good and easy to use for the participants. Second, as PU and BIU had a statistically significant correlation, it implies that participants of the workshop are likely to adopt the teaching-learning material in their teaching if they find the training workshop to be useful and relevant. Third, as Extrinsic motivation (EM) had a direct effect on Satisfaction, it implies that some kind of incentive, such as certification, sponsorship etc., would increase the satisfaction level of the participants. Fourth, apart from the six factors considered in the research model for blended synchronous mode training, coordinator effectiveness, facilitating conditions at RC, and peer group interactions could also influence participants’ satisfaction.

Acknowledgements

We would like to thank Prof. D. B. Phatak, Principal Investigator of the project, and Prof. A. Sharma and Prof. B. Puranik, Instructors of the workshop. Prof. P. Purang gave valuable inputs while writing the paper. Financial support
for conducting the workshop was made available by a grant from the NMEICT, Ministry of Human Resource Development, Government of India. AVIEW support was provided by Amrita Vishwa Vidyapeetham.

References


## Appendix A

<table>
<thead>
<tr>
<th>Construct</th>
<th>Type</th>
<th>Item</th>
<th>Statement</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor effectiveness (IE)</td>
<td>Independent variable</td>
<td>IE1</td>
<td>Instructor used the time effectively to meet workshop objectives</td>
<td>Young &amp; Lewis (2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IE2</td>
<td>The course material was well organized</td>
<td>-do-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IE3</td>
<td>Instructor gave effective examples, illustrations and guidelines</td>
<td>-do-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IE4</td>
<td>Instructor was able to explain difficult concepts</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IE5</td>
<td>Instructor used innovative teaching methods</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IE6</td>
<td>Instructor was responsive to participants' queries</td>
<td>Young &amp; Lewis (2008)</td>
</tr>
<tr>
<td>Technology effectiveness (TE)</td>
<td>Independent variable</td>
<td>TE1</td>
<td>The quality of audio transmission was good</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE2</td>
<td>The quality of video transmission was good</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE3</td>
<td>The live interaction session with the instructor was clear and understandable</td>
<td>Davis (1989)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE4</td>
<td>It was easy to interact with the instructor through ‘AVIEW’ technology (synchronous)</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE5</td>
<td>I found it easy to interact with fellow participants through “Moodle” (asynchronous)</td>
<td>self</td>
</tr>
<tr>
<td>Intrinsic motivation (IM)</td>
<td>Independent variable</td>
<td>IM1</td>
<td>To learn new teaching methodology</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IM2</td>
<td>To become an accomplished teacher</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IM3</td>
<td>To learn from subject expert</td>
<td>self</td>
</tr>
<tr>
<td>Extrinsic motivation (EM)</td>
<td>Independent variable</td>
<td>EM1</td>
<td>For certification</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EM2</td>
<td>It will help in promotion</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EM3</td>
<td>Management of your institute asked you to attend</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EM4</td>
<td>Because, the course is fully sponsored</td>
<td>self</td>
</tr>
<tr>
<td>Perceived usefulness (PU)</td>
<td>Dependent (Mediating)</td>
<td>PU1</td>
<td>I found the lectures to be useful for my understanding of the subject</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PU2</td>
<td>I found the lab/ tutorials to be useful for my understanding of the project</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PU3</td>
<td>The workshop would enable me to teach better</td>
<td>self</td>
</tr>
<tr>
<td>Behaviour intention to use (BIU)</td>
<td>Dependent (Mediating)</td>
<td>BIU1</td>
<td>I will adopt the material from the lectures in my teaching</td>
<td>self</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BIU2</td>
<td>I will adopt the material from the lab/ tutorial in my teaching</td>
<td>self</td>
</tr>
<tr>
<td>Satisfaction (SAT)</td>
<td>Dependent variable</td>
<td>SAT1</td>
<td>I enjoyed the distance mode workshop</td>
<td>Young and Lewis (2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SAT2</td>
<td>I would recommend such distance mode workshops to others</td>
<td>-do-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SAT3</td>
<td>The workshop provided a good learning experience</td>
<td>self</td>
</tr>
</tbody>
</table>
ET4ET: A Large-Scale Faculty Professional Development Program on Effective Integration of Educational Technology

Sahana Murthy*, Sridhar Iyer and Jayakrishnan Warriem
Indian Institute of Technology Bombay, Mumbai 400076, India // sahanamurthy@iitb.ac.in // sri@iitb.ac.in // jayakrishnan.m@iitb.ac.in

*Corresponding author

ABSTRACT

Educators have recommended that the affordances provided by ICTs should be used to promote student-centered constructivist learning. While the actual use of ICT in education has increased, not much change has occurred in terms of the pedagogical practices followed. Information transmission models of teaching are still being followed, albeit with the use of ICT. Efforts at the school level attempt to address this problem via teacher education programs that emphasize ICT integration with constructivist practices. In higher education settings, there are fewer reported studies on ICT integration at a systemic level, and most decisions related to the use of ICTs are left to the instructor. In this paper, we describe the design, implementation and evaluation of ET4ET, a large-scale faculty professional development program (1138 participants) on effective integration of educational technology for engineering college instructors in India. Guided by the Attain-Align-Integrate (A2I) model, the ET4ET program prepares instructors to implement ICT supported student-centric teaching strategies. To ensure engagement and learning of the participants, active learning strategies are implemented throughout the ET4ET program. This paper traces the development of ET4ET, from a small-scale face-to-face implementation to a large-scale mode mediated by technology, and suggests guidelines for similar training programs.

Keywords
Teacher professional development, ICT integration, Large-scale, Constructivist alignment, Active learning

Introduction

The proliferation of information and communication technologies (ICT) has led to its widespread use in classrooms around the world in the last two decades. Educators and researchers have argued that the affordances provided by these technologies should be used to promote student-centric collaborative learning environments based on constructivist approaches (Howland, Jonassen & Marra, 2012). While the actual use of ICT in education has increased (Greenhow, Robelia, & Hughes, 2009), a number of challenges related to ICT supported constructivist teaching have been reported, such as teachers feeling inadequately prepared to use technology and implement new instructional strategies (Mumtaz, 2000; Brown & Warschauer, 2006), teachers’ beliefs and attitudes towards technology (Ertmer, 2005) and constraints beyond an individual teacher’s scope, such as curriculum and examination demands (Lim & Chai, 2008). As a result, student-centric active learning approaches with ICT are still not common, and the use of technologies is often limited to information transmission (Gao, Choy, Wong, & Wu, 2009; Lim & Chai, 2008).

Most teacher education programs include ICT as part of their courses and there are recommendations to include ICT in professional development programs (Lawless & Pellegrino, 2007). Of these, many efforts focus on the teaching of the skills of handling tools (Friedman & Kajder, 2006), however such skills alone does not prepare teachers to successfully integrate technology (Mishra, Koehler, & Kereluik, 2009). Current efforts emphasize the integration of ICT with constructivist pedagogical practices (Howland, Jonassen & Marra, 2012), and some focus on teachers’ development of TPACK (Chai, Koh & Tsai, 2010). Most of these efforts are for pre- or in-service teachers at the school level, but there are fewer systemic efforts at the college and university level (Schaefer & Utschig, 2008).

At the tertiary education level, decisions for ICT integration is often left to the individual instructors, leading to problems such as ineffective use of the tool (Selwyn, 2007), isolation and inability of individuals to find know-how (Conole, Dyke, Oliver & Seale, 2004), and lack of percolation of good practices (Ebert-May, Derting, Hodder, Momsen, Long & Jardeleza, 2011).

The above problems are compounded when teacher education programs have to be implemented at a large-scale. Complexities due to the scale include the availability of infrastructure, diversity in operating conditions, and resources available and ensuring engagement and learning. An attempt to address this problem is the Teach 10000...
Teachers (T10KT) project in India, which promotes large-scale professional development programs for university faculty in various topics. The T10KT project provides the infrastructure for conducting such programs via synchronous remote classrooms, a blended learning solution. Within this project, we have developed Educational Technology for Engineering Teachers (ET4ET), a professional development program on effective integration of educational technology for engineering college faculty. ET4ET is a large-scale program guided by the A2I (Attain-Align-Integrate) model (Warriem, Murthy, & Iyer, 2014) that originated from a research project. The main objective of ET4ET is to prepare engineering college faculty across the country to implement ICT supported student-centric teaching strategies. The content and its sequencing in the ET4ET program are chosen to address this objective, while taking into account the diversity of conditions and resources. To ensure engagement and learning of the participants in ET4ET, active learning strategies are implemented throughout the program.

In this paper we begin with a brief background of the Teach 1000 Teachers project. We review related work on teacher training programs emphasizing integration of ICT in constructive learning environments, and large-scale projects on use of ICT in teaching. We then describe how the A2I model led to the design and deployment of the ET4ET training program, and present data and analysis for the evaluation of the program. We reflect on the key factors that led to the effectiveness of the program and pitfalls to avoid. The contributions of this paper are: A flexible design for training programs that share the similar goal of effective ICT integration, but may vary in duration or choice of specific content, and recommendations based on our experience, that might benefit others who wish to implement similar programs.

Background and national context

The Teach 10000 Teachers (T10KT, http://www.it.iitb.ac.in/nmeict/About_T10kT.html) project is part of a national initiative by the Indian government, the National Mission of Education through ICT (NMEICT, www.nmeict.ac.in). The goal of the T10KT project is to enhance the teaching skills of engineering college faculty. For this, 2-4 week training programs are conducted on the teaching of various engineering topics. To scale up participation, ICT is used to provide synchronous and asynchronous instruction. Participants attend the workshop at one of the 200+ remote centres across the country. The workshops contain lectures by experts, which are transmitted synchronously in the remote centres and include live two-way audio-visual interaction. Typical workshops also contain tutorials and labs conducted by “remote centre co-ordinators.” In addition, Moodle is used for asynchronous interaction, such as for assignments and quizzes. All workshop materials, including slides, assignments and videos of the lectures, are released in open source.

T10KT has conducted close to 20 workshops, each consisting of 1000 to 9000 participants, and has trained 85,000 participants till date. Most training programs under T10KT are on specific domain-based topics, in various fields of engineering (such as Thermodynamics for Mechanical Engineering, Electronics for Electrical Engineering). An important need recognized was for a training program for engineering faculty focusing on pedagogical practices, and use of ICT in their teaching. ET4ET fills this gap by providing training on research based student-centric teaching practices for effective integration of ICT.

Related work

Most pre-service teacher education programs in colleges and universities include courses on the use of educational technology. Pre-service education has been reported to be successful in teachers’ self-efficacy with ICT use (Lee, Chai, Teo, & Chen, 2008) and found to have a strong impact on teachers’ future use of ICT (Gao et al., 2009). However, problems with pre-service programs have been reported such as lack of exposure to pedagogical strategies to be used with ICT (Brown & Warschauer, 2006) and focus of programs mainly on ICT skills (Friedman & Kajder, 2006). Programs have been redesigned to address these limitations, especially with the help of theoretical frameworks like TPACK (Mishra & Koehler, 2006). The TPACK framework provides a means to understand and describe the different types of knowledge – technological, pedagogical and content knowledge - and their synthesis, which are needed by a teacher to achieve meaningful ICT integration.
In recent years, large-scale initiatives on ICT integration in education are being implemented all over the world. Some projects, like ITT in Chile are national or regional government led initiatives (Brun & Hinostroza, 2014). The Intel Teach to the Future project (Intel, 2014) is a privately funded project implemented in over 35 countries. It uses a Train-the-Trainer model for K-12 teachers to integrate technology effectively into classrooms and promote hands-on, collaborative and problem-solving methods.

Colleges and universities too are becoming increasingly aware of the need to train their faculty in effective ICT integration. For example, the Xanadu project at University of Turin uses a gradual approach to introducing university instructors to technology-enhanced learning at various levels from basic to advanced (Trentin, 2006). Another example is the National Effective Teaching Institute (NETI) program’s workshops at North Carolina University that are organized in basic and advanced stages, focusing on learning styles, outcomes, research based instructional strategies and evaluation (Brent & Felder, 2009).

Need for ET4ET

While there exist programs for training university faculty in ICT integration, their implementation designs are not suitable for direct adaptation in our setting, due to differences in operating contexts such as learner requirements or duration of program. For our context, we require a short-term program design whose goal is to train engineering college faculty across the country to implement ICT supported student-centric teaching strategies. The training program design should be able to address not only broadly recommended practices of ICT integration, but also specifically target technologies and pedagogical strategies suitable for engineering topics. This goal is addressed by the application of the A2I model to a large-scale setting.

Attain-Align-Integrate (A2I) Model

The A2I model was developed by the authors of this paper (Warriem et al., 2014) to realize the need for a framework to design short-term training programs whose goal is to enable student-centered learning with ICT integration. The A2I model prescribes features of the training program design, namely, key topics for the above goal, their organization and sequence, and format of activities to be done by the participants.

Theoretical basis

The major theoretical basis of the A2I model that helps decide the topics of the training program is constructive alignment (Biggs, 1996). This is achieved when the teaching-learning activities and evaluation are aligned with the intended student learning outcomes. Constructive alignment also ensures that instructors utilize more constructivist, learner-centered practices while performing this alignment. Constructive alignment has been successfully employed by instructors in course redesign and is known to promote deep learning among students (Wang, Su, Cheung, Wong, & Kwong, 2013).

Spiral curriculum (Bruner, 1977) forms the basis of the organization and sequence of topics in training programs guided by the A2I model. Spiral curriculum is characterized by an iterative process of revisiting content, with each successive visit addressed at a greater depth for learners to build on their initial understanding.

Active learning (Prince, 2004) forms the basis of the pedagogical strategies followed in training programs based on the A2I model. Based on constructivist teaching-learning philosophy, active learning encompasses several research-based strategies designed to engage students in the learning process, in which students go beyond listening, copying of notes, and execution of prescribed procedures (Meltzer & Thornton, 2012). The activities within the A2I-based training programs are designed using active learning strategies so that instructor-participants not only get engaged in the learning environment of the program, but also get exposed to student-centric strategies which they may then try in their own classrooms.
Features of A2I model

To facilitate constructive alignment by instructors, the A2I model prescribes three core modules in the training program - learning objectives, student-centered instructional strategies and assessment strategies related to ICT use. The actual topics for the training can then be selected based on these three core modules and the operating context.

The A2I model consists of a 3-phase design – “Attain,” “Align,” and “Integrate” - to provide a mechanism to connect the three core modules. In the Attain phase, participants are expected to attain preliminary knowledge of topics in each of the three modules, such as how to construct a learning objective, or what are the pedagogical affordances of the tool. The Align phase focuses on pairwise alignment between the modules, for example, writing assessment questions consistent with different levels of learning objectives, or choosing an ICT tool and associated teaching strategy for a higher-order learning objective. The Integrate phase consists of constructive alignment of the three modules. The focus is on inter-connection of the modules so that teachers create instructional and assessment strategies for their course, which are aligned with the intended student learning outcomes. The instructional strategies chosen are constructivist-oriented and are implemented along with the use of appropriate ICT.

While moving across the three phases, the core modules are visited at greater depth and complexity each time by following a spiral curriculum. For example, in the Attain phase participants first look at the hierarchy of cognitive levels for writing learning objectives. In the Align phase this is revisited when participants align assessment questions at different levels with learning objectives. In the Integrate phase, participants write an entire lesson plan using the same objectives and assessment questions from the previous phases.

Each phase of the A2I model follows a specific pedagogical format of the instructional activities, corresponding to the goals of that phase. In the Attain phase, activities are mostly instructor led, with a few being driven by participants at an individual level. In the Align phase, mastery in application of concepts is targeted. Hence most activities are participant-driven and are to be performed individually. Activities in the Attain and Align phase are 5 to 15 minutes long. In the Integrate phase, participants are primarily engaged in longer collaborative activities to solve real-life teaching-learning problems faced by them in class. Activities are mostly active-learning based, especially in the Align and Integrate phase, but also to a large extent in the Attain phase.

Some elements of the A2I model are derived from the TPACK framework (Mishra & Koehler, 2006). For example, the Attain phase looks at elements of Pedagogical Knowledge (PK) while Align phase stresses on the development of Pedagogical Content Knowledge (PCK) and the Technological Pedagogical Knowledge (TPK). The final Integrate phase addresses the synthesized TPACK of the instructor. (However, there is no explicit goal to increase instructors’ TPACK.)

Implementation of ET4ET training program

The A2I model was first applied and evaluated in a pilot small-scale implementation with 30 participants for a 5-day training program (Warriem et al., 2014). Results showed a conscious shift of the instructors towards including more student-centered learning strategies. Also, participants were successful in aligning their instructional strategies with their stated student learning objectives but faced difficulty in alignment of assessment strategies. This informed future implementations of the A2I model that more opportunity must be provided for practice of alignment of assessment strategies with the other modules.

Going from small to large-scale implementation, the key design features prescribed by the A2I model stayed the same, but the implementation was adapted to account for the scaling in terms of the number of participants, existence of remote classrooms and the duration of the program. The training program design also took advantage of the affordances of the ICT-enabled mode of implementation: the synchronous remote classroom mode provided by the T10KT project infrastructure, combined with the asynchronous interaction via Moodle. While scaling up, the following were the main design considerations:

- The three core modules of learning objective, instructional strategy and assessment strategy remained the same.
- Since the duration of ET4ET was longer than the pilot, more content was added. A larger number of ICT tools addressing different teaching-learning goals, and corresponding instructional strategies were discussed.
• The active learning pedagogy stayed the same, however this got adapted to the ICT-enabled mode of implementation.

The training program was conducted across 7 weeks during June-July 2014. The program began with 3 days of sessions in the Synchronous Remote Classroom (SRC) mode, followed by 5 weeks of asynchronous Moodle-based interactions (considered to be equivalent to 5 days of synchronous sessions) and concluded with 3 days of synchronous sessions where participants reassembled at their remote centres. Each synchronous day contained four 1.5-hour sessions, dealing with one of the core modules or their alignment. Each asynchronous session spanned a week. Figure 1 shows the distribution of content across the training program, categorized in terms of the three core modules and sequenced using the A2I model.

Content and organization of ET4ET

Figure 1 shows the specific topics and their distribution across the training program. The topics are categorized in terms of the three core modules and sequenced according to the three phases of A2I model. Many instructional strategies discussed in the program such as Peer-Instruction are explicitly based on theoretical and empirical research, and have shown objective evidence of improvement of learning under prescribed conditions (Crouch & Mazur, 2001).

![Figure 1. Topics and sequence of the ET4ET program (LO: Learning objective, IS: instructional strategy, AS: assessment strategy)](image)

We explain how the trainers utilized the A2I model using two examples of integrating technology: (i) a well-established technology of computer-based visualizations for conceptual and procedural understanding, and (ii) a more recent technology – wiki, for collaboration.
Visualizations such as animations and simulations are recommended for scientific topics since they facilitate learning for abstract, complex contents that are otherwise difficult to observe, allow learners to practice “what-if scenarios”, and foster their analytical skills (Rutten, van Joolingen, & van der Veen, 2012). While designing activities for integrating visualizations in classroom, participants first learnt about the need and types of visualization (Attain) on SRC-Day1 followed by aligning the different types of visualization corresponding to different learning objectives (Align) on SRC-Day2. The next few sessions targeted the rationale, features, and implementation know-how of active learning strategies such as Think-Pair Share and Peer Instruction (Align). This was followed by integrating visualizations in lesson plans (Integrate) using Lesson Module in Moodle during weeks 1-3.

The second example is that of teaching with wiki. The objective of the module was to get instructors to create a lesson plan for using a wiki in their own course for group projects. The Attain phase of the module focused on helping instructors identify the key pedagogical features of a wiki (like collaborative editing) that makes it a meaningful technology for group projects. Moodle was used to guide participants through two live wikis. In the Align phase, instructors used their knowledge to create a wiki implementation and evaluation plan for achieving specific objectives in their own course. This was done in synchronous mode and two TPS activities were used for participants to perform the planning. In the Integrate Phase, participants created a detailed implementation plan for a group project using wiki.

**Active learning pedagogy in ET4ET**

<table>
<thead>
<tr>
<th>Activity - Think-Pair-Share</th>
<th>Activity - Debate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong> -</td>
<td>Consider two teachers A and B, for a given topic. Teacher A gives a lecture on the topic in class, followed by problem-solving exercises for the students to practice at home. Teacher B asks students to watch a video of the lecture at home before coming to class, and direct problem-solving activities in class.</td>
</tr>
<tr>
<td>Think (−2 minutes):</td>
<td><strong>Instruction to Coordinators:</strong> Make two groups A and B</td>
</tr>
<tr>
<td>‘What will you do in class? What is your individual answer? Do not simply say I will do problem-solving. Be specific about what will happen during the activity.’</td>
<td><strong>Instruction to Participants:</strong> Those in group A should list points for why Teacher A’s strategy is “better” than Teacher B. Those in group B should list points for why Teacher B’s strategy is “better” than Teacher A.</td>
</tr>
<tr>
<td><strong>Pair (≈5 minutes):</strong></td>
<td><strong>Instruction to Coordinators</strong></td>
</tr>
<tr>
<td>Examine your neighbour’s answer. Does it help students to work on higher cognitive levels? Together, make your answers more specific so that your strategy develops the higher cognitive levels of your students.</td>
<td>Send two main points in favour of Teacher A and two points for Teacher B, through A-view chat.</td>
</tr>
<tr>
<td><strong>Share (≈5 minutes):</strong></td>
<td><strong>Activity – Group Problem Solving</strong></td>
</tr>
<tr>
<td>Share your answer with your colleagues.</td>
<td><strong>Goal:</strong> Create activities for Flipped Classroom</td>
</tr>
<tr>
<td><strong>Centre Coordinator – Share most common answer through A-view Chat.</strong></td>
<td><strong>INITIAL INSTRUCTIONS for participants:</strong></td>
</tr>
</tbody>
</table>

The activities in the ET4ET program were implemented using active learning strategies, many of which were similar to the strategies being discussed as part of the Instructional Strategy module. The rationale for this was two-fold:

- Active learning strategies have been tested repeatedly in various settings and have been shown to be effective to improve engagement and learning. Hence one reason for the use of active learning strategies was to get participants to be engaged with the program content, and ensure effective learning.

![Figure 2. Activities implemented in ET4ET based on active learning](image)
In order for participants to apply an instructional strategy in their own courses, it is essential that they experience the strategy themselves. This is especially true of constructivist-oriented active learning strategies, since participants may not be familiar with implementing such strategies. Once participants experience an active learning strategy first-hand in “student role” with the help of scaffolds like worksheets or assessment rubrics, it becomes easier for them to shift to “teacher role” and create activities. Additional scaffolds like activity constructors or templates were provided to design activities for their own course.

In the training program sessions, active learning strategies such as Think-Pair-Share, Peer-Instruction, debate, group problem solving and peer review were implemented, after adapting the strategies to the ICT-mediated Synchronous Remote Classroom mode (Warriem, Murthy & Iyer, 2013). Figure 2 shows examples of such activities conducted in the session on Flipped Classroom.

In the asynchronous mode, various modules available in Moodle were used to ensure that participants were engaged in the content and collaborated with each other. Specific forums were created for participants to share their ideas of applying ET4ET content in their courses. Participants also used forums to elicit help in the use of technology in their setting, thereby creating a community of ET4ET practitioners.

Assessment

The workshop had a total of 16 open-ended assignments submitted through Moodle. In the assignments, participants created material for their courses for active learning techniques, and lesson plans for integrating ICT. Participants were provided with formative assessment rubrics and trained to use them for both self- and peer-assessment. In addition, the program training faculty sampled 10-20% of submissions for each assignment, identified common errors and provided feedback in the following sessions or on Moodle. This provided closure and made it possible to provide meaningful assessment even though there were large number of submissions.

Evaluation of ET4ET

38 engineering colleges from across various geographic regions of India participated in the ET4ET program. A total of 1138 faculty members from these colleges registered for the program, out of which 603 (336 males - 55.7%, 267 females - 44.3%) participants participated in a survey questionnaire provided at the start of the workshop to collect demographic details. Participating instructors had diverse teaching experience: 23% had 10+ years of experience, 41% had 3-10 years of experience and 36% had < 3 years of experience. To evaluate the program, data were collected and analyzed for the following metrics: (i) Participation rate, (ii) perception of participants and (iii) learning outcome of participants.

Participation rate

Participation rates in the ET4ET program were measured via attendance in synchronous sessions and assignment submissions. The total number of registered participants were 1138, out of which 914 (80% of registered participants) attended on Day-1. The attendance on the last day of the program was 740, thus the attrition rate across workshop was 18.5% of the participation on Day-1. On each day, the gender distribution stayed constant: male – 60%, female – 40%. The details of attendance details of each day are given in Table 1 below:

<table>
<thead>
<tr>
<th>Synchronous day</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance total</td>
<td>914</td>
<td>903</td>
<td>878</td>
<td>780</td>
<td>760</td>
<td>740</td>
</tr>
</tbody>
</table>

Data related to assignment submissions (Table 2) were obtained from Moodle. While analyzing the assignments we looked at four levels of submissions – participants who submitted one assignment, 6 assignments (~40% of assignments), 12 (75% of assignments) and all 16 (100% of assignments).
Table 2. Frequencies of assignment submission

<table>
<thead>
<tr>
<th>Number of submissions</th>
<th>1</th>
<th>6</th>
<th>12</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants who submitted</td>
<td>840 (92%)</td>
<td>550 (60%)</td>
<td>311 (34%)</td>
<td>175 (19%)</td>
</tr>
</tbody>
</table>

In the program, participants who had completed 40% or more of the assignments were designated as having “successfully completed” the program, and were given a certificate of passing. These were considered to be participants who were actively engaged through most of the course. Participants who completed 75% or more successfully were considered to have “completed program with distinction.” We found that in ET4ET, 60% of the registered participants successfully completed the program, and 34% achieved distinction.

In many large-scale programs such as MOOCs, it is seen that as the course progresses, the attrition increases. Typical completion rates in MOOCs of comparable size are 10-12% (Jordan, 2014). Academicians have coined this as the funnel of participation (Clow, 2013) and have mentioned that funnelling occurs right from stage of awareness about the program till its completion. We found that in ET4ET, the completion rates were higher. A possible reason for higher participation and completion rates can be the blended mode of delivery and use of active learning strategies compared to the fully online delivery of MOOCs. We believe that the synchronous remote classroom mode in ET4ET facilitated a sense of community among participants, which is important not just for persistence but also for commitment towards cooperation and motivation (Rovai, 2002).

Participants’ perceptions

A questionnaire on participants’ perceptions was administered via Moodle at the end of program. Questions were based on the constructs of participants’ perception of their learning, and their intent to apply the knowledge and skills from ET4ET in their future courses. Each question was related to one topic discussed in the program (such as setting learning objectives, creating wiki-based activities etc.). The Cronbach $\alpha$ reliability coefficient was calculated for questions in each construct: $\alpha = 0.8383$ for the 6 questions on perceptions on learning; $\alpha = 0.7279$ for the 6 questions on intent to apply. The questions contained a 5-point Likert scale from Strongly Disagree to Strongly Agree. In addition, participants were asked to provide open-ended feedback on the program. We show data of 178 participants (83 males – 46.6% and 95 females – 53.4%) who completed the questionnaire and gave consent.

12 questions relevant to these constructs and corresponding data are shown in Tables 3 and 4. (The remaining questions were related to demographics or organization logistics). For the construct of perception of learning, 89% participants either strongly agreed or agreed that they learnt from the program. For the construct of intent to apply, 82% strongly agreed or agreed that they intend to use the knowledge and skills from the program in their own courses.

Table 3. Participants’ perceptions on learning ($N = 178$)

<table>
<thead>
<tr>
<th>Questions on participants’ learning</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setup learning objectives and matching assessment</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>85</td>
<td>82</td>
</tr>
<tr>
<td>Setup a wiki-based activity for my course</td>
<td>1</td>
<td>8</td>
<td>38</td>
<td>108</td>
<td>23</td>
</tr>
<tr>
<td>Set-up a Peer Instruction activity in my class</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>103</td>
<td>62</td>
</tr>
<tr>
<td>Set up a Think-Pair-Share activity in my class</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td>92</td>
<td>73</td>
</tr>
<tr>
<td>Setup a flipped classroom activity in my course</td>
<td>1</td>
<td>6</td>
<td>14</td>
<td>91</td>
<td>66</td>
</tr>
<tr>
<td>Use Visualizations along with an active learning strategy in my course</td>
<td>0</td>
<td>4</td>
<td>14</td>
<td>101</td>
<td>59</td>
</tr>
<tr>
<td>Average</td>
<td>3</td>
<td>4</td>
<td>14</td>
<td>97</td>
<td>60</td>
</tr>
</tbody>
</table>

In the open-ended responses, participants elaborated their views on their learning and intent to use the knowledge from ET4ET. For example, “I will practice at least two techniques in my next semester’s course.” The following quote is from a participant who found the program beneficial enough that she/he attempted to disseminate the learnings to his/her colleagues who did not attend the program: “Though only 10 faculty from our college initially participated in the program, we brought in others for the latter days. We conducted a 1-day workshop in our college that discussed the ideas that were presented in the [initial days of] the program.” This indicated a shift in the reform
ownership (Coburn, 2003), which is an important dimension of scaling, that goes beyond numbers. Many open-ended responses of participants provided useful feedback related to the design of the workshop. For example: “The online sessions were for too long period of time.” Such feedback will be considered in the next offering of the ET4ET program.

Table 4. Participants’ intention to use knowledge, skills and strategies from ET4ET ($N=178$)

<table>
<thead>
<tr>
<th>Questions on participants’ intention to apply</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify learning objectives and matching assessments</td>
<td>16</td>
<td>1</td>
<td>7</td>
<td>103</td>
<td>51</td>
</tr>
<tr>
<td>Use wiki</td>
<td>3</td>
<td>6</td>
<td>46</td>
<td>96</td>
<td>27</td>
</tr>
<tr>
<td>Use Peer-Instruction activity</td>
<td>1</td>
<td>3</td>
<td>23</td>
<td>83</td>
<td>68</td>
</tr>
<tr>
<td>Use Think-Pair-Share activity</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>91</td>
<td>70</td>
</tr>
<tr>
<td>Use Flipped Classroom</td>
<td>2</td>
<td>5</td>
<td>37</td>
<td>90</td>
<td>44</td>
</tr>
<tr>
<td>Use visualization based activities</td>
<td>1</td>
<td>4</td>
<td>18</td>
<td>102</td>
<td>43</td>
</tr>
<tr>
<td>Average</td>
<td>4</td>
<td>4</td>
<td>23</td>
<td>94</td>
<td>52</td>
</tr>
</tbody>
</table>

**Participants’ learning**

The assessment in the training program required participants to complete assignments that reflected the tangible outputs specified at each phase of the A2I model. These assignments were evaluated using performance rubrics. As an example, we provide the analysis of result from participants’ lesson plan assignment that was given at the end of workshop. A total of 391 participants submitted this assignment, out of which we selected 36 assignments based on the following sampling criteria:

- Participants had to submit more than 12 assignments.
- The remote center (of the participant) should have at least 2 submissions.
- Random sampling of 2 participants per remote center.

Table 5 shows the criteria for evaluation and number of participants attaining each performance level.

Table 5. Participants’ lesson plan rubric and number of participants in each performance level (Total $N = 36$)

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Target</th>
<th>Satisfactory</th>
<th>Inadequate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student-centeredness of instructional strategy</td>
<td>All strategies mentioned require active student participation, beyond mere listening or copying of notes or answering questions. (19 participants)</td>
<td>Majority of strategies require active student participation; however there are a few in which students are passive listeners or there is no clear description of student role. (9 participants)</td>
<td>Majority of strategies do not require active student participation or there is no clarity on the roles of students in these strategies. (8 participants)</td>
</tr>
<tr>
<td>Alignment between learning objectives and instructional strategy</td>
<td>There is a perfect alignment between all the learning objectives and the instructional strategies. (12 participants)</td>
<td>Most instructional strategies are aligned with learning objectives, however there are a few which are not aligned or not clearly explained. (20 participants)</td>
<td>Majority of the strategies are not aligned with the stated learning objectives or there is no clarity on how the strategies are going to be used. (4 participants)</td>
</tr>
<tr>
<td>Alignment between assessment and learning objectives</td>
<td>All assessment questions are aligned with the stated learning objectives. (14 participants)</td>
<td>Majority of assessment questions are aligned with stated learning objectives, however there are a few which are not aligned or unclear. (17 participants)</td>
<td>Majority of assessment questions are not aligned with the learning objectives or there is no clarity on how the assessment is to be implemented. (5 participants)</td>
</tr>
</tbody>
</table>
ICT Integration | ICT is used beyond routine procedural use in either instructional or assessment strategy. For example, use of wiki for Think-Pair-Share activity. (5 participants)
---|---
The teacher uses ICT for a useful but standard task. For example, submission of online assignments in Moodle, demonstration of videos. (20 participants)
The teacher does not use ICT at all. (11 participants)

From Table 5, we infer that:
- Majority of the participants have consciously used student-centered strategies in their final lesson plan.
- Most participants are able to align learning objectives with instructional strategies and assessment.
- Participants are primarily using ICT resources for conventional or procedural tasks like demonstration, file submission etc.

Reflections and recommendations

We found that the Attain-Align-Integrate approach of the A2I model is suitable for large-scale implementation of ET4ET. In this section we give some recommendations for the benefit of those who would like to undertake similar efforts.

- While A2I can be used to create face-to-face programs, we found that a blended approach worked better in large scale. The blended mode, especially interspersing asynchronous online sessions with the synchronous mode, made sure that participants had enough time to practice their learning and reflect on it, as well as had sufficient face-to-face interaction with their peers to keep up motivation and build a community.

- It is crucial that teaching-learning strategies and ICT integration techniques be illustrated via examples in one or more domains. Even though ET4ET is not about learning domain knowledge, it is necessary that participants be able to relate to the examples used for illustrating teaching-learning strategies. If participants are not familiar with a topic used in the example for a strategy, they find it harder to think about the teaching-learning aspects, and may stop being engaged. This is especially true when the topics are at the college-level and participants are remote. Hence, it is important that the examples in sessions and worksheet questions should be from the participants’ domain. If participants are from diverse domains, it is difficult to find examples from multiple domains. So it is desirable to conduct ET4ET for a single or related domains.

- For each instructional strategy being introduced, it is necessary to first implement the strategy as an activity that participants perform, before discussing the detailed explanation of the strategy. For example, before discussing Peer-Instruction as an instructional strategy, the participants are involved in Peer-Instruction activities in previous sessions. This provides them a first-hand experience and time to reflect on the activity.

- Participants not only have to learn new instructional strategies but also have to come up with plans to implement these strategies in their own class. So, it is important for them to be in “student” role before they move to “teacher” role. For each activity, it is useful to explicitly indicate to participants whether they are to be in “student” role or in “teacher” role. Not indicating the role explicitly causes mismatch of expectations.

- For each technology being introduced, it is necessary to equip participants not only with the skills to use the technology but also with the pedagogical affordances of the technology. For example, participants first learn about wiki from a student perspective by doing an assignment, followed by skills training on use and setup of wiki, as well as pedagogical affordances of wiki. This culminates in participants moving to “teacher” role and designing wiki assignments for their own students.

- The use of active learning strategies in the training program is a must. In order to adopt a strategy from ET4ET into their own courses, it is not sufficient for participants to listen about the strategy or see it being implemented. They need to do hands-on activities required of the strategy in “student” role, only then create instruction based
on that strategy in a “teacher” role. Moreover, such hands-on activities cannot be relegated to later lab sessions but need to be incorporated in a timely manner during discussion of a strategy.

- Sessions in the program which had a mix of individual and collaborative activities worked better than those that had only one or the other. For any activity being carried out by participants, it is useful to have a participant driven collaborative activity following an individual activity. This ensures that group work occurs and individual participants learn more.

- It is important to go beyond automated multiple-choice questions especially for “applied” topics such as ICT integration in teaching. Scaffolds such as “activity constructors” and templates were provided in open-ended questions. Performance rubrics are one method to assess responses to such open-ended questions, especially with large numbers. Peer- and self-assessment using such rubrics, ensures formative assessment for participants and are viable even in large-scale programs. In ET4ET, what was missing due to the scale, was individual expert feedback on participants’ work. But a well-designed rubric combined with structured peer-review and closure (such as a session reviewing common mistakes) compensated for it to a large extent.

**Conclusion**

In this paper, we have described the design, implementation and evaluation of ET4ET, a large-scale blended-mode professional development program on ICT integration for engineering college faculty. The design of ET4ET originates from the Attain-Align-Integrate (A2I) model, whose basis is constructive alignment. The A2I model was first implemented in a small-scale face-to-face training program for similar goals. The model informs the choice of content, its organization and pedagogy of the training program. At the same time, it is a flexible model, and can be used by others intending to create training programs for a similar goal, since the breadth and depth of individual topics in the training can be decided based on the implementation conditions.

Evaluation of ET4ET showed that participants had high perceptions of their learning, and more importantly, their intent to use the knowledge and skills of ET4ET in their own classes. Preliminary analysis of learning outcomes, measured via instructors’ lesson plans show favourable results. Future research involves the study of the impact of ET4ET on instructors’ actual practice. To study the extent of the scaling (Coburn, 2003), we plan to conduct in-depth case studies of some participants in their field settings via observations of teaching practices and interviews of stakeholders.

A key feature of the program that we feel was responsible for its success was the emphasis on practice and reflection, which have been recommended to scaffold teachers’ learning in systemic plans for ICT integration (Hsu & Sharma, 2006). In ET4ET, practice was best achieved through hands-on active learning strategies in which instructors work on activities relevant to their field setting. Reflection was provided via formative feedback on participants’ performance, which is important to close the teaching-learning loop. In this paper we have provided other such guidelines and lessons learnt, after reflecting on our experiences. We hope that our recommendations will benefit other educational technology training program designers.

**Acknowledgements**

We thank the T10KT Project and NMEICT, Government of India for their support.

**References**


The Factors and Impacts of Large-Scale Digital Content Accreditations

Tony C. T. Kuo¹, Hong-Ren Chen²*, Wu-Yuin Hwang³ and Nian-Shing Chen⁴

¹Management & Information, National Open University, Taiwan // ²Department of Digital Content and Technology, National Taichung University of Education, Taichung, Taiwan // ³Graduate School of Network Learning Technology, National Central University, Taoyuan, Taiwan // ⁴Department of Information Management, National Sun Yat-Sen University, Taiwan // tony@mail.nou.edu.tw // hrchen@mail.ntcu.edu.tw // wyhwang@cc.nctu.edu.tw // nschen@mis.nsysu.edu.tw

*Corresponding author

ABSTRACT
E-learning is an important and widespread contemporary trend in education. Because its success depends on the quality of digital materials, the mechanism by which such materials are accredited has received considerable attention and has influenced the design and implementation of digital courseware. For this reason, this study examined the mechanism by which digital materials are accredited in Taiwan by analyzing the effects of this process on the quality of the digital materials accredited between 2009 and 2014 and by exploring its influence on the design and production of these materials. This analysis was used to identify the advantages and disadvantages of e-learning courseware so as to enhance the quality of digital materials. Data were drawn from 240 applications submitted to the Taiwan E-Learning Quality Center between 2009 and 2014, which were evaluated using quality specifications designed by the Taiwan Ministry of Education. Both quantitative and qualitative methods were used to explore the key factors contributing to the success or failure of attempts to achieve accreditation. The results showed that the grading methods, design of learning activities, quantity and quality of teaching materials, use of different media, motivational techniques, and opportunities for students practice were the most important contributors to accreditation. Moreover, we found significant differences between applications that passed and failed the accreditation assessment along four dimensions: “teaching content and architecture,” “design of teaching materials,” “the use of computer-aided design,” and “media and interface design.” Further quantitative analysis revealed that the design of teaching materials was the key contributor to the overall quality of teaching materials. Qualitative analysis showed that designing exercises, examples, and teaching materials according to learners’ abilities was associated with successful applications for accreditation. We also found that designs that did not engage learners and low-quality media were associated with failure to obtain accreditation.

Keywords
E-learning, Teaching materials accreditation, Accreditation assessment, Large-scale analysis, Factor exploration

Introduction
The rapid advancement of information technology has led to the widespread adoption of educational technology known as e-learning. An increasing number of courses now incorporate digitalized learning materials and an e-learning approach (Jung, Wong, Li, Baigaltugs, & Belawati, 2011; Simelane, 2009). However, there is a difference between online learning and traditional physical classroom learning. For example, e-learning classes lack face-to-face interaction, and the learning materials they use are often designed for self-study. These differences may lead to undesirable learning outcomes (Stewart, Goodson, Miertschin, Norwood, & Ezell, 2013). Thus, it is critical to develop standards for the quality of digital learning materials to enhance learning and teaching. In terms of the development of digital materials, a framework incorporating a sequence of steps including analysis, design, development, implementation, and evaluation (ADDIE) can be used to systematically create high-quality digital material. This model can also be used to certify digital learning materials and select outstanding examples.

The mechanism by which digital content gains accreditation should include a variety of standards to ensure that instructional designs meet relevant requirements. The process should also be rigorous, yielding objective, fair, and consistent outcomes (Berta, 2013). Such an accreditation process should meet the following goals: (1) Recognize outstanding and high-quality digital content. (2) Provide criteria for the evaluation of instructional designs and principles to guide instructors and curriculum development. These standards will enable instructors who do not know how to improve their online teaching to adjust their teaching methods and create appropriate digital content (Jung, 2011). (3) Systematically format digital content. Because digital content is usually designed in various formats, it is not easy to share. However, the accreditation process can systematize digital content, thus making it available to other schools and even to the community. This practice can also prevent the development of identical and
overlapping materials by different schools. (4) Promote the development of e-learning courses that can be easily implemented in rural areas to overcome problems related to geography and the availability of space (Ehlers, 2012).

The process of accrediting e-learning materials can be implemented by the government or by private organizations. Since 2006, the Taiwanese government has accredited e-learning materials using clear standards for online course production and for digital content. During the past decade, this certification process has been recognized by the majority of universities as well as the community in Taiwan. It also has had positive effects on the quality of e-learning. For example, national projects to promote teaching excellence have adopted the accreditation of e-learning as one of the criteria used to evaluate project outcomes.

Although the availability of e-learning accreditation has elicited numerous applications for accreditation, these data have not yet been analyzed for research purposes. In terms of accredited digital content, there are 13 datasets with an average of two sets from each year between 2006 and 2014; these data were drawn from 308 courses. The aim of this study was to analyze these data in terms of the following research questions: (a) what are the general findings about and difficulties encountered during the development of digital content? (b) does the success of application for accreditation differ significantly among subject areas? (c) are the various dimensions relevant to e-learning, i.e., teaching content and architecture, the design of teaching material, the use of computer-aided design, the media and interface design of accredited digital content, correlated with accreditation success? and (d) what are the key standards used in determinations of accreditation?

Literature review

An accreditation process for digital content is crucial for enhancing both teaching and learning (Agariya, & Singh, 2013; Wang, Deng, Huang, & Ding, 2011). To enhance the quality and competitiveness of its educational system, the Taiwanese government has embraced ideas originating in the domain of business that emphasize performance management, and it has already developed a series of performance standards for use in the certification process (Chiu, Hsu, Sun, Lin, & Sun, 2005; Yang, 2005). Unlike the accreditation process used for typical educational settings, “digital learning quality assurance” emphasizes a well-planned and systematic procedure for evaluation and maintenance whereby the teaching content must meet certain standards (Pond, 2002; Wang & Chen, 2009).

Pawlowski (2007) proposed that quality should be viewed from different perspectives. Indeed, governments and non-governmental organizations in numerous countries have recognized the need to control the quality of digital learning and have initiated national digital learning plans, established agencies to ensure quality control, and developed standards for evaluation and certification (Jung, 2012). These institutes include the Council for Higher Education Accreditation (CHEA), the United State Department of Education (USDE), the Quality Assurance Agency for Higher Education (QAA), Open Universities Australia, the Australasian Council on Open, Distance and e-Learning (ACODE), the Certificate in Adult and Continuing Education (CACE), eQcheck, the Korea e-Learning Industry Association (KELIA), and eLC (Concannon, Flynn, & Campbell, 2005; MacDonald & Thompson, 2005).

In 2006, the Center for Quality Assurance of Taiwan Digital Learning was founded under the national projects of e-learning. The quality accreditation is different from the digital learning accreditation given by MOE in that the quality accreditation is composed of digital learning service quality, three facets, eight specifications, two types of applicants, three categories of accreditation and three levels of accreditation. The purpose of the establishment of the Center for Quality Assurance of Taiwan Digital Learning is to provide e-learning industries with references to the development of digital content, its implementation and management and the application for evaluation. Through the quality control and its specifications, it is expected that these industries can meet the requirements.

American Society for Training & Development (ASTD) founded in 1994, is the biggest training institute for human resource training and development. It points out that the use of conventional training by the industries will decrease while technology will be used as a tool to gradually replace the conventional one. So far with the rise of information technology, training can be conducted through virtual classroom. However, online learning lacks the face-to-face interaction, forcing learners to take a rather self-study approach. Thus, the design of the digital content is particularly important in whether the content can be clearly presented and whether learners can be facilitated to comprehend and acquire knowledge.
Accordingly, ASTD have derived accreditation standards according to suggestions made by instructional designers and information system specialists from Northern America (ASTD, 2001). These standards are the criteria for the development of digital content (Sanders, 2001) and the evaluation of ECC (E-Learning Courseware Certification) and ADL (Advanced Distributed Learning) (Liu, 2000; Xiong & Wu, 2001). The goal of these standards is to evaluate the applications, the design of online courseware, and multimedia and asynchronous digital content.

The E-Learning Courseware Certification has four evaluating dimensions: interface, compatibility, production quality and instructional design, including 19 standards (ASTD, 2002, 2006) with each standard having 2 to 8 specifications. The first 3 dimensions, namely interface, compatibility, production quality, have 11 standards, of which 3 mandatory standards (navigation function, installation and activation, and readability of words and graphics) should be met and the rest 8 standards should be met up to at least 5 standards. As to instructional design, it has 8 standards, of which 2 mandatory standards should be met (illustration and demonstration) and the rest 6 standards should be met up to half. These required standards should be met in order to be certified. In online courses, instructors are unable to facilitate learning personally. Therefore, instructional design should be carefully designed, including learning objectives, learning content, learning tool, course introduction, learning assessment and learning guidance.

Founded in 1969 by the British government, ODLQC (Open and Distance Learning Quality Council) pursues a goal to enhance educational and training quality as well as protect learners’ interests. ODLQC accredits providers of home study, distance learning, online learning, open courses and courses provided by schools (Council, 2006). It adopted a series of standards in 1999 to guarantee quality in online learning (ODLQC, 2012) and then revised them in 2005. The standards are subdivided into 10 sections: course objectives and outcomes, learning content, promotion and recruitment, enrollment process, learning support, open learning centers, learners’ interests and collaborative providers and accreditation (Adamson, Becerro, Cullen, González-Vega, Sobrino, & Ryan, 2010; Ehlers, 2012).

ODLQC standards were set out in 1998 and then went through several revisions. New standards were put into practice in April 2006. They are subdivided into six components: outcomes, resources, support, selling, providers and collaborative provision. Every component has four to ten assessment specifications based on which providers are assessed, ranging from administrative and tutorial method to educational materials and publicity (Council, 2009). The emphasis of ODLQ standards is on education quality. Once providers are accredited, the monitor will begin to ensure good service and reassessment will take place once every three years.

Yang (2004) points out that ODLQC accreditation not only aims to assess e-learning content and educational activities but also works on monitoring the development of e-learning process. It appears that ODLQC standards include both the aspects of management and education system (Lee & Lee, 2008; Sun, Fingerc, Chen, & Yeh, 2008). Through accreditation, it is expected the quality in e-learning education in UK can be enhanced.

Shyu and Liao (2004) points out that there are issues with the current digital content designed by most of the colleges or universities in Taiwan. These digital contents oftentimes fail to motivate learners. Besides, as e-learning courses pervade the higher education in Taiwan, most of these courses are taught in the traditional approach as a face-to-face class would do (Chen, Chen, & Chen, 2008). This makes learning rather unengaging and ineffective. In response to these issue, five types of e-learning knowledge—remembering, conceptual learning, problem solving, motivation and procedural learning—are derived to guide the design of the digital content. Rong, Lee, and Liao (2006) and Chen, Chen, and Wu (2011) proposed a reciprocal representation instructional strategy model to better design digital content where learning content is integrated into multimedia and learning activities which feature interactive or hands-on exercises. This way, learners can learn at their own paces and feel more engaged and motivated at the same time. Horton (2005) proposed a process of assessment consisting of four phases: analysis, design, implementation and evaluation. Analysis phase emphasizes the achievement of the business goal and performance so as to identify learning objective. Design phase includes design of learning method, resource integration, material development and technology adoption. Implementation phase seeks appropriate e-learning design according to the organization. Evaluation phase assesses learning outcomes to ensure whether individual and organizational objectives are achieved.

In order to guarantee e-learning service quality and accreditation, Taiwanese government initiated “National Science & Technology Program for e-Learning” and set up “eLearning Quality Certification Center, eLQCC,” executed by Ministry of Economic Affairs. The mission of eLQCC has been to promote and provide accreditation of digital
content and e-learning service. It plays a role as information provider, standards designer and promoter, and accreditation provider. In 2005 the latest e-learning quality standards were revised and set out to provide evaluation service for digital content as well as promote such quality certification (Chang, Chen, Yeh, & Fang, 2004; Kuo, Ke, Ho, & Chen, 2013).

Quality Accreditation for Digital Content lays its emphasis on digital content, aiming to facilitate learners with appropriate learning content. From the learner’s perspective, the design of digital content should take into consideration instruction, interface design, content, interaction, strategies and outcomes. From the instructor’s perspective, the main point is to supply a cohesive instructional design in terms of objective, content, activities, guidance and evaluation involved during learning. From the material developer’s perspective, the design should include stable functions, convenient communication channel, friendly operation interface and high media quality. As far as management is concerned, it should provide tracking and managing functions as well as progress report for tracking overall and individual learning process. Based on these view points, the standards for digital content accreditation should subsume four dimensions: learning content, learner support, instructional design and instructional media (Sung, Chang, Lin, Lee, & Chen, 2009).

Another e-learning accreditation for e-learning has been led by Taiwanese Ministry of Education since 2006. This accreditation is open for colleges and universities to apply for assessment of e-learning courses and digital content. Through rigorous process of assessment, the quality of e-learning courses and digital content can be assured. After tested run for three years, the application of e-learning accreditation has received recognition from the schools and community and thus became official operation. Using the large amount of data collected from the e-learning certification service, this study attempts to analyze e-learning materials

The accreditation for e-learning

The establishment of E-learning accreditation indicators in Taiwan uses the research and development approach and the Delphi technique. Based on existing digital content accreditation systems and literature analysis of relevant research papers, a prototype for accreditation indicators was developed and submitted to experts for evaluation. This process was repeated until satisfactory results were obtained. This repetitive procedure provided prototype documents to experts and scholars for review, then their opinions were collected, and the modified documents underwent a new round of review. After retrieval, if there were no consistent results, review and adjustments would be made repeatedly until the experts reached a consensus. In the first step of this study, various domestic and foreign documents, reports, and journal articles on distance open universities were collected. In the second step, the collected documents underwent critical analysis, screening for useful information. The important indicators of the quality of distance open universities were broken down into individual specifications. In the third step, all of the above indicators were reclassified, as based on characteristics, into certain dimensions in order to form a draft of quality indicators with a comprehensive scope. In the fourth step, such information was delivered to at least 20 scholars, experts, and practical personnel involved in online instruction in Taiwan, who made review recommendations for the items in the draft of indicators. In the fifth step, after compiling the opinions of the scholars, experts, and practical personnel, the various indicator descriptions and weights were revised in order to adjust the quality indicators draft. In the sixth step, the modified results were delivered to the experts, scholars, and practical personnel for confirmation; and based on the results, steps four through six were repeated. If consistent results were obtained from the experts and scholars, then the accreditation specifications and standards for e-learning in Taiwan were promulgated.

A process for the accreditation of E-learning was implemented by the Taiwanese Ministry of Education, and its assessment standards include four dimensions: instructional design, teaching content and architecture, facility design, and media and interface design. Each standard is subdivided into several specifications of essential elements to guarantee quality. The grading system for each specification includes three levels: A, A+, and B. The achievement of accreditation requires grades of at least A in all essential specifications. In other words, one grade of B assigned for an essential specification prevents accreditation. In terms of numeric averages, an A+ is equivalent to 3 points, an A
to 2 points, and a B to 1 point. For instance, the average of an A+ and a B is an A. In brief, this accreditation process uses a pass/fail standard. The current accreditation standards can be found in Appendix A.

In terms of the assessment process, once a case is qualified for assessment, it will be assessed by five examining committees, three e-learning specialist and two field experts. The assessment process normally takes about 2 to 3 weeks to complete. Upon the completion of the assessment, a meeting will be held to go over the results again to ensure they reach consensus. Only one grading remark will be released so that the assessment results are objective and rigorous. The standards for E-learning Accreditation have gone through several revisions since 2006 with a revision taking place every two years as shown in Table 1. Since there was a test-run period between 2006 and 2008, the standards adopted before 2008 were significantly different from those adopted afterwards.

<table>
<thead>
<tr>
<th>Years of revision</th>
<th>2006 ~ 2008</th>
<th>2009 ~ 2010-1</th>
<th>2010-2 ~ 2012</th>
<th>2013 ~ 2014-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Description</td>
<td>35 specifications (23 required, 12 optional)</td>
<td>30 specifications (20 required, 10 optional)</td>
<td>29 specifications (19 required, 10 optional)</td>
<td>28 specifications (20 required, 8 optional)</td>
</tr>
</tbody>
</table>

The latest fourth edition includes 28 specification indicators, including 20 required specifications and eight optional specifications. Between the second and third editions, 1 required specification (the system provides clear learning activity guidelines) was deleted; between the third and fourth editions, one required specification (the instructional materials clearly explain the corresponding relationship between the content and the unit under instruction) was added, and 2 optional specifications (the digital instructional materials clearly explain the copyright of cited materials, and the digital instructional materials provide a FAQ section and a list of keywords) were deleted. Please refer to the e-learning accreditation indicators in the appendix A.

**Research method**

The study firstly introduces the accreditation standards for the Center for Quality Assurance of Taiwan E-Learning and then analyzes the significant aspects included in the accreditation standards based on which the quality of digital content are examined. How the examining procedure and committee were formed will be also investigated. Then the revisions of the accreditation standards taking place between 2006 and 2014 will be described. Finally, the classification of subject areas from 2009 to 2014 will be presented to display its impacts on the accreditation itself. This secession is divided into four sections, introduction to accreditation standards for digital content, research questions, research sampling and research methods.

**Research design**

Taiwan has been accrediting digital content for many years, and this process has yielded a substantial amount of relevant data. However, these data have not yet been examined in terms of accreditation rates in the service of providing an overarching understanding the assessment process. Thus, an in-depth investigation is warranted. It is hoped that such an analysis of the data will answer the following four questions: (a) what is the most important contributor to achieving or not achieving accreditation? (b) do different subject areas differ significantly in their success or failure in achieving accreditation? (c) is accreditation success correlated with four standards noted above? and (d) what is the most important contributor to the quality of digital material according to the accreditation committee?

Having these research questions answered, we can come to conclusions and suggestions for applicants’ references which can be used to promote the development of digital content. During the making of digital content, the school-based providers would benefit from the findings on the key factors to good quality and then learn how to achieve it. They can also learn how to better design digital content by taking the varying nature of subject areas into account.
Sampling

A total of 240 cases were collected as data sets from the application pool. The analyzed samples came from the applications for E-learning course accreditation with the Ministry of Education of Taiwan. The applications from university teachers for E-learning course accreditation are accepted in February and July each year. These data sets were collected at two time points each year from 2009 to 2014 as shown in Table 2. Since the period between 2006 and 2008 was the test-run period, the standards (version 1) and data sets were excluded from the analysis. Three versions of standards were adopted to assess these applications at different time with 18 applications assessed by version 2, 132 cases assessed by version 3 and 90 cases assessed by version 4.

Table 2. The application status

<table>
<thead>
<tr>
<th>Timeline</th>
<th>Applications</th>
<th>Version</th>
<th>Application in total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>10</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>2010-1</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-2</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-1</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-2</td>
<td>23</td>
<td>3</td>
<td>132</td>
</tr>
<tr>
<td>2012-1</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012-2</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-1</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013-2</td>
<td>25</td>
<td>4</td>
<td>90</td>
</tr>
<tr>
<td>2014-1</td>
<td>51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>240</td>
</tr>
</tbody>
</table>

In addition, to compare the outcomes of these applications for different subject areas, the subject areas are classified into ten categories as shown in Table 3.

Table 3. Categorizations of subject matters

<table>
<thead>
<tr>
<th>Number</th>
<th>Subject areas</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Computer and information</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>Economic and financial management</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>Language learning</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>Art and Literature</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>Education studies</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>Health and care</td>
<td>22</td>
</tr>
<tr>
<td>7</td>
<td>Social science</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>Science and life</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>Natural science</td>
<td>52</td>
</tr>
<tr>
<td>10</td>
<td>Practical life</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>240</td>
</tr>
</tbody>
</table>

Data analysis

The 240 sets of data were analyzed using categories of subject areas as independent variable, and fail or pass remark and grading remark (A+, A and B) of each standard as dependent variable. These data sets were used to present the distribution of and correlation among these data. Moreover, the suggestions made by the assessment committee were analyzed using qualitative method.

- Basic analysis: Descriptive statistics analysis was show the percentage of the fail or pass counts and grading remarks (A+, A and B) of each standard. This analysis can sufficiently identify the key factors to the failure of passing the accreditation. In addition, to further compare whether there is any difference in passing rates for different the subject areas, the data were first ranked by subject areas and then summed up the number of fail or pass counts and grading remarks.
Qualitative analysis: Pearson correlation coefficient analysis was used to examine whether there is any relationship among the four accreditation standards for cases that failed and passed the accreditation assessment.

**Results**

**Basic analysis**

To find out the general drawbacks in the making of digital content and concrete strategies and methods for improving e-learning quality, failure rates are analyzed. First, the overall failure rates are higher than 25 percent as shown in Table 4. As to the failure rates ranking by Assessment Standards, it appears that 36 percent of the cases failed in the standard of “Appropriate Provision of Assessments,” 37 percent of the cases failed in the standard of “Appropriate Provision of Feedback,” 36 percent of the cases failed in the standard of “E-Learning Materials in Appropriate Quantity,” 31 percent of the cases failed in the standard of “Good Media Design,” and 29 percent of the cases failed in the standard of “Engaging Design” and “Appropriate Exercises” respectively.

<table>
<thead>
<tr>
<th>Assessment specification label</th>
<th>Assessment standards</th>
<th>Failure rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-5</td>
<td>Appropriate Provision of Assessments</td>
<td>38%</td>
</tr>
<tr>
<td>2-6</td>
<td>Appropriate Provision of Feedback</td>
<td>37%</td>
</tr>
<tr>
<td>1-7</td>
<td>E-Learning Materials in Appropriate Quantity</td>
<td>36%</td>
</tr>
<tr>
<td>4-1</td>
<td>Good Media Design</td>
<td>31%</td>
</tr>
<tr>
<td>2-1</td>
<td>Engaging Design</td>
<td>29%</td>
</tr>
<tr>
<td>2-4</td>
<td>Appropriate Exercises</td>
<td>29%</td>
</tr>
</tbody>
</table>

From the findings shown above, the major factors that potentially led to failure of the accreditation can are: (1) Standard 2 is very important in that the learning unit failed to provide necessary evaluation and feedback. (2) Standard 1 is important for the digital content were not properly prepared in quantity. (3) Standard 4 is important for the quality of the media design is unsatisfactory. It appears that the failure of the accreditation assessment can be attributable to the lack of instructional design. Besides, the lack of necessary technology literacy for making e-learning materials might be another issue. All the digital materials used in these cases can be broadly categorized in three groups: (1) Classroom video-taping with basic editing (2) Video-based PowerPoint presentation where voice recording (3) Other diverse digital materials.

<table>
<thead>
<tr>
<th>Subject areas</th>
<th>Applications</th>
<th>Number of applications passed</th>
<th>Passing rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer and information</td>
<td>60</td>
<td>31</td>
<td>49.18%</td>
</tr>
<tr>
<td>Economic and financial management</td>
<td>27</td>
<td>8</td>
<td>29.63%</td>
</tr>
<tr>
<td>Language learning</td>
<td>9</td>
<td>3</td>
<td>33.33%</td>
</tr>
<tr>
<td>Art and Literature</td>
<td>16</td>
<td>8</td>
<td>50.00%</td>
</tr>
<tr>
<td>Education studies</td>
<td>8</td>
<td>3</td>
<td>37.50%</td>
</tr>
<tr>
<td>Health and care</td>
<td>22</td>
<td>10</td>
<td>45.45%</td>
</tr>
<tr>
<td>Social science</td>
<td>24</td>
<td>13</td>
<td>54.17%</td>
</tr>
<tr>
<td>Science and life</td>
<td>11</td>
<td>6</td>
<td>54.54%</td>
</tr>
<tr>
<td>Natural science</td>
<td>52</td>
<td>25</td>
<td>48.08%</td>
</tr>
<tr>
<td>Practical life</td>
<td>11</td>
<td>5</td>
<td>45.45%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>240</strong></td>
<td><strong>112</strong></td>
<td><strong>46.67%</strong></td>
</tr>
</tbody>
</table>

To understand whether there is any difference among these subject areas, the pass rates were analyzed by subject areas. Of the 240 cases, 112 passed and the overall pass rate is 46.67 percent. The passing rates of each subject matter are, in order, 54.54% for Science and Life, 54.17% for Social Science, 50% for Art and Literature, 49.18% for Computer and Information, and 48.08 for Natural Science, 29.63% for Economic and Financial Management, 33.33% for Language Learning and 37.50 for Education Studies as shown in Table 5. Note that the passing rates of the first five subject areas were above the average passing rates.
In Table 6, the symbol m(n) represents the number of failed the accreditation in the E-learning accreditation indicators. For instance, 2-5(11) shows that the number of failed the accreditation is eleven for the 2-5 of E-learning accreditation indicator.

Table 6. Comparisons between subject matters in standard grading

<table>
<thead>
<tr>
<th>Subject Area</th>
<th>Required</th>
<th>Required</th>
<th>Required</th>
<th>Optional</th>
<th>Optional</th>
<th>Optional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer and information (29)</td>
<td>2-5 (13)</td>
<td>1-7 (11)</td>
<td>2-6 (10)</td>
<td>2-7 (20)</td>
<td>2-9 (12)</td>
<td>3-3 (12)</td>
</tr>
<tr>
<td>Economic and financial management (19)</td>
<td>2-5 (6)</td>
<td>2-6 (6)</td>
<td>2-4 (5)</td>
<td>2-7 (11)</td>
<td>1-10 (7)</td>
<td>2-9 (6)</td>
</tr>
<tr>
<td>Language learning (6)</td>
<td>1-5 (3)</td>
<td>1-7 (3)</td>
<td>2-4 (2)</td>
<td>2-7 (4)</td>
<td>3-3 (3)</td>
<td>1-9 (2)</td>
</tr>
<tr>
<td>Art and Literature (8)</td>
<td>1-6 (5)</td>
<td>1-5 (4)</td>
<td>1-7 (4)</td>
<td>2-7 (7)</td>
<td>2-9 (3)</td>
<td>3-3 (2)</td>
</tr>
<tr>
<td>Education studies (5)</td>
<td>2-5 (3)</td>
<td>2-4 (2)</td>
<td>2-6 (2)</td>
<td>2-7 (3)</td>
<td>3-3 (2)</td>
<td>1-10 (1)</td>
</tr>
<tr>
<td>Health and care (12)</td>
<td>2-1 (4)</td>
<td>2-5 (4)</td>
<td>2-6 (3)</td>
<td>2-7 (7)</td>
<td>3-3 (4)</td>
<td>2-8 (3)</td>
</tr>
<tr>
<td>Social science (11)</td>
<td>2-5 (4)</td>
<td>1-7 (3)</td>
<td>2-1 (3)</td>
<td>2-7 (7)</td>
<td>3-3 (4)</td>
<td>2-9 (2)</td>
</tr>
<tr>
<td>Science and life (5)</td>
<td>1-7 (3)</td>
<td>2-5 (3)</td>
<td>1-2 (2)</td>
<td>2-7 (3)</td>
<td>2-9 (2)</td>
<td>1-10 (2)</td>
</tr>
<tr>
<td>Natural science</td>
<td>4-1 (9)</td>
<td>2-1 (8)</td>
<td>2-6 (6)</td>
<td>2-7 (17)</td>
<td>3-3 (11)</td>
<td>1-10 (10)</td>
</tr>
<tr>
<td>Practical life</td>
<td>2-1 (2)</td>
<td>4-1 (2)</td>
<td>1-3 (1)</td>
<td>2-7 (3)</td>
<td>1-10 (2)</td>
<td>1-8 (1)</td>
</tr>
<tr>
<td></td>
<td>1-7 (1)</td>
<td></td>
<td>2-4 (1)</td>
<td></td>
<td></td>
<td>2-9 (1)</td>
</tr>
<tr>
<td></td>
<td>2-5 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-6 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Essential specifications by the top three failed subject matters

<table>
<thead>
<tr>
<th>Specifications Label</th>
<th>Essential specification</th>
<th>Total number of failed subject areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-6</td>
<td>Appropriate Provision of Feedback</td>
<td>8</td>
</tr>
<tr>
<td>2-5</td>
<td>Appropriate Provision of Assessments</td>
<td>7</td>
</tr>
<tr>
<td>1-7</td>
<td>Learning Materials in Appropriate Quantity</td>
<td>5</td>
</tr>
<tr>
<td>2-1</td>
<td>Engaging Design</td>
<td>5</td>
</tr>
<tr>
<td>2-4</td>
<td>Appropriate Provision of Exercises</td>
<td>5</td>
</tr>
</tbody>
</table>

In terms of essential specification, the top 3 assessment specifications that are difficult to pass were identified as Table 7 shows. “Appropriate Provision of Feedback” (2-6) was found to be top factor that failed in the assessment among these eight subject areas except Language Learning and Life and Science. “Appropriate Provision of Assessments” (2-5) was found in 7 subject areas, ranking second. This finding suggests that feedback and assessments are the mostly commonly neglected aspects of the digital content design. “Engaging Design” (2-1) was found to in Health and Care, Social Science, Life Science, Natural Science and Practical Life. As far as subject area is concerned, it was found that for Computer and Information, specifications 2-5, 1-7, and 2-6 are the areas that should be enhanced while for Economic and Financial Management, specifications 2-5, 2-4 and 2-6 are difficult to pass.

In terms of selective specifications, the ranking of the selective specifications by the top three failed subject areas is shown in Table 8. “Appropriate Remedial Learning” (2-7) was found to be the key factor that resulted in B in these 10 subject areas, implying this area should receive great attention. Following that, “Appropriate Provision of FAQ
and Index” (2-9) and “Appropriate Provision of Learning-relevant Materials” (1-10) were found not easy to pass. “Report on the Learned Content” and “Appropriate Provision of Keyword Search and Definition” ranked No. 4 and No. 5. Subject areas wise, 2-7, 2-9 and 3-3 were not easy to pass for Computer and Information, 2-7, 2-9, 1-10 were difficult for Economic and Financial Management, and 2-7, 3-3 and 1-10 were not easy for Educational Studies and Natural Science.

**Table 8. Selective specifications by the top three failed subject areas**

<table>
<thead>
<tr>
<th>Specification Label</th>
<th>Essential specification</th>
<th>Total number of failed subject areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-7</td>
<td>Appropriate Provision of Remedial Learning</td>
<td>10</td>
</tr>
<tr>
<td>2-9</td>
<td>Appropriate Provision of FAQ and Index</td>
<td>6</td>
</tr>
<tr>
<td>1-10</td>
<td>Appropriate Provision of Learning-relevant</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Materials</td>
<td></td>
</tr>
<tr>
<td>3-3→3-2</td>
<td>Report on the Learned Content</td>
<td>4</td>
</tr>
<tr>
<td>3-3 (newly added)</td>
<td>Appropriate Provision of Keyword Search and</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Definition</td>
<td></td>
</tr>
</tbody>
</table>

Generally, Computer and Information and Natural Science have the highest applications rates 46.7 percent, coming to a total of 112 cases out of 240. The pass rate of Computer and Information is 49.18 percent and that for Natural Science is 48.08 percent. For the cases of Computer and Information that failed in the assessment, three assessment specifications, including the essential and selective ones, were identified and found to be the potential factor that leads to such failure: (1) Appropriate Provision of Feedback (2) Appropriate Provision of Assessments (3) Learning Materials in Appropriate Quantity. Although being different in ranking, they were the common ones.

**Quantitative analysis**

This section first addresses whether applications that did and did not achieve accreditation differed significantly in terms of their scores for the four accreditation standards. The results of independent-sample t-tests revealed significant differences between successful and unsuccessful applications for all four standards: teaching materials and architecture \( t = -14.628, p = .007 < .05 \); design of teaching materials \( t = -17.939, p = .000 < .05 \); use of computer-aided design \( t = -8.578, p = .000 < .05 \); and media and interface design \( t = -10.226, p = .000 < .05 \).

Next, we explored the correlation between accreditation success and each standard, i.e., teaching content and architecture, design of teaching materials, use of computer-aided design, and media and interface design, to identify the factor with the greatest influence on the quality of learning materials. The evaluation of teaching content and architecture focused on teaching goals, content, and learning level; the design of teaching materials was evaluated in terms of opportunities for practice (e.g., remedial learning and assessment activities); that evaluation of use of computer-aided design attended to teaching manuals, guidelines, the search function, and the feedback mechanism; and the evaluation of media and interface design focused on the quality of the media employed by the teaching materials and the user-friendliness of the interface design. The Pearson correlation analysis (Table 9) revealed that only the design of teaching materials was significantly correlated with accreditation success. Thus, achieving or not achieving accreditation was most strongly associated with the design of teaching materials.

**Table 9. Pearson correlation analysis of failed certification courses**

<table>
<thead>
<tr>
<th>Certification demission</th>
<th>Teaching contents and architecture</th>
<th>Design of teaching materials</th>
<th>Use of computer-aided design</th>
<th>Media and interface design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teaching contents and</td>
<td>1</td>
<td>.245*</td>
<td>.199</td>
<td>.114</td>
</tr>
<tr>
<td>architecture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching design</td>
<td>.245*</td>
<td>1</td>
<td>.342**</td>
<td>.250*</td>
</tr>
<tr>
<td>Computer-aided design</td>
<td>.199</td>
<td>.342**</td>
<td>1</td>
<td>.277*</td>
</tr>
<tr>
<td>Media and interface</td>
<td>.114</td>
<td>.250**</td>
<td>.277*</td>
<td>1</td>
</tr>
<tr>
<td>design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. \( p < .05 \), **\( p < .001 \).*

To further investigate the correlation results, some implications can be obtained; first, the content of learning materials, examples, practice, assessment activities may have different effects on learners due to their different abilities. It is critical to provide suitable content to help learners understand. Secondly, the operation and user
interface of the system for the learners to use examples, practice, assessment activities may be affected by the design of different media and digital content for learners. Therefore, it is important to consider providing a friendly operation interface for learners using teaching examples, exercises or assessment activities, or suitable digital content and media matching learners’ ability when designing learning materials; otherwise it may lead to the failing of the accreditation. These results are consistent with the views of the following scholars. Kassim, Nicholas and Ng (2014) believed that high-quality multimedia animation in learning can help students’ comprehension, which can benefit their creative expression. Chen and Huang (2012) believed that interface design in recommending learning service systems is very important, and is sufficient to influence student satisfaction with the digital instructional materials and their learning achievements.

Qualitative analysis

**Factor analysis of successful certification of e-learning materials**

This section presents an analysis of the standards by focusing on applications that achieved accreditation between 2009 and 2014. Figure 1 shows that three qualities were associated with successful accreditation each year. The first was that the instructional materials clearly specified themes, unit names, study hours, and eligible students; the second was that the instructional materials provided detailed navigation instructions; and the third standard was that the instructional materials provided a clear explanation of learning activities. In other words, reviewers treated these three factors as the most important contributors to the quality of digital content. Thus, to achieve accreditation, the digital content must clearly specify the content themes, unit names, study hours, and eligible students. In terms of detailed navigation and explanations of the instructional materials, the digital content submitted for review must clearly explain layout, buttons, and site maps. With regard to learning activities, students must be provided with explanations about how to implement the learning activities as well as with design suggestions for and feedback in response to their participation in the learning activities. We also identified three factors that contributed to the successful certification of instructional materials. In 2009 and 2013, the instructional materials that were accredited included descriptions of the appropriate practices for their implementation. In 2010 and 2012, the successful instructional materials provided suitable practical examples or models. In 2012 and 2014, the successful instructional materials were suitable for the learning level of the students.

**Figure 1. Analysis of indicators for successful accreditation of digital content between 2009 and 2014**

**Factor analysis of failed certification of e-learning materials**

This section is an analysis of standards for the applications that failed the accreditation of digital content between 2009-2014 years. Figure 2 shows that there are three standards that were the same for the failed applications. These
are (1) suitability of the content of instructional materials, (2) design to elicit learning motivation, and (3) quality of media in the instructional materials.

For the suitability of the content of instructional materials, reviewers believed that the instructional time of instructional materials was too short, the average class time per week was insufficient for credit requirements, or the instructional materials did not provide weekly class hour information, the demonstration files were too short in terms of time, and the distribution to units in the instructional materials was uneven.

In terms of the instructional materials eliciting learning motivation, reviewers believed that the digital instructional materials lacked diverse designs to inspire learning motivation. In terms of quality of instructional materials, there are three main opinions, which were that the slides for instruction lacked audio explanations or videos to accommodate design of the instructional media, improvements were needed for the video quality, some parts needed text explanations, and that the sounds explaining some units were not clear, and that some online instructional units had no sound or were blurry.

There are also 3 different factors for the successful review of digital instructional materials each year. In 2009 year, the reviewed digital content did not provide clear explanations on learning activities; specifically, it failed to provide students with explanations on how to implement the learning activities in the instructional materials.

In 2010 and 2012 years, the media in the instructional materials submitted failed to help learners understand the content. For instance, most of the video content was reading the slides aloud, and did not provide additional explanations, which made it hard for students to understand. The multimedia presentation of digital instructional materials did not help learners understand the content, thus the media design needed improvement. The content of the digital content did not cover standards of instructional objectives in the digital content that failed review in 2011 year, including that the digital content failed to provide detailed learning objectives for chapters or that there were no correlations among the content, knowledge, techniques, and attitudes of the instructional materials.

In 2013, units in instructional materials failed to provide suitable evaluation activities; for instance, after learning units there were no tests, so students could not know the learning outcomes, or that the quality of evaluation activities was poor. For example, all answers were correct or there were insufficient evaluation activities.

In 2014 year, instructional materials did not provide suitable feedback for learning activities. For example, after-class evaluation feedback did not show the process of problem-solving, which made it difficult to achieve practice effects, and the instructional materials used discussion forums to interact with students but did not give students significant feedback content.

![Factor analysis of successful certification of E-learning materials](image)

**Figure 2.** Analysis of standards for failed certification of digital content between 2009 and 2014

Implications and suggestions

Our analysis revealed that the main cause of failure to receive accreditation was the failure to address fundamental principles of instructional design such as adaptive feedbacks to lead students to learn why they are wrong and what
to study more, appropriate evaluation to help students what they know and what they have not known and learning contents with high quality and related to familiar content. Regarding “the analysis of media types showed that the coordination between instructional design and digital content was an essential contributor to the quality of the content,” I append the followings, the further exploration shows that multimedia content in instruction design had better consider the authentic examples or stories surrounding instead of unfamiliar context and some questions will give to students to follow up. After that, students will be lead to explore the questions with multimedia content surrounding them and later on collaborative instruction design like peer assessment or team work can proceed effectively due to students’ familiar questions and contents (Jung & Latchem, 2012; Kassim, Nicholas, & Ng, 2014).

As for “institutions devote attention to providing adequate professional training to enhance teachers’ ability in the area of the design of instructional materials,” which is a rather well-known fact in e-learning. Besides institutions devote attention to providing adequate professional training to enhance teachers’ ability in the area of the design of instructional materials, it was found that the successful cases usually have been applied to real classes a couple of times and get useful feedbacks from students and teachers as well for further improvement, therefore their designs in instructions and contents become flexible and suitable for students’ needs. Moreover, after examining the evidences provided by the documents of successful cases for evaluation criteria, it also shows that their teachers have done good activity design to motivate discussions using forums among students and provided self-evaluation and tests with meaningful feedbacks and referring the related content as well for further study if students answer incorrectly.

Discussion and conclusion

Our analysis revealed that the main cause of failure to receive accreditation was the failure to address fundamental principles of instructional design. Furthermore, the analysis of media types showed that the coordination between instructional design and digital content was an essential contributor to the quality of the content. It is recommended that institutions devote attention to providing adequate professional training to enhance teachers’ ability in the area of the design of instructional materials. The following common weaknesses were identified in our analysis of subject areas:

- **Appropriate provision of adequate feedback.** Many examples of digital content were directly edited or reproduced from lectures that had been videotaped in physical classrooms. Most involved a unilateral mode of delivery, making it difficult to add interactive learning activities or adequate feedback while a video recording was playing.

- **Appropriate assessment.** Digital content is used primarily for online self-learning; thus, it is important to provide assessment mechanisms for tracking learning outcomes during the learning process. Most digital content failed to meet this standard, as the instructors did not know how to embed exercises and quizzes into a learning unit. One possible solution to this problem would be to use digital-content authoring tools with features that support assessment.

- **Provision of appropriate digital content in quantity.** Most applicants were not clear about much digital content should be created for a complete learning unit. For example, it is easy to calculate the elapsed time of a video-based learning content; however, it is not easy for other types of digital content like HTML/PDF format, PPT slides or interactive animations. Normally, accreditation review of E-learning courses does not strictly review instructional material hours, but for instructional materials that are clearly insufficient; for example, materials with 3 credits, and the applicant provides only 10-20 minutes instructional materials per week, would be given a B grade by experts and scholars, who request additional digital instructional materials. Another critical evaluation concern is how to design a balanced amount of digital content for the whole semester.

In the area of natural science subjects, many applicants failed to pass standard 4-1 due to the poor audio quality of recordings or inadequate use of multimedia materials. For the teachers who were trained in a general education discipline, their technology literacy for making digital content might be weaker than those trained in other disciplines. If their institutions do not provide good hardware equipment and authoring tools, teachers would have encounter obstacles and have difficulty coping with this standard. Another major issue was the lack of engaging design. To have good motivational designs, vivid learning activity design or varieties in media presentations would be required. This again has to do with whether a good institutional support could be provided or not (Zygouris-Coe, Swan, & Ireland, 2009). In the areas of language learning and literature and art, the main challenge is the
arrangement of the created digital learning materials. Furthermore, the structure and sequence among learning units for literature and art related area poses another challenge, that is the rationale of organizational structure should be justified and the sequence of how learning units should be clearly presented.

In terms of quantitative analysis, the study explores whether there is any difference in the four dimensions of certification between the applications that passed or failed the accreditation assessment. Digital content accreditation includes four aspects: teaching contents and architecture, instructional design, facility design, and media and interface design. It is learned from Pearson correlation analysis that these four dimensions influence one another; design of teaching materials, use of computer-aided design, and media and interface design influence one another. Teaching contents and architecture and design of teaching materials are correlated with each other. The design of teaching materials is correlated with other three aspects. It can be inferred that the success or failure of digital content accreditation is possibly highly related with the design of teaching materials. The results are consistent with the findings in the descriptive statistics analysis that the Standard 1 and Standard 2 are the key standards. These research outcomes are similar to adaptive learning, as stated by Yang, Hwang and Yang (2013), by providing students with customized learning for individual personalities and needs. Since the special digital instructional materials for different students are matched with differentiated learning content, every student can successfully complete their learning (Hwang & Chang, 2010).

Qualitative analysis shows that the most valued three indicators held by the committee are: (1) The instruction materials can clearly present the content subject, unit name, learning hours and intended learners, (2) The instruction materials can provide detailed content guidelines and description, and (3) The standards can provide clear instructions on learning activity. Moreover, there are also three factors that can lead to successful accreditation of digital content: Instructional materials provide appropriate exercises; Instructional materials provide suitable examples or models; Instructional materials are designed according to learners’ abilities.

From the analysis of the failure of accreditation, three standards were found to be the common indicators in which the submitted digital content failed every year as follows: (1) suitability of the content of instructional materials, (2) design to elicit learning motivation, and (3) quality of media in the instructional materials. That is to say the committee pays attention on the key indicators such as Standard 1 and Standard 2, which is consistent with the finding in the quantitative analysis above. Besides, the good quality indicator of instructional media is highly correlated with Standard 1 and Standard 2. What we have to be aware of is that no matter digital content accreditation is approved or not, the distinguishable indicators are mostly related to learners’ understanding and interaction with digital content. These findings confirm consistency with the following academic researches. The entertaining characteristics of game-based learning are used to design the English course content, with practice combined with gaming, to realize the advantages of context simulations, repeated practice, and feedback; this can achieve the elevation of learning benefits (Berns, González-Pardo, & Camacho, 2013; Council, Stansfield, & Hainey, 2011).

The study has analyzed the data of the digital content accreditation executed by Taiwan Ministry of Education and explored the key factors and impacts of E-learning quality evaluation. Future research can literally adopt accredited digital content to gain feedback from the teachers and learners as well as to examine the learning outcomes. Whether the learners would have same viewpoints as the assessment committee is worthy of further research.

Acknowledgements

This study was partially supported by the Ministry of Science and Technology under grant numbers, MOST-103-2511-S-110-002-MY3, MOST-103-2511-S-008-017-MY3, MOST-103-2511-S-142-022-MY2, NSC-101-2511-S-110-003-MY3, NSC-101-2511-S-008-012-MY3 and NSC-101-2511-S-008-013-MY3.

References


### Appendix A. Standard of teaching materials accreditation

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Property</th>
<th>Content for quality assessment</th>
</tr>
</thead>
</table>
| Dimension 1: Teaching contents and architecture (required: 8 optional: 2) | Required | 1-1 The teaching materials clearly describe the learning topics, unit titles, learning hours, and target students.  
A+: All four of the learning topics, unit titles, learning hours and target students are listed on the web pages of the teaching materials with clear description.  
A : Three of the learning topics, unit titles, learning hours and target students are listed on the web pages of the teaching materials with clear description.  
B : Two, one, or none of the learning topics, unit titles, learning hours and target students are listed on the web pages of the teaching materials, or the description is not clear.  

Required | 1-2 The teaching materials clearly describe the learning goals such as the knowledge, skills, and attitudes learners can acquire from the teaching materials.  
A+: All three learning goals of knowledge, skills, and attitudes are listed in the overall learning goals and the separate learning goals of every unit with clear description.  
A : Two of the learning goals of knowledge, skills, and attitudes are listed in the overall learning goals and the separate learning goals of every unit with clear description.  
B : Only one or none of the learning goals of knowledge, skills, and attitudes is listed in the overall learning goals and the separate learning goals of every unit, or the description is not clear.  
The learning goals in this criterion include the overall learning goals and the separate learning goals of every unit.  

Required | 1-3 The teaching materials clearly describe the corresponding relation between the content and the teaching subject.  
A+: The corresponding relation between the content and the teaching subject is listed in more than three fourths of the units in the teaching materials with clear description.  
A : The corresponding relation between the content and the teaching subject is listed in more than two thirds of the units in the teaching materials with clear description.  
B : The corresponding relation between the content and the teaching subject is listed in less than two thirds of the units in the teaching materials, or the description is not clear.  

Required | 1-4 The content of the teaching materials covers the learning goals.  
A+: The content of the teaching materials totally covers the overall and the separate learning goals listed in each unit.  
A : The content of the teaching materials covers 90% of the overall and the separate learning goals listed in each unit.  
A : The content of the teaching materials covers under 90% of the overall and the separate learning goals listed in each unit.  

Required | 1-5 The content of the teaching materials is suitable for the ability of the target students.  
A+: The difficulty and breadth of the content of teaching materials are suitable for the ability of the target students.  
A : The difficulty and breadth of the content of teaching materials are approximately suitable for the ability of the target students.  
B : There is a rather gap between the difficulty and breadth of the content of teaching materials and the ability of the target students.  

|
Required 1-6 The content of teaching materials is correct.
   A+: All the content of teaching materials is correct.
   A : There are minor wrong parts in the teaching materials.
   B : There are serious or too many wrong parts.
   “Wrong parts” in this criterion means the knowledge itself is wrong.
   Typos are not included. Self-examination records should be provided by
   schools applying for this certification.

Required 1-7 The presented sequence of the units is appropriate.
   A+: The presented sequence of the units in teaching materials is appropriate
   for the target students.
   A : The presented sequence of the units in teaching materials is approximately
   appropriate for the target students.
   B : The presented sequence of the units in teaching materials is inappropriate
   for the target students.

Required 1-8 The quantity of the content of teaching materials is
   appropriate.
   A+: The quantity of the content in every unit of the teaching materials is
   appropriate.
   A : The quantity of the content in more than two thirds of the units is
   appropriate.
   B : The quantity of the content in less than two thirds of the units is
   appropriate.
   The quantity of the content of teaching materials in this criterion should
   meet the requirements of credits.

Optional 1-9 The teaching materials include relatively-new contents.
   A+: The teaching materials include materials within two years.
   A : The teaching materials include relatively-new materials.
   B : The teaching materials are old and need to be supplemented.

Optional 1-10 Relevant supplementary teaching materials and learning resources are
   provided in the teaching materials.
   A+: Relevant supplementary teaching materials and learning resources are
   provided in more than two thirds of the units in teaching materials.
   A : Relevant supplementary teaching materials and learning resources are
   provided in more than half of the units in teaching materials.
   B : Relevant supplementary teaching materials and learning resources are
   provided in less than half of the units in teaching materials.

Dimension 2: Design of teaching materials

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Property</th>
<th>Content for quality assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension 2: Design of teaching materials (required: 6 optional: 2)</td>
<td>Required 2-1</td>
<td>The units in teaching materials include the design to motivate learning.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A+: More than two thirds of the units in teaching materials include the design to motivate learning.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A : More than half of the units in teaching materials include the design to motivate learning.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B : Less than half of the units in teaching materials include the design to motivate learning.</td>
</tr>
<tr>
<td></td>
<td>Required 2-2</td>
<td>The units in teaching materials provide clear explanation of learning activities.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A+: Every unit in the teaching materials provides clear explanation of learning activities.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A : More than half of the units in the teaching materials provide clear explanation of learning activities.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B : Less than half of the units in the teaching materials provide clear explanation of learning activities.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The learning activities of teaching materials in this criterion mean the activities that learners are required to do in the teaching materials, such as watching supplementary videos, simulations, self-evaluations, and filling out</td>
</tr>
<tr>
<td>Requirement</td>
<td>Credit</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
<td>-------------</td>
</tr>
</tbody>
</table>
| Required 2-3 | | Appropriate examples or practices are provided in the units in teaching materials.  
A+: Appropriate real-life examples or practices are provided in more than two thirds of the units in teaching materials.  
A: Appropriate real-life examples or practices are provided in more than half of the units in teaching materials.  
B: Appropriate real-life examples or practices are provided in less than half of the units in teaching materials. |
| Required 2-4 | | Appropriate exercises are provided in the units in teaching materials.  
A+: Appropriate exercises are provided in more than two thirds of the units in teaching materials.  
A: Appropriate exercises are provided in more than half of the units in teaching materials.  
B: Appropriate exercises are provided in less than half of the units in teaching materials. |
| Required 2-5 | | Appropriate evaluation activities are provided in the units in teaching materials.  
A+: Appropriate evaluation activities are provided in every unit in the teaching materials.  
A: Appropriate evaluation activities are provided in more than half of the units in teaching materials.  
B: Appropriate evaluation activities are provided in less than half of the units in teaching materials. |
| Required 2-6 | | Appropriate feedbacks on learning activities are provided in the units in teaching materials.  
A+: Appropriate feedbacks on learning activities are provided in more than two thirds of the units in teaching materials.  
A: Appropriate feedbacks on learning activities are provided in more than half of the units in teaching materials.  
B: Appropriate feedbacks on learning activities are provided in less than half of the units in teaching materials.  
The learning activity feedbacks on the units in teaching materials in this criterion mean the interaction between the teaching materials and learners, such as grading after self-evaluation and statistical analyses after filling out questionnaires. |
| Optional 2-7 | | Appropriate remedial learning is provided in the units in teaching materials.  
A+: Appropriate remedial learning is provided in more than half of the rather difficult units in the teaching materials.  
A: Appropriate remedial learning is provided in more than one third of the rather difficult units in the teaching materials.  
B: Appropriate remedial learning is provided in less than one third of the rather difficult units in the teaching materials.  
Remedial learning in this criterion means, when there are difficulties in learning, the learners have at least one different way of learning. |
| Optional 2-8 | | The teaching materials clearly describe the learning stages and the contents and learning suggestions in each learning stage.  
A+: Different learning stages are listed in the teaching materials with clear learning contents and learning suggestions.  
A: Different learning stages are listed in the teaching materials with almost clear learning contents and learning suggestions.  
B: Different learning stages are not listed in the teaching materials. Learning contents or learning suggestions are not provided. |
### Dimension 3: Use of computer-aided design

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Property</th>
<th>Content for quality assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension 3: Use of computer-aided design (required: 1 optional: 3)</td>
<td>Required 3-1</td>
<td>Clear tutorials are provided in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A+: Complete tutorials such as layout arrangements, HTML button explanations, and site maps are provided in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A : Some of the tutorials such as layout arrangements, HTML button explanations, and site maps are provided in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B : Tutorials such as layout arrangements, HTML button explanations, and site maps are not provided in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td>Optional 3-2</td>
<td>The parts that have already been learned can be shown in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A+: The parts that have already been learned can directly and clearly be shown in the teaching materials themselves or through system platforms.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A : The parts that have already been learned can be searched in the teaching materials themselves or through system platforms.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B : The parts that have already been learned can be shown neither in the teaching materials themselves nor through system platforms.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In this criterion, “the parts that have already been learned can be shown” means that the color of the parts that have already been learned can directly be changed in the on-line teaching materials, or the learned parts can be looked up through the search function in the system.</td>
</tr>
<tr>
<td></td>
<td>Optional 3-3</td>
<td>The explanations of key words are provided in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A+: The explanations of key words are provided in more than two thirds of the units in teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A : The explanations of key words are provided in more than half of the units in teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B : The explanations of key words are provided in less than half of the units in teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The explanations of key words in this criterion mean the explanations of key words in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td>Optional 3-4</td>
<td>Feedback mechanisms are provided in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A+: Learners can give feedbacks on the teaching materials through the official built-in feedback mechanisms in the teaching materials themselves or system platforms.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A : Learners can give feedbacks on the teaching materials through unofficial feedback mechanisms aside from the teaching materials themselves or the system platforms.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B : There is no feedback mechanism.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The feedback mechanisms in this criterion mean, for example, built-in teaching material evaluation questionnaires. Unofficial feedback mechanisms mean feedback channels like phone consultations, email boxes, or on-line discussion groups.</td>
</tr>
</tbody>
</table>

### Dimension 4: Media and interface design

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Property</th>
<th>Content for quality assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension 4: Media and interface design (required: 5 optional: 1)</td>
<td>Required 4-1</td>
<td>The quality of teaching material media is excellent.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A+: The quality of the media in teaching materials is excellent.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A : The quality of the media in teaching materials is mostly excellent.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B : The quality of the media in teaching materials needs to be improved.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The media in teaching materials in this criterion include words, audios, images, videos, animations, etc.</td>
</tr>
<tr>
<td></td>
<td>Required 4-2</td>
<td>The media in teaching materials can help learners understand the contents.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A+: More than two thirds of the media in teaching material units can help learners understand the contents.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A : More than half of the media in teaching material units can help learners.</td>
</tr>
</tbody>
</table>
understand the contents.

B : Less than half of the media in teaching material units can help learners understand the contents.

In this criterion, “the media in teaching material units can help learners understand the contents” means the media can match the contents of teaching materials, so they can help learners understand the contents.

<table>
<thead>
<tr>
<th>Required</th>
<th>4-3 The layout design of the teaching materials is appropriate.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A+: The outer appearance, color, function, and position of the layout of the teaching materials are all appropriate.</td>
</tr>
<tr>
<td></td>
<td>A : The outer appearance, color, function, and position of the layout of the teaching materials are mostly appropriate.</td>
</tr>
<tr>
<td></td>
<td>B : The outer appearance, color, function, and position of the layout of the teaching materials need to be improved or are disputable.</td>
</tr>
</tbody>
</table>

The layout design in this rule means the design of the outer appearance, color, function, and position of the layout.

<table>
<thead>
<tr>
<th>Required</th>
<th>4-4 The interface operation of the teaching materials is convenient and consistent.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A+: The convenience of using and consistency in the interface operation design in teaching materials are both appropriate.</td>
</tr>
<tr>
<td></td>
<td>A : The convenience of using and consistency in the interface operation design in teaching materials are mostly appropriate.</td>
</tr>
<tr>
<td></td>
<td>B : The convenience of using and consistency in the interface operation design in teaching materials need to be improved.</td>
</tr>
</tbody>
</table>

The interface operation in this rule means the design of the outer appearance, color, function, and position of the layout.

<table>
<thead>
<tr>
<th>Required</th>
<th>4-5 The reading tools of the teaching materials are appropriate.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A+: The function and operation of the reading tools of the teaching materials are both appropriate.</td>
</tr>
<tr>
<td></td>
<td>A : The function and operation of the reading tools of the teaching materials are mostly appropriate.</td>
</tr>
<tr>
<td></td>
<td>B : The function and operation of the reading tools of the teaching materials need to be improved.</td>
</tr>
</tbody>
</table>

The reading tools in this criterion mean the software tools that can play and show the contents of the teaching material media, such as webpage browsers and customized application software.

<table>
<thead>
<tr>
<th>Optional</th>
<th>4-6 Friendly download functions are provided in the teaching materials.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A+: Friendly download functions are provided in more than two thirds of the units in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td>A : Friendly download functions are provided in more than half of the units in the teaching materials.</td>
</tr>
<tr>
<td></td>
<td>B : Friendly download functions are provided in less than half of the units in the teaching materials.</td>
</tr>
</tbody>
</table>

Friendly download in this criterion means, in order to let users read the teaching materials off-line in a convenient way, the contents are rearranged into an appropriate form for download, and unnecessary materials are deleted in this simplified or mobile form.
Assessing a Collaborative Online Environment for Music Composition

Michele Biasutti
FISPPA Department, Padova University, Italy // michele.biasutti@unipd.it

(Submitted January 17, 2014; Revised December 11, 2014; Accepted December 13, 2014)

ABSTRACT
The current pilot study tested the effectiveness of an e-learning environment built to enable students to compose music collaboratively. The participants interacted online by using synchronous and asynchronous resources to develop a project in which they composed a new music piece in collaboration. After the learning sessions, individual semi-structured interviews with the participants were conducted to analyze the participants' perspectives regarding the e-learning environment’s functionality, the resources of the e-learning platform, and their overall experience with the e-learning process. Qualitative analyses of forum discussions with respect to metacognitive dimensions, and semi-structured interview transcriptions were performed. The findings showed that the participants successfully completed the composition task in the virtual environment, and that they demonstrated the use of metacognitive processes. Moreover, four themes were apparent in the semi-structured interview transcriptions: Teamwork, the platform, face-to-face/online differences, and strengths/weaknesses. Overall, the participants exhibited an awareness of the potential of the online tools, and the task performed. The results are discussed in consideration of metacognitive processes, and the following aspects that rendered virtual activity effective for learning: The learning environment, the platform, the technological resources, the level of challenge, and the nature of the activity. The possible implications of the findings for research on online collaborative composition are also considered.

Keywords
E-learning, Online collaborative learning, Asynchronous and synchronous resources, Online music creativity

Introduction
Several studies in the field of online learning have investigated technical solutions to online learning, and methodological issues related to online learning research, and many ideas for improving learning approaches have been provided (Karvounidis, Chimos, Bersimis & Douligeris, 2014; Ng, 2014; Tseng & Yeh, 2013; Wu & Huang, 2013). In early online learning approaches, an instructional model was used in which online tools were merely considered a means for practice. More recently, however, this instructional model has been updated with more interactive didactic methods based on socio-constructivism, and collaborative activities have been developed. These methods aim to respect the learner’s experience, and to stimulate divergent thinking in participants. Prior research has also highlighted relevant aspects of the online learning process related to, for instance, pedagogy, didactic methods, online environments, tools, organization and creativity, and has offered suggestions for designing high-quality learning environments (Biasutti & EL-Deghaidy, 2012; 2014).

The growth of e-learning products has also affected music education, and several music-related tools, and software programs have been developed (Akoumianakis, 2011; Hadjileontiadou, Nikolaidou, Hadjileontiadis & Balafoutas, 2004; Yu, Lai, Tsai & Chang, 2010). Technological advancements have supported the development of e-learning products by providing technological solutions for activities that were previously impossible, such as interacting online in real time to perform, and compose music. The potential of the Internet thus expanded, and one can find through any web search engine thousands of online tools that are now available for music learning. However, many of these online tools have not been evaluated. Thus, the effectiveness of the online tools, and resources needs to be tested, and the pedagogical and didactic approaches to the online learning activities need to be assessed (Seddon & Biasutti, 2009). Moreover, many of the didactic approaches related to the music learning resources available on the Internet are based on an instructional model that involves simple practice, whereas interactive didactic methods based on collaboration are used less often.

The current study presents a project in which participants worked online to collaboratively compose a new piece of music. The learning environment also allowed the participants to interact synchronously, which was a challenging, and complex task. Special software was used to allow more than two participants to interact in real time, and powerful technological solutions were adopted to minimize the latency of the signal. The learning experience was then tested to assess the effectiveness of the online environment.
Theoretical background

In the present study, principles related to collaborative online learning, and online music learning were combined to form the theoretical background.

Online collaborative learning

There is a growing interest in the factors that contribute to the effectiveness of online collaborative activities (Donnelly & Boniface, 2013; Tseng & Yeh, 2013). Prior research has addressed several issues related to online collaboration, such as the strengths and weaknesses of online collaborative activities, the process of collaborative knowledge construction, and the dimensions of metacognition. A positive relationship and collaboration during online activities are relevant factors in online collaboration, since students’ satisfaction with online activities is connected with their perceived levels of collaborative learning. The strengths of online collaborative learning include the ability to compare ideas, collaborate, learn from peers, share knowledge, and skills to support other participants, analyze and integrate different points of view, plan in a group, manage the workload, and use an effective platform (Biasutti, 2011). Moreover, crucial factors for building trust for teamwork include individual accountability, familiarity with team members, commitment toward quality work, and team cohesion (Tseng & Yeh, 2013). Positive interdependence (i.e., the perception that participants are linked with others) developed during online activities is also important, as well as aspects such as establishing a positive group environment, and creating a sense of community. Other aspects include technology competence, technology utility, and technology resourcing (Donnelly & Boniface, 2013). Conversely, factors that impede online collaboration include insufficient ability in workload management, different levels of engagement, insufficient coordination and organization, and technical issues (Biasutti, 2011).

Group coordination and dynamics have been addressed in several previous studies. For instance, prior research has examined the process of collaborative knowledge construction, which describes learners’ cognitive processes that occur during collaborative learning, with a focus on the exploration of their processes rather than considering the mere products. Collaborative knowledge construction depends on specific social, and cognitive processes, as well as the interaction between these processes. Regarding social interaction, Wu, Hwang, and Kuo (2014) demonstrated that highly interactive students have higher learning achievements than less interactive students, indicating the importance of the social dimension of the learning process, and the importance of interaction among group members for knowledge construction. Furthermore, the discourse among the learners in a group is important during collaborative learning, which relates to the cognitive dimension of learning, and the participants’ knowledge construction during collaborative activities. Regarding other online activities, Anderson and Simpson (2004) argued that while discussions in online forums induce basic processes such as the exchange of information, and the investigation of ideas, more articulated processes such as higher-order cognitive skills are not activated. By contrast, Biasutti (2011) has demonstrated that wiki activities can induce higher-order cognitive skills, such as the evaluation of various elements, and subsequent decision making.

Higher-order cognitive skills involve metacognition and reflection on performed actions, and metacognitive processes have been recognized as a crucial factor for enhancing group coordination, and fostering effective learning. Sharing cognitive experiences is another fundamental aspect of the development of metacognitive skills, allowing participants to control and assess one another’s behaviors, cognitive processes, and feelings (Kwon, Hong & Laffey, 2013). Activities such as evaluating group activities, reflecting on the results, and considering approaches to collaboration enhance group coordination and performance. In this way, a group can set realistic goals and select the proper strategy to achieve them. Regarding metacognition, Akyol and Garrison (2011) developed a metacognitive construct with the following three metacognitive dimensions: Knowledge of cognition, monitoring of cognition, and regulation of cognition. Knowledge of cognition refers to the awareness of oneself as a learner and the awareness of one’s knowledge, and skills concerning personal cognitive processes. Monitoring of cognition refers to the willingness to reflect upon the learning process, and involves understanding progression, assessing tasks, and making judgments about content validity. Regulation of cognition refers to the interactive aspect of metacognition when students are engaged in providing or asking for help from others to mutually improve their learning experience. This theoretical framework can be used to explore cognitive and metacognitive development, and it was validated by Akyol and Garrison (2011) for assessing metacognition in online discussions.
Online music learning

Several studies, primarily in school settings, have examined the application of information and communication technology (ICT) in music composition (for a review see Biasutti, 2012). Although prior studies in the music domain have focused on the use of technology, they were generally conducted in face-to-face environments. Furthermore, the few studies that were conducted in a virtual environment primarily pertained to instrumental practice rather than music composition. For instance, Seddon and Biasutti (2009) investigated the efficacy of a music e-learning resource for playing improvised blues. A qualitative research approach was adopted that consisted of videotaped observations and semi-structured follow-up interviews. The findings showed that all the participants successfully performed the musical activity in the online setting, and that the online activities facilitated the development of abilities such as planning, organizing, monitoring, and assessing one's competencies. In addition, the following aspects of the e-learning setting were considered helpful: The use of particular topic themes, the synergy between theory and practice, the flexibility of the work schedule, the easy use of the platform, and the full access to essential tools.

Summary of the theoretical background

The reviewed literature indicates several factors that influence the effectiveness of online collaborative learning and metacognitive processing. Regarding online collaborative learning, factors such as the sharing of knowledge and skills, the analysis and integration of different points of view, group planning, workload management, the establishment of trust for teamwork, individual accountability, and commitment toward quality work have been identified. In addition, social aspects such as familiarity with team members, positive interdependence, team cohesion, and a sense of community have been considered. Other aspects of online collaborative learning include technology competence, technology utility, and technology resourcing.

Regarding prior research conducted in the music domain, ICT tools have been considered suitable for facilitating musical creativity, and for enabling online instrumental music learning (Seddon & Biasutti, 2009). Because previous studies on collaborative music composition were primarily conducted in face-to-face environments (Biasutti, 2012), the current study aimed to address this gap in the literature by examining collaborative music composition in a virtual environment. In so doing, the study aimed to identify the successful aspects of the online collaborative activities, and to examine the metacognitive dimensions of online music collaboration.

The current research

Research design and questions

The current research is a pilot study that aimed to test the efficacy of an e-learning environment built to enable participants to collaboratively compose music online. The research design is qualitative: Individual, semi-structured interviews were conducted with the participants after the learning sessions to analyze the participants' perspectives regarding the functionality of the e-learning environment, the resources of the e-learning platform, and their overall experience with the e-learning process. In addition, the forum discussions were analyzed with respect to the metacognitive dimensions. Although the study is exploratory by nature, the following research questions were considered:

- Did the virtual environment enable the participants to compose a satisfactory music piece?
- Did the virtual environment enable the participants to utilize the metacognitive dimensions?
- What were the participants' perspectives regarding composing music through the online learning activity?

Method

Participants

Experienced musicians were involved in the online learning activity. In this way, it was possible to compare previous face-to-face composition experiences with the current online composition experience. The participants (n = 3)
included the following three musicians: Marco (guitar), Matteo (bass guitar), and Paolo (computer/keyboards/vocals). All the participants were Italian, and had prior experience with formal instrumental music instruction in private music school, and a music conservatory. The participants had approximately twenty years of experience in performing in rock bands, and their ages ranged from 37 to 39 years (mean age 38). The participants are members of the band Reeta Pawone, which was formed in 2001. The music genre performed by the participants was electronic rock, a form of rock music that involves the use of computers, and other electronic instruments to generate sounds. The group has recorded two CDs, and is working on a third. One participant was studying at a university in the northern Italy at the time of the research.

Equipment

Moodle was used as a platform for the online activities, and an e-learning environment was designed to work both asynchronously and synchronously. The tools that were used for working asynchronously included a database, a blog, a diary set up as a wiki, and several forums for discussing ideas related to composing music, the music content, and technical issues. In addition, the following software was used for the synchronous activities: ooVoo for video and eJAMMING for audio. ooVoo software was adopted because it allows for a real-time video connection of more than two people. However, the audio quality of ooVoo is poor for music making, and for this reason, eJAMMING was used for the audio. A Dropbox database was used to upload, and share the live rehearsals. The participants used PCs with webcams to interact with one another.

Procedure

The task assigned to the participants was to collaboratively compose a new piece of music that could be used for their repertoire within the online environment. No style or genre constraints were imposed. The proposed task is an authentic activity for musicians, and is not an artificial experimental task. This design provided ecological validity to the study, and offered a meaningful and motivating activity for the participants (Kump, Moskaliuk, Dennerlein & Ley, 2013).

An online tutor was available on the platform. During the virtual work, the participants had to respect the following rules:

- Compose a new piece of music, not a rework of a previous piece of music.
• Perform the work online and in the designed platform only. Do not discuss the new piece of music on the phone or in the presence of another.
• If you are working individually, please take notes and inform the other participants of your work through the platform. Please use the forums and the other tools for these purposes, and consider that you can also upload additional multimedia files to the platform if necessary.
• Keep a weekly diary of the work progress in which all the participants can contribute with their observations. The diary is set up as a wiki tool, so everybody can integrate or modify the text.
• Immediately inform the tutor about any inconvenience or technical problem.

The participants worked asynchronously on the platform to define general aspects, exchange ideas, and develop the composition process. In addition, the participants interacted synchronously to experiment with their ideas in four real-time sessions of approximately ninety minutes each, as shown in Figures 1 and 2. The online activities took place over a period of approximately two and a half months.

The semi-structured interviews

At the end of the activities, the researcher conducted individual semi-structured interviews with the participants to collect their reflections on the learning experience. Broad questions were proposed to the participants to provide them the opportunity to describe, and assess their experiences. Any further problems raised by the participants were investigated through additional questions. The following questions were proposed:
• How did you collaborate during the online activities?
• What role did the members have during the online activities?
• Please indicate the differences between composing music online, and composing music face-to-face.
• What were the benefits of the online collaborative activities?
• What were the disadvantages of the online collaborative activities?
• What were the most, and least useful tools provided by the online platform?
• In what way were the blogs, diaries, and forums helpful?
• What kind of problems did you have, and how did you resolve them?
• What new perspectives did you develop from collaboratively working online?

The semi-structured interviews were audio recorded and transcribed verbatim for analysis. Before the results are presented, it may be helpful to provide a description of the virtual activities to explain the tasks involved in the online collaborative activities.

Description of the online activities

During the online activities, the participants actively collaborated through experimentation by sharing and discussing ideas, composing and performing music, assessing the composition, and making decisions together. The participants were engaged in a collaborative process in which their ideas were rehearsed and discussed. The work started with listening to music by other rock bands online followed by a discussion. The aim was to draw inspiration from listening to the atmosphere of music by other bands, and to define general aspects of the musical piece to be
composed. The live sessions started with improvisation, which provided a way for the participants to experiment with their ideas, and musical material. Paolo then proposed a backing track facilitated by a computer, and the participants improvised over the backing track. Paolo manipulated the backing track on the computer during live composition sessions, and the other musicians developed their ideas for their tracks. All the video and audio recordings of the live sessions were uploaded on the platform and reviewed by the participants. The participants exchanged ideas on the platform about the improvisations and the musical material developed during the sessions, offering new ideas and discussing other aspects related to the music. The participants then evaluated the various sessions, and collectively decided what to select and how to develop the musical material to be included in the new music piece.

The participants interacted online via the platform among several different forums, a diary (activated within a wiki resource), and a blog. A screenshot of the composition discussion forum is provided in Figure 3.

![Figure 3. Screenshot of the composition discussion forum](image)

The platform was useful to plan the work, organize virtual meetings, discuss technical issues, develop the composition process, and share feelings. Regarding the forums, the participants established the following seven forums: Listening, software for online performance, session planning, composition, instruments, technical, and events. Although the forums had different names, the topics overlapped because, e.g., composing issues were discussed in both the composing forum, and the listening forum. A detailed description of the online interactions is reported in Table 1.
The final piece was recorded live during the last session. It was constructed with layers of sound, consisting of musical “loops” over which the participants played their instrumental parts. The final piece was not written down in a formal way but was recorded on multi-track software.

<table>
<thead>
<tr>
<th>Online resource</th>
<th>Interventions</th>
<th>Content description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forum: Listening</td>
<td>59</td>
<td>Links to music by other rock bands for drawing inspiration, and listening to the atmosphere of other music; detailed comments on the music; proposals of demos.</td>
</tr>
<tr>
<td>Forum: Software for online performance</td>
<td>7</td>
<td>Characteristics of the software for online performance; passwords; comments on the functionality of the software; and technical problems such as the latency of Internet streaming.</td>
</tr>
<tr>
<td>Forum: Session planning</td>
<td>29</td>
<td>Planning of the synchronous (real-time) sessions; appointments for the sessions (synchronous activity); discussions of technical resources; suggestions for improving the organization of the composition activity.</td>
</tr>
<tr>
<td>Forum: Composition</td>
<td>20</td>
<td>Comments on and detailed evaluation of the live sessions performed by the group; proposals for selecting and organizing the material; proposals for modifying sounds, and parts of the music piece.</td>
</tr>
<tr>
<td>Forum: Instruments</td>
<td>3</td>
<td>The musical instruments that were available for the overall project, and the instruments that the musicians intend to use during individual sessions.</td>
</tr>
<tr>
<td>Forum: Technical</td>
<td>11</td>
<td>Discussions about the technical set up and technical problems.</td>
</tr>
<tr>
<td>Forum: Events</td>
<td>7</td>
<td>Events and concerts by other rock bands.</td>
</tr>
<tr>
<td>Blog</td>
<td>1</td>
<td>Personal feelings.</td>
</tr>
<tr>
<td>Diary</td>
<td>12</td>
<td>Descriptions of how the activities proceeded, including personal comments.</td>
</tr>
</tbody>
</table>

**Results**

The data consisted of discussions on the forums, transcripts of the semi-structured interviews, and video recordings of the synchronous sessions. For the purposes of this paper, only the forum discussions, and semi-structured interview transcripts were considered.

**Analysis and results of the forum discussions with respect to the metacognitive dimensions**

The Akyol and Garrison’s (2011) framework has been used to analyze the forum discussions. Accordingly, the analysis focused on the following three metacognitive dimensions described in the theoretical background section: Knowledge of cognition, monitoring of cognition, and regulation of cognition. The coding process was later validated by an independent researcher who separately checked the coding to ensure that the coding actually reflected the metacognitive dimensions considered. The following quotes are coding examples for each of the dimensions of metacognition.

*Knowledge of cognition*

...I think, the use of the computer especially when you do electronic music, allows you to create music more focused on the sound rather than music designed to be arranged.

...Conversely, the advantage of the PC is to give you an opportunity to experiment with sounds.

...I would use this tool in the future, at least in the process of composing music, and for preparing the material.

...Music on SoundCloud more and more allows you to break down the song form, to get a sense of sound pieces but without a clear beginning and end.
Monitoring of cognition

...Pleasant creation of a carpet of sounds with a minimal rhythm section.
...Yes, but we are adding ideas on ideas...
...Next time, it would be better to do a more analytical job, because improvising goes well, but then you have to synthesize.
...The rhythm section is still quite rugged.
...It would be worth considering in detail various parts because the general idea is clear, but we have to arrange it in detail.

Regulation of cognition

...Why you do not like the drum at 3’ 57”?
...Drum: First and second parts. Paolo, can you change the sound of the snare? Something more soft,...
...But, I am more interested in collecting your impressions...
...I must admit that I did not get your points...
...Listening to various music pieces with you, I understood that...

Analysis and results of the semi-structured interviews

The transcripts of the semi-structured interviews were analyzed by using an inductive method based on grounded theory. An adaptation of the constant comparative method, which has been successfully used in previous qualitative studies (Biasutti, 2011, 2013), was used. The following five phases were adopted: (1) immersion, in which all the discernibly different behaviors are recognized; (2) categorization, in which categories appear from the discernibly different behaviors; (3) phenomenological reduction, in which themes come out from the categories; (4) triangulation, in which supplementary elements are used to sustain the researchers’ interpretations; and (5) interpretation, the final step in which a complete explanation of outcomes is provided in connection with previous research and/or models. In the immersion phase, the transcripts were read several times to acquire a high level of familiarity with the raw data, and then, the discernibly different answers were identified. In the categorization phase, similar behaviors were sorted into similar categories, and 17 distinct categories emerged from the discernibly different answers. Four themes were derived from the 17 categories during the phenomenological reduction phase: teamwork, the platform, face-to-face/online differences, and strengths/weaknesses. A diagrammatic version of the first three phases of analysis for the semi-structured interviews is provided in Figure 4.

Support for the researcher’s interpretations of the themes was provided through the process of triangulation. Examination of the semi-structured interview transcripts and forum interactions indicated that the participants independently referred to the researcher’s interpretations of the themes. The data supporting the process of triangulation are reported in Table 2.

The coding of the semi-structured interview data was later validated by an independent researcher who separately checked the data coding. The original researcher and the independent researcher discussed any possible disagreements related to the coding. Changes to the original coding were made accordingly. Below, detailed descriptions of the emergent themes are reported.

Table 2. Supporting quotations for the themes of the semi-structured interviews

<table>
<thead>
<tr>
<th>Themes</th>
<th>Supporting quotations</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teamwork</td>
<td>“Performing the instrument for the electronic programming, Paolo has a structural role…”</td>
<td>These quotations support the interpretation of a teamwork theme because the participants refer to roles and collaboration during the collaborative activities.</td>
</tr>
<tr>
<td></td>
<td>“We have revised the various parts in a sequence that was shared by all of us.”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Working in groups is positive for creativity”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“The most interesting thing was the use of the forum…”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“…the diary was used primarily by me,…”</td>
<td></td>
</tr>
<tr>
<td>Platform</td>
<td>“The most interesting thing was the use of the forum…”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“It was more important to talk about…”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“…we should have been more organized…”</td>
<td></td>
</tr>
</tbody>
</table>

56
“No problem with the platform, it is easy to use”
“Certainly, the online work was more flexible.”
“The online work was more systematic, better organized.”
“Working online made the process followed during the development of the musical material more explicit.”

These quotations support the interpretation of a face-to-face/online differences theme because the participants compared the face-to-face modality with the online activities and processes.

Strengths and weaknesses
“The advantage of working online is the possibility to operate within my own home environment,...”
“From the point of view of time, it is a big advantage to work online.”
“There were no disadvantages for my part, with the exception of the feeling of a lack of something compared to the normal sessions in a physical place,...”

These quotations support the interpretation of a strengths/weaknesses theme because the participants referred to advantages and disadvantages of the online work.

---

**Figure 4. Diagram of the three phases of analysis for the interview data**

**Teamwork**

The participants considered teamwork to be essential during the online activities, and the coding process revealed the following categories: roles and collaboration. Regarding roles, the participants acknowledged that they have particular roles in making music, and each participant had his/her own specialty. Paolo uses a computer, and his role is to define the structural part of the piece and to design the backing track, while Matteo and Marco focus more on the sound and the music phrases to be arranged onto the backing track. Matteo is more proactive than Marco in composing the music, while Marco focuses more on the sound effects of the guitar. According to Matteo,
“Performing the instrument for the electronic programming, Paolo has a structural role. The construction of the music piece then revolves around the various proposals and ideas. I think Paolo and myself were more proactive, while Marco was less purposeful because he spends more time on the sounds.”

Collaboration was a key aspect at all stages of the online activities from the generation of ideas to the composition of the music. The collaboration developed through virtual discussions and activities, and the participants exchanged ideas and shared decisions in a democratic setting. “Together we contributed to the production of the musical material... (Paolo).” “The arrangements were developed following a series of discussions between the three of us in relation to the material produced with the improvisation... (Marco).” In addition, the collaboration involved sharing ideas and principles, “The samples produced during improvisation were then reprocessed... on the basis of exchanged impressions in the forum after listening (Paolo).” “We have revised the various parts in a sequence that was shared by all of us... (Paolo).” These quotes highlight the collaboration that occurred during the online activities, and the creative learning environment that was established. The discussions enabled the participants to clarify their thoughts, and to make relevant decisions when composing the music piece. The participants were considered equals, and the virtual learning environment facilitated egalitarian participation, as reported by Paolo, “The online situation has permitted a more horizontal process, and less binding to the original proposal.”

**Platform**

The platform was deemed functional, and several useful tools for the online work were highlighted by the participants. The related categories included more suitable functions, less suitable functions, and problems.

The discussion forum and the database used for uploading the session recordings were considered to be the more suitable functions of the platform. The forum gave continuity to the composition process, as reported by Paolo, “The most interesting thing was the use of the forum that gave me a sense of continuity while working, which is lost in a face-to-face situation. Face-to-face, when we finish the sessions, “we pull the plug,” and everyone reworks the material on his own.” Thus, the forum constituted a tool that helped the participants share their opinions, and more explicitly reflect on the process. The platform further strengthened the sense of community through practice based on collaborative work. Sharing recordings was important because it enabled the participants to listen, and identify fragments for potential revision.

The participants identified the diary as a less suitable function of the platform. Specifically, the diary was not considered particularly useful for articulating the work, as reported by Matteo, “Well, the diary was used primarily by me, then, seeing that the others were not very attracted to this thing...I have not used it a lot.”

Regarding problems, the encountered issues were mostly technical, e.g., related to the availability of broadband Internet, as a very powerful connection was required for the synchronous sessions.

**Face-to-face/online differences**

The participants shared their opinions regarding face-to-face versus online interaction for composing music. Specific aspects of the online work, for instance, related to the organization of the work and work method, the composition process, time management, technical resources, and the achieved results, emerged.

Regarding the organization of the work and work method, it was reported that working online is more systematic than working face-to-face. As Paolo stated, “The online work was more systematic, better organized.”

Regarding the composition process, the participants stated that it closely resembled face-to-face collaboration with respect to the basic operations, although there were differences in the overall process. Moreover, the basic principles that were followed to compose the music piece and the efforts to produce the musical material were considered to be similar in between two different settings, “I think we set up the online work more or less like the one we set up face-to-face, where a large part of the exchange of opinions, ideas and suggestions is via email, SMS, or phone. Here, we have been working exclusively online (Marco).” However, it was reported that the online tools facilitated the comprehensive composition process, and the development of a metacognitive dimension. As Paolo asserted, “Working online made the process followed during the development of the musical material more explicit.”

statement demonstrates that the online environment enhanced the participants’ reflection on, and awareness of the performed task.

Working online also increased the flexibility of the work, and facilitated time management. As Matteo reported, “There was definitely better time management due to the fact that everyone needed to make fewer trips with equipment, tools, etc.”

The technical resources and technology involved in working online were deemed to be more developed than those in a face-to-face setting because the online setting required a fast Internet connection to minimize the latency of the signal transmission during the synchronous sessions. As Paolo stated, “Good communication requires a telecommunications system that can offer a good capacity for data transmission. For the rest, it does not require additional technical resources than those used in face-to-face situations.”

With regard to the achieved results, the participants expressed satisfaction with the work performed. As Marco noted, “I am personally satisfied both with the music piece, and with the online experience....The music piece is a bit different from what we have composed so far, more ambient, almost as a kind of soundtrack....Interestingly, the music offers good suggestions on what to experiment with in the coming months.” Matteo, however, encountered difficulties in making aesthetic comparisons between the composed piece, and their previous pieces, “It is difficult to compare this music piece with the previous ones...; we should make others in this way, and then you could make a comparison with those made previously in a face-to-face situation....Anyway, I am pleased with the finished composition, and also with the experience. Results fully achieved.”

Strengths/weaknesses

The participants indicated that the online process entailed logistical, organizational, emotive, and compositional strengths, as well as technical, communicative, and emotive weaknesses.

Regarding the logistical strength of the online process, Paolo stated that “the advantage of working online is the possibility to operate within my own home environment, without having to move my equipment.” Regarding resources, Marco remarked about “the opportunity to have a whole range of instruments and materials that you may not have in an external rehearsal because they are at your home, and not in the rehearsal room.” Similarly, Matteo recognized “the opportunity to have access at any time to the archives, and your instruments. Because we do not have a dedicated rehearsal room, at the end of each session, we must move all the stuff. This is the reason that many times we do not bring with us things that maybe could have been useful during the session.”

In discussing the organization of the online process, the participants unanimously expressed the advantage of being able to individually manage their time, “Surely the possibility to optimize time, and space (Matteo);” “the possibility to make the most use of your time (no movement, no downtime) (Marco);” “the possibility to work without space-temporal or practical constraints.”

Regarding the emotive strength of the online process, Marco reported that a more comfortable situation occurred by rehearsing at home, “working at home, I have developed a greater feeling of tranquility.”

Regarding the composition process, the participants noted that working online was more reflexive than working face-to-face. Indeed, the online environment affected the overall composition process by facilitating the stylistic elaboration of the material, as reported by Matteo, “The impression is that there is a continuity of style while composing online.” Moreover, the participants developed a more logical approach to composing music based on systematic reflection. As Paolo reported, “I had the feeling of working in an arranged, perhaps even more logical way: improvisation, discussion, and processing.” The online collaborative work further opened several operational perspectives. For example, Paolo stated that “there is the possibility of enlarging in a theoretically infinite way the possibility of exchanging, and processing the musical material....From a compositional point of view, this was a new perspective that opened an exciting new dimension for me.” In addition, Matteo reported that “having a virtual interaction opens a new window for creativity development; this is very positive.”
The technical disadvantages that were noted in the online collaborative process included some latency in the signal transmission. Such latency created an aerial effect that was appreciated by participants. However, it was not possible to work on very precise rhythmical solutions within such circumstances. As Matteo asserted, “...while in a situation like this one, where maybe you have even a little latency, you cannot be very precise, although the program works very well...but you may be stimulated to work on the atmosphere of the music...I mean, it developed better...Face-to-face..., you have more possibilities to use the rhythm than here...” Another technical disadvantage related to the audio volume, which could not be as loud in the online sessions as in face-to-face sessions, and this lack of power influenced the participants’ feelings associated with the music. As Matteo reported, “Concerning the rhythm, for example, in a face-to-face situation, you feel it more, because the sound comes out more aggressively since you have a more powerful amplifier in front of you.”

Regarding communicative disadvantages, Marco noted an aspect of visual communication, “Sometimes, it was difficult to interact visually or gesturally. In addition, when listening to what was done previously, it was sometimes hard to understand when the other participants were listening or when they were interacting with the music, perhaps re-improvising on ideas already expressed.”

Regarding how to improve online collaborative work, the participants identified the systematic nature of the work, “We must be more systematic during the elaboration of the material; some ideas are often left out too quickly. However, this is also related to a compromise between the need to produce the material, and the time available (Paolo).”

**Discussion**

Regarding the first research question, the online environment was considered effective because the participants were able to compose a new piece. Moreover, the participants positively evaluated the experience, and the new music piece, indicating that collaborative composition activities can be performed in virtual environments. This result extends findings reported by previous research regarding the use of ICT for composing music.

Regarding the second research question, the results showed that the participants engaged in metacognitive processing during the online composition activities. The online work proceeded at a metacognitive level, and the participants were stimulated to reflect on their actions, and decisions. As noted in Matteo’s diary, “The work goes on, and we are now mastering the technological resources; therefore, we focus on creativity.” This quote provides evidence that the online learning activities stimulated the development of higher-order cognitive abilities in the participants, which is consistent with previous findings reported by Biasutti (2011), and Lin and Jou (2013). In addition, the metacognitive analysis conducted with the framework developed by Akyol and Garrison (2011) provided evidence that the participants reflected on their knowledge, evaluated the progress of the activities, and regulated their cognitive resources based on the roles of their bandmates. These findings are in line with Akyol and Garrison’s (2011) results regarding the distinct metacognitive dimensions. The online learning activities also stimulated cognitive processes in the participants, as they made aesthetic evaluations, synthesized different perspectives, and developed critical thinking skills. This result is consistent with the findings of Tseng and Yeh (2013), who reported that processes associated with teamwork can promote critical thinking skills in students. In the current research, the participants reported that the online environment allowed them to develop an even more logical approach to music composition based on orderly reflection than the face-to-face setting, as the online environment allowed them to organize their activities, and supported them in composing the music piece in a sequential fashion.

Regarding the third research question, the participants identified components relevant to the online composition experience. These included teamwork, the platform, face-to-face/online differences, and strengths/weaknesses. The platform was considered useful for planning the work, organizing online meetings, discussing technical issues, developing the composition process, and sharing feelings. Indeed, collaboration and teamwork were considered key components of the online work (Donnelly & Boniface, 2013; Tseng & Yeh, 2013). Other comments related to the
organization of the work and work method, time management, and technical resources. The participants also noted strengths and weaknesses of the online experience. These strengths and weaknesses can be exploited to provide guidelines on how to utilize online tools more efficiently to support future research projects involving composing music in an e-learning setting.

The findings of the present study contribute to a better understanding of the usefulness of collaborative online environments and resources, and reveal certain aspects of the virtual setting that render online collaborative activities efficacious for learning. These aspects include the opportunity for collaboration, the dynamics of the group, and the use of appropriate technology (Biasutti, 2011). In addition, in line with Tseng and Yeh’s (2013) findings, other aspects, such as individual accountability, familiarity with team members, commitment toward quality work, and team cohesion, were found to affect the online collaboration process. Overall, several factors affected the online collaboration process, including the learning environment, the platform, the technology resources, the level of challenge, and the nature of the activity.

The learning environment was important for the online collaboration process because it was simple yet effective, and because it offered the participants an opportunity to collaboratively work online. The participants interacted continuously while composing a piece of music, showing commitment to the task.

The platform provided support for the online collaboration process and facilitated collaborative work and actions such as sharing during listening situations, as well as discussing and developing a cooperative decision process. The platform fostered a sense of community based on cooperation and collaboration, as reported by Donnelly and Boniface (2013). The participants felt reflective, and actively engaged in the composition process, demonstrating commitment and responsibility (Tseng & Yeh, 2013).

Regarding technological resources, the software programs that were used for the synchronous sessions performed satisfactorily, and full access to necessary resources was a key aspect of the effectiveness of the online activities for composing music. Indeed, previous studies have demonstrated that adequate technical solutions are crucial for achieving high-quality results (Valentín, et al., 2013).

The participants noted the challenge of the activities, and they were actively engaged in performing the activities. As Matteo noted, “Working online is a new and exciting experience...; performing online is particularly challenging and compelling for its employment of technology.” This quote demonstrates that the online activities were considered challenging, and that they effectively stimulated the involvement of all the participants, as occurred in the study by Seddon and Biasutti (2009).

With regard to the nature of the activity, the assigned task — to collaboratively compose a music piece regardless of style or genre constraints — was simple and useful, and allowed the participants to interact naturally in the virtual environment. The task was an authentic activity for musicians, and the use of authentic activities is crucial for facilitating collaboration within virtual settings. The nature of the learning environment was also important, as online activities should be performed in a realistic setting (Kump et al., 2013). The authenticity of the activities strengthened the ecological validity of the study, and stimulated the involvement of the participants. Moreover, the open-ended nature of the task facilitated the development of collaborative work, motivating the participants to achieve at least a functional level of cooperation. The flexible work schedule was another relevant aspect, as it allowed the participants to feel more comfortable in making decisions about the work activities, which is consistent with Seddon and Biasutti’s (2009) results. Other elements such as constant interaction, deep involvement, and collaboration among the participants were crucial for rendering the online activities effective, as reported by Tseng and Yeh (2013). In addition, Anderson and Simpson (2004), and Seddon and Biasutti (2009) have reported similar findings, highlighting the effects of the instructional design on the interactions among participants.

Conclusions and future work

There are several limitations to the current study. The results must be considered in relation to the study’s exploratory nature because only one music group was involved. Because of the limited number of participants, it was not
possible to include quantitative methods, so only qualitative methods were used. The findings provide some insights into a possible methodology for evaluating the online tools within a virtual setting. However, the generalizability of the results is limited by the study’s methodological design. In the future, it would be interesting to design a study with a larger number of participants to increase the validity of the results. Another limitation pertains to the assessment of the final product, which was not evaluated by external experts. The music piece was only evaluated by participants, who were satisfied with the final output.

The results of this study have implications for research on online music composition, and suggest the need for further research on the nature of the processes involved in online music composition. The virtual environment used in the study was effective for fostering online music composition, and could be applied to develop online music composition projects involving participants from different parts of the world. Regarding the technical implications of the findings, the synchronous resources were important tools for developing ideas through improvisation, and they allowed participants to collaboratively compose their music piece. In further research, it would be interesting to expand the use of asynchronous software for collaborative online music composition in wiki environments, and to evaluate their effectiveness. In addition, other music genres, such as contemporary and electronic music, could be considered.

References


Examining the Effect of Academic Procrastination on Achievement Using LMS Data in e-Learning

Ji Won You
Department of Early Childhood Education, Gachon University, Korea // uimagine@gachon.ac.kr

(Submitted February 11, 2014; Revised September 29, 2014; Accepted November 26, 2014)

ABSTRACT
This study aimed to investigate the effect of academic procrastination on e-learning course achievement. Because all of the interactions among students, instructors, and contents in an e-learning environment were automatically recorded in a learning management system (LMS), procrastination such as the delays in weekly scheduled learning and late submission of assignments could be identified from log data. Among 569 college students who enrolled in an e-learning course in Korea, the absence and late submission of assignments were chosen to measure academic procrastination in e-learning. Multiple regression analysis was conducted to examine the relationship between academic procrastination and course achievement. The results showed that the absence and late submission of assignments were negatively significant in predicting course achievement. Furthermore, the study explored the predictability of academic procrastination on course achievement at four points of the 15-week course to test its potential for early prediction. The results showed that the regression model at each time point significantly predicted course achievement, and the predictability increased as time passed. Based on the findings, practical implications for facilitating a successful e-learning environment were suggested, and the potential of analyzing LMS data was discussed.

Keywords
Academic procrastination, e-Learning, Achievement, LMS data analysis

Introduction
Procrastination refers to the delay of initiation or timely completion of a task (Lay, 1986). Procrastination is a prevalent phenomenon especially among college students due to the relatively flexible learning environment compared to K-12 circumstances. Some studies (e.g., Ellis & Knaus, 1977; Schouwenburg, 2004) reported that approximately 70% of college students considered themselves procrastinators, and academic procrastination has been reported across the world (e.g., Ferrari, Diaz-Morales, O’Callaghan, Diaz, & Argumedo, 2007; Klassen et al., 2010; Seo, 2011). In addition, as e-learning is becoming popular, the likelihood of procrastination is also growing because of the self-directed nature of the e-learning environment. For example, an annual report (Allen & Seaman, 2013) that surveyed over 2,800 colleges and universities in the U.S. reported that a total of 6.7 million students had taken at least one online course, but many students expressed difficulties in planning and regulating their studies.

Studies have shown that procrastinators delay or participate less in learning activities due to a failure in self-regulation (Dewitte & Lens, 2000; Steel, 2007), and procrastination leads to poor achievement (Akinsola, Tella, & Tella, 2007). Because procrastinators had a relatively short period of time for completing tasks or preparing for exams, they hurried to finish their work or crammed the study material. Furthermore, high procrastinators have reported that they feel a compulsion to dropout frequently throughout a course and experience difficulty in maintaining motivation (Michinov, Brunot, Le Bohec, Juhel, & Delaval, 2011). Therefore, procrastination has been considered a proxy for disengaged learning, and it negatively impacts achievement (Asarta & Schmidt, 2013; Melton, 1970). Delays in studying weekly learning materials, participating in activities, and completing tasks are typical procrastination in e-learning, and the consequence of procrastination in e-learning is considered more damaging on achievement than procrastination in traditional classrooms (Tuckman, 2005) and often results in withdrawals (Doherty, 2006). Therefore, special attention should be paid to procrastination in e-learning.

However, there is not enough empirical research on academic procrastination in the e-learning environment. Furthermore, most procrastination studies have used self-reported questionnaires to measure the behavioral tendencies of procrastination either in academic tasks or in daily tasks. Although these instruments (e.g., Lay, 1986; Tuckman, 1991) have been validated and used in many studies, they are intrusive to students and limited in capturing actual procrastination behaviors displayed during a course. This issue can be resolved in an e-learning course because all student behaviors are recorded in a learning management system (LMS). The generated log data convey
information regarding who did it, what was taken, and when it was taken (Hershkovitz & Nachmias, 2009). More than 80% of surveyed institutions in the U.S. (Abdous, He, & Yen, 2012) and Korea (Korea Education and Research Information Service, 2012) have reported using LMS as a medium for e-learning course delivery, and the majority of traditional universities have also used LMSs to support their regular classes. This trend indicates that it is not difficult to trace and identify students’ learning patterns throughout a course in a field. Therefore, it would be valuable to clarify the effect of academic procrastination on achievement using LMS data.

Several studies have explored the significant indicators from log data to e-learning achievement. Participation indices, including the number of pages to read, time spent reading materials, and the number of original postings, were commonly used in the research. However, recent studies (Asarta & Schmidt, 2013; Jo & Kim, 2013) have claimed that consistent learning behaviors, such as avoiding cramming or regularity of learning pace, are more critical than the frequency of access or amount of time spent viewing. This argument implies that regular study reflecting self-regulation is vital for successful learning.

This study aimed to investigate the effect of academic procrastination as measured in LMS data on e-learning course achievement. Among the indices from log data, the present study selected two procrastination indicators that reflected a failure of self-regulation: delays in weekly scheduled study and late submission of assignments. Because each individual’s access to specific materials and submission of an assignment were time-stamped, delays in accessing scheduled learning materials and submitting assignments after the due dates could be easily identified. Additionally, the present study attempted to test the predictability of academic procrastination on course achievement at different time points as a course proceeded. This approach aims to validate the potential of using log data as a feasible research tool for improving pedagogical effectiveness.

Theoretical background

Academic procrastination

Academic procrastination is generally described as a failure of completing an academic task within the expected time frame (Senécal, Koestner, & Vallerand, 1995) or delaying academic work that must be completed (Schraw, Wadkins, & Olafson, 2007). Past research on procrastination has tended to focus on investigating the relationships among personal traits or psychological factors and procrastination. Perfectionism (e.g., Ferrari, 1992), self-handicapping (e.g., Beck, Koons, & Milgrim, 2000; Ferrari & Tice, 2000), personality (Schouwenburg & Lay, 1995), fear of failure (e.g., Solomon & Rothblum, 1984), etc., have often been discussed as leading to procrastination. However, some studies have claimed that procrastination is a failure in self-regulation (Dewitte & Lens, 2000; Klingsieck, Fries, Horz, & Hofer, 2012; Steel, 2007; Wolters, 2003). For example, Ferrari (1991) reported that there was no intellectual difference between procrastinators and non-procrastinators, but the level of regulation skills was different between them (Dewitte & Lens, 2000; Rabin, Fogel, & Nutter-Upham, 2011; Wolters, 2003). Furthermore, Wolters (2003) emphasized that frequent procrastinators had characteristics opposite those of self-regulated learners. Wolters described that self-regulated learners had a wide range of cognitive strategies and were able to choose effective ones for their learning. Additionally, self-regulated learners utilized metacognitive skills and possessed adaptive motivation and attitudes, responding to a variety of academic contexts. He argued that these attributes all together allowed self-regulated students to actively engage in their learning and that such students were expected to procrastinate less on their academic work.

Many studies have illustrated a negative relationship between procrastination and academic outcomes in the traditional classroom (Akinsola, Tella, & Tella, 2007; Ferrari, Johnson, & McCown, 1995; Fritzsche, Young, & Hickson, 2003; Klingsieck et al., 2012; Tice & Baumeister, 1997). Melton (1970) argued that students who crammed study material at the last minute showed poorer long-term retention of the material than those who studied regularly. Additionally, frequent procrastinators often underestimated the time required to complete a task and did not allocate sufficient time to complete the task (McCown, Petzel, & Rupert, 1987), so they easily failed if an unexpected event happened (Ferrari, Johnson, & McCown, 1995). These findings suggest that procrastination should be avoided in the pursuit of academic success.
Academic procrastination in e-learning

People consider that procrastination in the e-learning environment can be more detrimental because learning depends more on the individual learner. Elvers, Polzella, and Graetz (2003), who described students’ behaviors during e-learning, reported that students in the course exhibited heavy access to the course web pages either the day before or the actual day of an exam. The authors discussed how students easily postponed studying until the last minute because they did not adhere to a strict course schedule. Such cramming behavior has been commonly reported in other studies (Asarta & Schmidt, 2013; Donnon, Violato, & DesCoteaux, 2006; Levy & Ramin, 2012), and the regularity of studying has been revealed to be a stronger predictor of achievement than the amount of study time (Jo & Kim, 2013). Song, Singleton, Hill, and Koh (2004) reported that students gave priority to time management to achieve success in e-learning, in addition to the design of the course, comfort with online technology, and motivation. Similarly, You (2012) empirically verified that self-regulated learning negatively influenced academic procrastination in e-learning.

In addition, several studies (e.g., Rakes & Dunn, 2010; Wighting, Liu, & Rovai, 2008) have examined the relationship between engaged learning and procrastination. Rakes and Dunn (2010) found that intrinsic motivation and effort regulation contributed to decreased academic procrastination among online graduate students. Wighting, Liu, and Rovai (2008) articulated that the level of participation or delay in the performance was a significant indicator of engaged learning, and they emphasized that students in the e-learning environment required more intrinsic motivation to prevent procrastination. Tuckman (2005) confirmed not only that procrastinators in e-learning tended to perform worse than non-procrastinators but also that the negative relationship between procrastination and achievement in the e-learning environment was stronger than that in the traditional learning environment. Therefore, more research is required to enhance the understanding of procrastination in e-learning.

Analysis of log data as a research tool

Most studies (e.g., Michinov et al., 2011; Rakes & Dunn, 2010; Wighting, Liu, & Rovai, 2008) on academic procrastination, however, have measured academic procrastination based on self-reported questionnaires instead of observed behaviors. In the e-learning environment, students’ timely learning behaviors can be easily obtained by accessing log data. Analyzing the students’ log data and identifying valuable patterns in a large dataset is called Educational Data Mining (EDM). Data mining is used extensively in e-commerce to increase sales and profit (Hershkovitz & Nachmias, 2009), whereas in higher education institutions EDM is used to advance educational outcomes.

Delavari, Beikzadeh, and Phon-Amnuaisuk (2008) argued that EDM enables instructors and institutions to make better decisions in administration, to predict students’ performance with more accuracy, to develop better strategies, and to allocate resources effectively. In particular, analyzing LMS data helps instructors identify at-risk students while students are learning and provide prompt support and guidance to them (Campbell, DeBlois, & Oblinger, 2007). For example, Blackboard, a popular LMS, provides an early prediction function of students’ performance by analyzing LMS students’ data (Wang & Newlin, 2002). Similarly, Course Signal, an early warning system implemented at Purdue University, provides three different colored signals similar to those of a traffic light to students in courses, with each color visually indicating the status of the students’ learning. Research showed that employing Course Signals helped at-risk students improve their performance and decrease dropout rates (Purdue University, 2013). This approach is called learning analytics, which consists in collecting data from LMS and using data mining techniques to identify instructional problems or students who suffer academically. Consequently, learning analytics in the academic domain enables instructors to provide interventions for at-risk students and enhance learning outcomes (Campbell, DeBlois, & Oblinger, 2007; van Barneveld, Arnold, & Campbell, 2012).

Log data, engagement, and academic achievement

In previous research, indicators of participation among log data have been frequently used to explain individual learning differences. For example, question response time and answer correctness were used to measure online learners’ engagement (Beck, 2004); the number of pages read, time spent reading pages, the number of tests, and time spent on tests were also used to demonstrate engagement (Qu & Johnson, 2005). Moreover, Morris, Finnegan,
and Wu (2005) empirically analyzed students’ behaviors in e-learning and examined the difference in online participation between withdrawers and completers and between successful completers and non-successful completers. Among log data, they used the number of content pages viewed, the number of discussion posts viewed, the number of original posts, and the number of follow-up posts as frequency of participation, and hours spent viewing the discussion pages, hours spent viewing content, hours spent creating original posts, and hours spent creating follow-up posts as duration of participation. The results showed significant differences in participation between withdrawers and completers and between successful completers and non-successful completers. Further, the number of discussion posts viewed, the number of content pages viewed, and the time spent on viewing the discussion pages were revealed as significant predictors of student participation and final grade. The authors reported that participation indicators accounted for approximately 31% of the variability in achievement. This study indicated that active participation in the learning process was essential for e-learning success and that log data could be useful in predicting the learning outcomes.

Some studies have analyzed procrastination and the observed participation patterns in e-learning and discussed their relationship to academic performance. For example, Michinov et al. (2011) divided students into high and low procrastination groups based on a self-reported procrastination questionnaire, and they investigated the tendency of participation in a discussion forum. The authors found that high procrastinators tended to participate less in a discussion forum and perform lower. They argued that procrastination negatively influenced participation in the discussion forum and performance and that participation mediated the relationship between procrastination and performance. Additionally, Michinov and colleagues compared the behavior trends between high and low procrastinators during a 10-week course. Unlike low procrastinators, high procrastinators reported the urge to drop out over time, which implied that they experienced difficulty in time management. Furthermore, the level of motivation between the two groups was different. Whereas low procrastinators maintained a high level of motivation over the course, high procrastinators lost their motivation shortly after beginning the course, and motivation increased abruptly at the end. This study revealed the contrasting learning patterns between high and low procrastinators over time during the course and confirmed the relationship between procrastination, engaged learning, and achievement.

Furthermore, Asarta and Schmidt (2013) examined the effects of students’ access patterns on course achievement in a blended course. The course was arranged to deliver online lectures and to offer supplementary learning activities offline. The authors collected data from 179 college students and used various access measures, including the amount of access, intensity, pacing, anti-cramming, reviewing, completeness, and consistency. As a result, pacing, anti-cramming, and consistency were found to be positive and significant predictors of course achievement after controlling for grade point average and math skills. It should be noted that anti-cramming had the strongest effect on course achievement, and the results supported the benefits of regular study. Warnock, Bingham, Driscoll, Fromal, and Rouse (2012) investigated whether an early posting on a course message board was related to the final grade among 12 classes. The results demonstrated that the average grade of the first posters was consistently higher than that of the last posters, and the length of the posts was also correlated with course performance. In summary, these findings evidenced that procrastination had a negative effect on achievement by showing that regular studying and timely participation were strong determinants of academic success and that observed learning behaviors were plausible predictors of learning outcome.

Although students’ recorded data in LMSs have been shown to be useful, it is challenging for instructors to browse and handle such a large amount of raw data (Zaïane, 2001). Not every number is meaningful, and there is still a lack of standard methods and measures in log data that can be used to comprehend students’ learning (Mahajan, Sodhi, & Mahajan, 2012). Therefore, more empirical research is required to identify and validate meaningful log-based indices for predicting students’ success and failure in learning.

**Purpose of the study and research questions**

This study aimed to investigate the effects of academic procrastination behaviors on e-learning course achievement by using LMS data. Delays in regular study and submitting assignments were used to represent academic procrastination in this study, and the absence score and frequency of late submission among LMS data were used accordingly as input variables. Furthermore, the study explored the predictability of academic procrastination on
course achievement at four different time points of a 15-week course (Time 1, Time 2, Time 3 and Time 4), if the predictability was significant enough to predict achievement. The research questions were as follows:

- Do academic procrastination behaviors significantly predict e-learning course achievement?
- Do the academic procrastination behaviors recorded at certain intervals (Time 1, Time 2, Time 3, and Time 4) throughout an e-learning course significantly predict course achievement?

Method

Participants and procedure

The study was conducted at a mid-sized 4-year university near Seoul, South Korea. Although the university is a traditional university, it offers a few e-learning courses every semester. A course was selected at the researcher’s convenience, titled “Introduction to Color.” The course was a three-credit course, and there was no prerequisite. A total of 600 undergraduate students were registered for the course, but 31 students dropped out within four weeks. Therefore, only the data from the 569 students who completed the course were used for the present study. The e-learning course continued for 15 weeks. The students were required to study the weekly scheduled online learning materials, complete four assignments, and take two offline examinations. Online lecture videos served as a primary medium for providing the learning materials, and questions and answers through the message board were used. All learning activities and communication occurred and were recorded in the LMS.

Instruments

Academic procrastination

The absence score and the frequency of late submission were chosen as indicators of academic procrastination. The absence score illustrated the delay in regular study, and late submission represented the failure to submit the assignment on time. The accumulated absence and late submission scores were obtained at four different points during the course (Time 1 ~ Time 4).

Absence

The participating school had a specific e-learning attendance rule. Because the course lasted for 15 weeks, the course materials were arranged into 15 weekly segments, and each week’s course materials were subsequently available. One point per week was assigned for virtual attendance, and weekly attendance was checked automatically by LMS. To earn a full weekly attendance point, students needed to meet the following two conditions. First, they had to access the weekly lecture video within a designated time period. If a student accessed and completed the content after the specified period of time, then it accounted as an absence. Students were prevented from accessing the upcoming content before the scheduled time, but they could review the past contents freely. Second, because the online lecture videos were the main type of the learning material in this study, students were required to spend sufficient time watching them. If the students left in the middle of playing the content or skipped part of the content, then they earned only half of an attendance point (.5) due to a shortage of studying time. The attendance score recorded in the LMS ranged from 0 to 13, excluding the scores for the two offline examination weeks (the midterm and the final exams), and the total absence score was subtracted from the attendance score. The absence score in this study indicated delay in regular access to the learning content and incomplete viewing of the lecture videos; thus, a high absence score represented high academic procrastination. Furthermore, to test the predictability of the cumulative absence score at four different times, the cumulative absence score from weeks 1 to 4 was used for the Time 1 absence; the cumulative absence score from weeks 1 to 8 was used for the Time 2 absence; and the cumulative absence score from weeks 1 to 12 was used for the Time 3 absence score. The Time 4 absence score was the same as the total absence score.
Late submission

The course instructor assigned four assignments. The fixed due dates for each submission were announced at the beginning of the class with the description of each assignment. The due dates of the assignments were at the end of weeks 4, 8, 12, and 15. As each assignment due date approached, the instructor reminded the students about the assignment and its due date. In the LMS, when a student submitted his or her assignment to the assignment folder, a time stamp was generated. One point was assigned to each late submission (ranging from 0 to 4), and a high score indicated high academic procrastination.

Achievement score

Course evaluations were based on attendance, midterm and final scores, and assignments. Two achievement scores were used as academic performance variables. One was the final achievement raw score for the course (ranging from 0 to 100), and another was the sum of the midterm and final exam scores (ranging from 0 to 60). Because attendance was a part of the evaluation in the course, the sum of the two exam scores was used to control the confounding relation between attendance and final course score.

Data analysis

Descriptive statistics and correlation analysis were conducted using SPSS 18.0. Multiple regression analyses were employed to test the proposed hypotheses. To check the assumptions of the multiple regression analysis, scatter plots, histograms of the standardized residuals (Stevens, 2009) and a P-P plot with normal distribution were examined. The variance inflation factor (VIF) between the predictors was also examined to identify multicollinearity. No critical violation of assumptions was identified. A significance level of .05 was used for the analysis.

Results

Descriptive statistics and correlation analysis

The descriptive statistics of absence ($M = .96, SD = 1.57$), late submission ($M = .56, SD = 1.00$), final course score ($M = 79.16, SD = 13.94$), and exam score ($M = 42.50, SD = 9.43$) as well as correlation analysis results among variables are summarized in Table 1. The low mean of absence indicated that the participants studied on a regular basis fairly well. Absence was negatively correlated with the final course score ($r = -.71$) and the exam score ($r = -.55$), and late submission was also negatively related to the final course score ($r = -.68$) and the exam score ($r = -.47$).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Absence</td>
<td>.96</td>
<td>1.57</td>
<td>.00</td>
<td>10.50</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Late submission</td>
<td>.56</td>
<td>1.00</td>
<td>.00</td>
<td>4.00</td>
<td>.60**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>3. Final course score</td>
<td>79.16</td>
<td>13.94</td>
<td>10.00</td>
<td>97.60</td>
<td>-.71**</td>
<td>-.68**</td>
<td>1.00</td>
</tr>
<tr>
<td>4. Exam score</td>
<td>42.50</td>
<td>9.43</td>
<td>.00</td>
<td>57.6</td>
<td>-.55**</td>
<td>-.47**</td>
<td>.94**</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.

Regression analysis on course achievement

First, absence and late submission were entered as predictors for the regression model of the final course score. According to the multiple regression analysis results (see Table 2), absence ($B = -4.12, p < .001$) and late submission ($B = -5.62, p < .001$) were negatively significant in predicting the final course score. The regression model explained 59.7% of the variance in the final course score ($R^2 = .597, F(2, 567) = 420.12, p < .001$).
Another multiple regression analysis was conducted to test the predictability of absence and late submission on the exam score, and the results are presented in Table 3. Absence ($B = -2.54, p < .001$) and late submission ($B = -2.00, p < .001$) were negatively significant in predicting the exam score. Academic procrastination accounted for 32.5% of the variability in the exam score ($R^2 = .325, F(2, 567) = 136.66, p < .001$).

Furthermore, multiple regression analyses were performed to test the predictability of the determined model on course achievement at four different points of the course. As shown in Table 4, absence ($B = -7.18, p < .001$) and late submission ($B = -4.16, p < .01$) were significant in predicting the final course score in model T1, and the model explained 20.8% of the variance in the final course score ($R^2 = .208, F(2, 567) = 74.36, p < .001$). In model T2, absence ($B = -5.86, p < .001$) and late submission ($B = -12.37, p < .001$) were significant in predicting the final course score, and model T2 explained 39.8% of the variance in the final course score ($R^2 = .398, F(2, 567) = 187.14, p < .001$). The difference in $R^2$ between models T1 and T2 was .19. In model T3, absence ($B = -4.76, p < .001$) and late submission ($B = -7.33, p < .001$) were significant in predicting the final course score, and model T3 explained 54.1% of the variance in the final course score ($R^2 = .541, F(2, 567) = 333.76, p < .001$). From model T2 to model T3, $R^2$ increased by .143. Model T4 explained 59.7% of the variance in the final course score ($R^2 = .597, F(2, 567) = 420.12, p < .001$), and from model T3 to model T4, $R^2$ increased by .056. These results demonstrate that absence and late submission predicted the final course score significantly at different points during the course. Over time, the extent of predictability increased from 20.8% to 39.8% to 54.1% to 59.7%.

Similarly, in model T1, absence ($B = -3.73, p < .001$) and late submission ($B = -4.16, p < .01$) were significant in predicting the exam score, and absence ($B = -3.28, p < .001$) and late submission ($B = -4.80, p < .001$) were significant in predicting the exam score in model T2. In model T3, absence ($B = -2.87, p < .001$) and late submission ($B = -2.58, p < .001$) were significant in predicting the exam score. In the final model, T4, absence ($B = -2.54, p < .001$) and late submission ($B = -2.00, p < .001$) were significant in predicting the exam score. Over time, the magnitude of the model’s predictability increased from 8.5% to 18.4% to 28.7% to 32.5%.

### Table 2. Multiple regression analysis results on final course score ($n = 569$)

<table>
<thead>
<tr>
<th>Variables</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\beta$</th>
<th>$R^2$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence</td>
<td>-4.12</td>
<td>.30</td>
<td>-.46***</td>
<td>.597</td>
<td>420.12***</td>
</tr>
<tr>
<td>Late submission</td>
<td>-5.62</td>
<td>.47</td>
<td>-.40***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***$p < .001$.

### Table 3. Multiple regression analysis results on exam score ($n = 569$)

<table>
<thead>
<tr>
<th>Variables</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\beta$</th>
<th>$R^2$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence</td>
<td>-2.54</td>
<td>.26</td>
<td>-.42***</td>
<td>.325</td>
<td>136.66***</td>
</tr>
<tr>
<td>Late submission</td>
<td>-2.00</td>
<td>.41</td>
<td>-.21***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**$p < .001$.**

### Table 4. Multiple regression analyses results ($n = 569$)

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>Final course score</th>
<th>Exam score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$B$</td>
<td>$SE$</td>
</tr>
<tr>
<td>T1</td>
<td>Absence</td>
<td>-7.18</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>Late submission</td>
<td>-13.55</td>
<td>2.21</td>
</tr>
<tr>
<td>T2</td>
<td>Absence</td>
<td>-5.86</td>
<td>.66</td>
</tr>
<tr>
<td></td>
<td>Late submission</td>
<td>-12.37</td>
<td>1.09</td>
</tr>
<tr>
<td>T3</td>
<td>Absence</td>
<td>-4.76</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td>Late submission</td>
<td>-7.33</td>
<td>.66</td>
</tr>
<tr>
<td>T4</td>
<td>Absence</td>
<td>-4.12</td>
<td>.30</td>
</tr>
<tr>
<td></td>
<td>Late submission</td>
<td>-5.62</td>
<td>.47</td>
</tr>
</tbody>
</table>

**$p < .001$.**
Discussion

The present study was conducted to investigate the effects of academic procrastination on course achievement. To measure academic procrastination in e-learning, delays in weekly scheduled learning and late submission of assignments were identified as academic procrastination behaviors, and the absence score and the frequency of late submission were calculated from log data. The final course score and the exam score were obtained from the class as course achievement. The multiple regression results showed that absence and late submission negatively influenced the final course score and the exam score. These results are in line with those reported in previous research (Akinsola, Tella, & Tella, 2007; Asarta & Schmidt, 2013; Ferrari, Johnson, & McCown, 1995; Michinov et al., 2011; Tuckman, 2005) and empirically proved the negative effects of academic procrastination on course achievement in e-learning. In addition, the predictability of the model at four different time points was significant in predicting course achievement, and the extent of the predictability increased accordingly.

The findings obtained from the present study have several implications. First, the study suggests the importance of regular study in achieving success in e-learning. In this study, the e-learning course was arranged over 15 weeks, and the learning materials were accessible according to the course schedule. Absence in current study indicated that a student either did not access the scheduled learning materials within the predetermined time period or accessed the materials during the designated period but did not fully finish reviewing the materials during the period. The results showed that the failure to study regularly negatively influenced students’ achievement. Although students who were marked as absent might have studied the weekly materials later, the irregular study behavior led to poor achievement. This interpretation could be supported by the previous research finding indicating that students who postponed and crammed the materials at the last minute showed poorer long-term retention or achievement (Asarta & Schmidt, 2013; Elvers, Polizella, & Graetz, 2003; Melton, 1970; Tuckman, 2005). Therefore, instructors need to not only encourage students to study regularly but must also try to detect low self-regulated learners as early as possible and provide substantive guidance to them.

Second, in addition to studying regularly, submitting an assignment on time was significant in achieving success in the course, and this finding confirmed the importance of time management. Britton and Tesser (1991) discussed that time was considered a learning resource and that the capability of managing resources effectively was essential to improving achievement. Furthermore, they stressed the effects of time management practices by showing that time management skills accounted for more variance of college GPA than SAT scores. In the present study, the two academic procrastination behaviors accounted for 59.6% of the variance in the final course score and 32.5% of the variance in the exam score; these results clearly demonstrate the detrimental effects of academic procrastination. Therefore, providing help for students to keep their learning schedule should be sought. Although most LMSs offer several functions that facilitate students’ participation and interaction, Dunlap and Lowenthal (2009) noted that the accessibility and promptness of LMS functions are limited. They recommended using social media, such as Twitter, to enhance participation and prompt communication among students and instructors because they are connected to social media most of the time. Therefore, the supplementary use of social media or other learning tools that facilitate prompt interaction would be helpful in keeping students on track.

Third, the selected log-based procrastination behaviors in this study predicted course achievement considerably well. The results demonstrate the potential of analyzing LMS data to detect academic risks and predict students’ performance during e-learning. As mentioned earlier, however, it is demanding to utilize such a large amount of data effectively and efficiently. The number of students in e-learning courses is often large, and it is not easy for instructors to monitor and identify each student’s progress and level of understanding (Maedadyen & Dawson, 2010). Therefore, new tools such as dashboard interfaces are necessary to summarize students’ progress visually and intuitively. If instructors could easily recognize the performance tendencies of both the class as a whole and of individual students, they would be able to adjust their teaching pace and provide additional materials for elaboration. From this perspective, researchers must strive to explore and validate critical indicators for learning, and institutions and educational system developers must make efforts to implement learning analytics in the field.

Limitations and future research

Although this study provides insight into using LMS data to investigate the relationship between procrastination and performance, some suggestions for future study should be noted. First, because the data in this study were collected
from a university in Korea, more empirical studies on academic procrastination in e-learning in various educational contexts would help to generalize the research findings. Second, absence and late submission were chosen as indicators of academic procrastination behaviors in this study, but it should be noted that absence in this study reflected a failure to study regularly and was scored under specific conditions; thus, careful interpretation is required. Moreover, behaviors such as last-minute cramming or submissions that barely meet the deadline are also deemed procrastinating behavior, but were not considered in this study. Investigating other indicators of academic procrastination would expand the understanding of procrastination.

References


Gamification in Education: A Systematic Mapping Study

Darina Dicheva¹*, Christo Dichev¹, Gennady Agre² and Galia Angelova²

¹Department of Computer Science, Winston Salem State University, Winston Salem, North Carolina, USA // ²Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Sofia, Bulgaria // dichevad@wssu.edu // dichevc@wssu.edu // agre@iinf.bas.bg // galia@lml.bas.bg

*Corresponding author

(Submitted July 21, 2014; Revised November 16, 2014; Accepted November 22, 2014)

ABSTRACT

While gamification is gaining ground in business, marketing, corporate management, and wellness initiatives, its application in education is still an emerging trend. This article presents a study of the published empirical research on the application of gamification to education. The study is limited to papers that discuss explicitly the effects of using game elements in specific educational contexts. It employs a systematic mapping design. Accordingly, a categorical structure for classifying the research results is proposed based on the extracted topics discussed in the reviewed papers. The categories include gamification design principles, game mechanics, context of applying gamification (type of application, educational level, and academic subject), implementation, and evaluation. By mapping the published works to the classification criteria and analyzing them, the study highlights the directions of the currently conducted empirical research on applying gamification to education. It also indicates some major obstacles and needs, such as the need for proper technological support, for controlled studies demonstrating reliable positive or negative results of using specific game elements in particular educational contexts, etc. Although most of the reviewed papers report promising results, more substantial empirical research is needed to determine whether both extrinsic and intrinsic motivation of the learners can be influenced by gamification.

Keywords
Gamification in education, Game design elements, Systematic mapping study, Literature review

Introduction

Traditional schooling is perceived as ineffective and boring by many students. Although teachers continuously seek novel instructional approaches, it is largely agreed that today’s schools face major problems around student motivation and engagement (Lee & Hammer, 2011). The use of educational games as learning tools is a promising approach due to the games’ abilities to teach and the fact that they reinforce not only knowledge but also important skills such as problem-solving, collaboration, and communication. Games have remarkable motivational power; they utilize a number of mechanisms to encourage people to engage with them, often without any reward, just for the joy of playing and the possibility to win. Creating a highly engaging, full-blown instructional game however is difficult, time consuming, and costly (Kapp, 2012a), while typically targeting only a single set of learning objectives as chosen by the game designer. In addition, their effective classroom adoption requires certain technical infrastructure and appropriate pedagogical integration. As opposed to using elaborate games requiring a large amount of design and development efforts, the “gamification” approach suggests using game thinking and game design elements to improve learners’ engagement and motivation.

Gamification, defined by Deterding et al. (2011) as the use of game design elements in non-game contexts, is a fairly new and rapidly growing field. The concept of gamification is different from that of an educational or serious game. While the latter describes the design of full-fledged games for non-entertainment purposes, “gamified” applications merely employ elements of games. The term “gamification” is quite recent: According to (Deterding et al., 2011) its first documented use is in 2008 but it did not see widespread adoption before the second half of 2010. Nevertheless, the concept itself is not new. For example, badges and ranks have been long used in the military, in the early Soviet era, game elements were used by the Soviet Union leaders as a substitute for monetary incentives for performing at work, etc.

In recent years gamification has seen rapid adoption in business, marketing, corporate management, and wellness and ecology initiatives. This is driven by its potential to shape users’ behavior in a desirable direction. Loyalty programs such as the frequent-flyer programs, Foursquare, and Nike+ are often given as examples of successful gamified mass-market products. Stackoverflow.com provides another example in which users’ reputations increase as
they answer questions and receive votes for their answers. Online education sites such as codeacademy.com and khanacademy.org use game elements to better engage users. The more courses and lessons that users complete, the more badges they earn. Sites like eBay and Fitocracy use game elements to keep people engaged and to encourage friendly competition between users.

Gamification is still rising in popularity. According to Gartner’s Hype Cycle (Gartner, 2013), a research methodology that outlines an emerging technology’s viability for commercial success, gamification is at the peak of the Hype Cycle in 2013, with an expectation for reaching the productivity plateau in five to ten years. This position, however, mainly reflects its use in business contexts. The penetration of the gamification trend in educational settings seems to be still climbing up to the top, as indicated by the amount and annual distribution of the reviewed works.

This paper presents the results of a study of the published works on the application of gamification to education, which aims to shed light on the tendencies and emerging practices in this area. There are few literature reviews on gamification (see Xu, 2012; Hamari, Koivisto, & Sarsa, 2014; Nah, Zeng, Telaprolu, Ayyappa, & Eschenbrenner, 2014), with only the last one focusing on education. This study differs from the latter by presenting a thematic analysis instead of narrative summaries that focus on a qualitative review.

**Systematic mapping study design**

The main research questions behind this study were: “What educational contexts has gamification been applied to?” and “What game elements have been used in gamifying educational systems?” We used a systematic mapping design for the study. Systematic mapping studies are similar to systematic reviews, except that they employ broader inclusion criteria and are intended to map out topics rather than synthesize study results. A systematic mapping study provides a categorical structure for classifying the published research reports and results. The study presented here covers the existing work in the field of gamification in education: articles and conference papers published and indexed until June 30, 2014. The recency of the interest in conducting research on this topic is demonstrated by the distribution of the studied papers by year of publication, presented in Figure 1.

![Figure 1. Work distribution by year of publication](image)

**Inclusion, search, and screening**

The inclusion criterion for the papers was to discuss explicitly the use of game elements in educational contexts. Note that motivation is a very central and fundamental topic in education (different from other contexts of application of gamification), and a lot of research has been done on it. Also, techniques such as feedback, ordering learning tasks by their complexity, personalization, etc., are as fundamentally essential for games as they are for education. Therefore, from an educational point of view, it would be unnatural to consider them as “game mechanisms” making their way to education. There is substantial motivation-related research, for example, on pedagogical methods such as inquiry-based learning, psychological research on intrinsic and extrinsic motivation, and self-regulation (see, for example, Deci & Ryan, 1985; Lei, 2010), on motivation for participation in social networks (see, for example, Vassileva, 2012), or technological approaches, such as course sequencing (see, for example, Brusilovsky & Vassileva, 2003), or adaptive learning systems (see, for example, Brusilovsky, 1999), etc. Consequently, papers presenting research on such topics (although related to principles and techniques considered by the traditional computer game theorists as game elements) are not included in this study. We are targeting a more
holistic approach to the use of game design elements in education and consider them from the perspective of gamification: Can their game-like implementation motivate learners and enrich the educational experiences?

We searched seven major scientific databases: ACM Digital Library, IEEE Xplore, ScienceDirect, SCOPUS, Springer Link (books), ERIC, and Google Scholar. After searching the databases (in this order) with keywords “gamification,” “gamify,” and “gameful,” and removing the duplicates, we obtained the following search results: ACM Digital Library (376 papers), IEEE Xplore (100 papers), ScienceDirect (119 papers), SCOPUS (405 papers), Springer Link (86 papers), ERIC (7 papers), and Google Scholar (554 papers). Based on abstracts, we first filtered out all publications that are not related to education or are not published in peer-reviewed conferences or journals (e.g., technical reports and master theses). This was followed by a second round of filtering in which, based on the full text, we removed the publications that are concerned with applying gamification for tasks that are not directly related to learning, such as university orientation for freshmen, library orientation, academic advising, etc., and those related to full-fledged educational games. We also removed early papers that only explain the concept of gamification and suggest very general possible uses in education. Meanwhile, we investigated the references of the found papers and discovered several papers relevant for the review that were not included in the databases. The resulting set contained 34 papers presenting empirical studies to be analyzed and classified (see Appendix I).

Categorization criteria

In order to answer the research questions, we performed a concept-centric review focusing on categories related to the context of use and game elements employed for gamification of education. The review of the papers provided us with information allowing the classification of the current research and work in the field along the following dimensions:

- Game elements
- Context: type of application
- Context: education level
- Context: academic subject
- Implementation
- Reported results from evaluation

With regard to the categorization of the game elements, we first surveyed the existing seminal, conceptual, and literature-review publications on gamification (not included in the 34 papers reporting empirical research). However, we discovered that there is not a commonly agreed classification of game design elements. For example, the popular game element “badges” is considered as a game interface design pattern in (Deterding et al., 2011), a game mechanic in (Zichermann & Cunningham, 2011), a game dynamic in (Iosup & Epema, 2014), a motivational affordance in (Hamari, Koivisto, & Sarsa, 2014), and a game component (a specific instantiation of mechanics or dynamics) in (Werbach & Hunter, 2012). Nevertheless, all authors define the game design elements at several levels of abstraction. For example, Zichermann and Cunningham (2011), following traditional computer game theorists, categorize game elements into mechanics, dynamics, and aesthetics. Mechanics define the way games (as systems) convert specific inputs into specific outputs. Dynamics guide how players and the game mechanics interact during the game. Aesthetics refer to the way the game mechanics and dynamics interact with the game designer’s artistry, to produce cultural and emotional outcomes. Differently, Deterding et al. (2011) categorize game design elements at five levels of abstraction. Ordered from concrete to abstract, these are: interface design patterns; game design patterns or game mechanics; game design principles, heuristics or “lenses”; conceptual models of game design units; and game design methods and design processes.

For the purpose of reviewing the use of game elements in gamified educational contexts, we use a two-level framework. The first level combines the first two levels of Deterding’s classification and, as most of the authors in the field, we refer to it as game mechanics. We further combine Levels 3 and 4 of Deterding’s classification (game design principles and conceptual models) and call them educational gamification design principles. We use the term gamification design principles instead of game design principles to stress the fact that a number of these are not specific to games. In the education domain, some have been used in instructional systems as long as those have existed. These two categories roughly correspond to the first two components of the framework in (Zichermann &
Cunningham, 2011). The last Deterding’s category “game design methods and processes,” as well as Zichermann’s “aesthetics,” are essential for the game elements’ implementation but are not relevant to this mapping study.

To further identify the second level of the classification structure, we collected game mechanics and game design dynamics, patterns, and principles used in the 34 reviewed case studies on using gamification in education. We identified the use of the following game mechanics: points, badges, levels, progress bars, leaderboards, virtual currency, and avatars. Point systems manage the acquisition and spending of points that quantify user performance. Badges are given for special achievements. Based on the received points and badges, users are ranked on leaderboards that reflect their performance in comparison to other users. Levels show the user’s expertise and progress and where the player is in the game. Progress bars provide a percentage-based graphical representation of the players’ progress. Virtual currency is used for purchasing in-game (virtual) goods.

Table 1 below presents the identified educational gamification design principles with, where appropriate, the game mechanics typically used to implement them. For each principle, corresponding references are presented. Some of the listed educational gamification design principles are fundamental and always present in educational systems but may need to be adapted to fit the gamification paradigm. For example, the feedback should be immediate or with shortened cycles (not as in the current educational practices). Others have been used individually and sporadically by some instructors but still need re-thinking in light of gamification, and some are new design elements borrowed from video games.

<table>
<thead>
<tr>
<th>Design principles</th>
<th>Used game mechanics</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goals: specific, clear, moderately difficult, immediate goals</td>
<td></td>
<td>Lee &amp; Hammer, 2011</td>
</tr>
<tr>
<td>Challenges and quests: clear, concrete, actionable learning tasks with increased</td>
<td></td>
<td>Kapp, 2012b</td>
</tr>
<tr>
<td>complexity</td>
<td></td>
<td>Lee &amp; Hammer, 2011</td>
</tr>
<tr>
<td>Customization: personalized experiences, adaptive difficulty; challenges that are</td>
<td>Points, progress bars,</td>
<td>Zichermann &amp; Cunningham, 2011</td>
</tr>
<tr>
<td>perfectly tailored to the player’s skill level, increasing the difficulty as the</td>
<td>levels, virtual goods/currency</td>
<td>Deterding, 2013</td>
</tr>
<tr>
<td>player’s skill expands</td>
<td></td>
<td>Simões, Díaz, &amp; Fernández, 2013</td>
</tr>
<tr>
<td>Progress: visible progression to mastery</td>
<td></td>
<td>Gordon, Brayshaw, &amp; Grey, 2013</td>
</tr>
<tr>
<td>Feedback: immediate feedback or shorten feedback cycles; immediate rewards instead</td>
<td>Points</td>
<td>Zichermann &amp; Cunningham, 2011</td>
</tr>
<tr>
<td>of vague long-term benefits</td>
<td></td>
<td>Lee &amp; Hammer, 2011</td>
</tr>
<tr>
<td>Competition and cooperation/social engagement loops</td>
<td>Badges, leaderboards,</td>
<td>Zichermann &amp; Cunningham, 2011</td>
</tr>
<tr>
<td></td>
<td>levels, avatars</td>
<td>Iosup &amp; Epema, 2014</td>
</tr>
<tr>
<td>Accrual grading</td>
<td>Points</td>
<td>Zichermann &amp; Cunningham, 2011</td>
</tr>
<tr>
<td>Visible status: reputation, social credibility and recognition</td>
<td>Points, badges, leaderboards,</td>
<td>Deterding, 2013</td>
</tr>
<tr>
<td></td>
<td>avatars</td>
<td>Simões, Díaz, &amp; Fernández, 2013</td>
</tr>
<tr>
<td>Access/unlocking content</td>
<td></td>
<td>Iosup &amp; Epema, 2014</td>
</tr>
</tbody>
</table>

78
allowing students to choose their own sub-goals within the larger task

| Freedom to fail: low risk from submission, multiple attempts | Iosup & Epema, 2014 |
| | Deterding, 2013 |
| | Simões, Díaz, & Fernández, 2013 |

| Storytelling | Avatars |
| | Nah et al., 2014 |
| | Kapp, 2012b |
| | Simões, Díaz, & Fernández, 2013 |

| New identities and/or roles | Avatars |
| | Lee & Hammer, 2011 |
| | Simões, Díaz, & Fernández, 2013 |

| Onboarding | Avatars |
| | Zichermann & Cunningham, 2011 |

| Time restriction | Countdown clock |
| | Kapp, 2012b |

Each of the 34 papers presenting empirical studies was evaluated to examine which of these defined categorization criteria were discussed.

**Mapping study results**

This section describes the distribution of published work on each classification criterion. As proposed above, the criterion of game elements is divided into two: gamification design principles and game mechanics.

*Gamification design principles.* Figure 2 shows the number of papers discussing each of the identified educational gamification design principles (see Table 1).

![Figure 2. Work distribution by gamification design principles](image)

As we can see, the most used gamification design principles in educational context are visual status, social engagement, freedom of choice, freedom to fail, and rapid feedback. Papers that discuss the principles of goals and personalization are rare. The likely reason for this was mentioned before: these are fundamental principles for instruction and educational applications being the target of a long-standing pedagogical and educational computing research. So advancements related to them would not be considered a result of gamifying education.

Examples of applying the principle “freedom of choice” include the possibility for students to choose. For example, what type of challenges to complete: writing traditional essays, completing an open-ended group project, completing an open-ended individual project, or contributing to the class blog (Holman, Aguilar, & Fishman, 2013); writing academic papers, creating an instructional YouTube video, or developing an educational game design (De Schutter & Abeele, 2014); and taking tests or completing artistic assignments (Mak, 2013). Other examples include choices of specific challenges to complete (e.g., Barata, Gama, Jorge, & Gonçalves, 2013; Haaranen, Ilantola, Hakulinen, &
Korhonen, 2014), the order and/or speed of completing the challenges (e.g., Berkling & Thomas, 2013; Todor & Pitica, 2013), the choice of selecting skill goals, how the challenges or their types are weighted (e.g., Holman, Aguilar, & Fishman, 2013; Gibbons, 2013), customizing assignment deadlines (Gibbons, 2013), and voting on the extent of the marks deduction for penalties for absences or non-completion of assigned tasks by a team member (Caton & Greenhill, 2013).

The principle “freedom to fail” presumes no penalties on poor task performance and typically includes allowing students to revise and re-submit assignments (e.g., Haaranen, Ihantola, Hakulinen, & Korhonen, 2014; Berkling & Thomas, 2013; de Byl & Hooper, 2013; Hentenryck & Coffrin, 2014) or re-take quizzes (O’Donovan, Gain, & Marais, 2013). Although this principle is perhaps one of the most controversial for applying in a conventional classroom, there are no empirical studies carrying out specifically its controlled evaluation.

Social engagement includes individual and team competitions (e.g., O’Donovan et al., 2013; Li, Grossman, & Fitzmaurice, 2014), taking part in group “guild” learning activities and work on team projects (e.g., Mak, 2013; Caton & Greenhill, 2013; Mitchell, Danino, & May, 2013; Burkey, Anastasio, & Suresh, 2013), cooperation and interaction with other students (e.g., Giannetto et al., 2013; Landers & Callan, 2011), etc.

Only six studies were found to investigate the impact of the use of a single game technique: one of a leaderboard (Hentenryck & Coffrin, 2014) and all the others of badges (Anderson, Huttenlocher, Kleinberg, & Leskovec, 2014; Hakulinen & Auvinen, 2014; Haaranen, Ihantola, Hakulinen, & Korhonen, 2014; Denny, 2013; Abramovich, Schunn, & Higashi, 2013). Only one study was identified to evaluate the effect of different type of game elements (badges) on different type of learners (Abramovich et al., 2013).

Game mechanics. Figure 3 shows the number of papers reporting the use of each of the identified game mechanisms. It confirms that the most popular game mechanisms are points, badges, and leaderboards.

Regarding the use of badges, in some of the case studies their assignment does not affect student grading, but is aimed at triggering competitive motivation (Pirker, Riffhaller-Schiefer, & Gütl, 2014). Badges are given for different achievements, for example, for challenge achievements and participation achievements (Domínguez et al., 2013), for learning, time management, and carefulness (Hakulinen & Auvinen, 2014), for contributing to threads and reading/voting on content (Anderson et al., 2014), or for performance and fun (Bartel & Hagel, 2014). As to levels, (Kapp, 2012b), for example, considers three types of levels: game levels, playing levels, and player levels. Goehle (2013) recommends choosing levels so that initially levels are earned quickly but become increasingly difficult to obtain later on. Examples for using virtual (in-game) currency include spending it on puzzle hints, assignment extensions, quiz do-overs (allowing the buyer another three chances at a quiz) (O’Donovan, Gain, & Marais, 2013), or getting help on certain homework problems, extending a due date with no penalty, using a larger index card for notes on a test (Goehle, 2013), etc.

Type of application. This criterion is about the context of the gamification application, that is, where gamification is applied. The papers were grouped in the following categories: for gamifying courses without online gamification support, for gamifying MOOCS or online courses, for gamifying blended learning courses, for gamifying e-learning
sites, and for developing gamification support platforms. Figure 4 shows the number of papers in each category. As it can be seen, the majority of the reported case studies are on gamification of blended learning courses.

Education level. This criterion is about the targeted educational level. Only two papers consider gamification for the K12 education (Abramovich, Schunn, & Higashi, 2013; Morrison & DiSalvo, 2014), while the remaining articles target higher education and training.

Subject. This criterion is related to the subject domain of the application of gamification. The following categories were identified here: computer science (CS); information technology (IT); game programming, math/science/engineering, and subject-neutral (see Figure 5). Most of the papers report gamifying of computer science or IT courses.

Implementation. A spectrum of implemented support for the instructors introducing gamification approaches in their teaching framework was identified, varying from no automated support at all to the use of standalone gamification platforms. The papers were grouped in the following categories (see Figure 6, where the first and second options are combined):

- No e-learning platform or other software used (Mak, 2013; Caton & Greenhill, 2013; Mitchell, Danino, & May, 2013; Burkey, Anastasio, & Suresh, 2013). For example, Mitchell et al. (2013) report that only teacher efforts and a leaderboard have been used.

- Manual collection of data on student performance and processing it with a computer program. Barata et al. (2013) report collecting data from lectures and labs by faculty on Excel sheets and downloading data logs from Moodle followed by running a Python script to process the data and generate the leaderboard webpage (two to three times a day to track major updates with low response time).
• Software for supporting gamification implemented as a plug-in or extension of a learning management system (LMS) or other online learning environment at the university. Examples include extending Moodle (Pirker, Riffnaller-Schiefer, & Gütl, 2014), A+ (Haaranen, Ihattola, Hakulinen, & Korhonen, 2014), Vula Sakai environment (O’Donovan, Gain, & Marais, 2013), Blackboard 9 (Domínguez, et al., 2013), QizBox (Giannetto, Chao, & Fontana, 2013), and the online homework platform WeBWorK (https://github.com/openwebwork, Goehle, 2013).

• Third-party software used to support some aspect of gamification. Examples include using Moodle (Thomas & Berkling, 2013); the Diagnosys tool for assessment of basic mathematical skills, which includes lives, time limits, and adaptive difficulty (Gordon, Brayshaw, & Grey, 2013); the collaborative learning environment Curatr (http://www.curatr3.com/), which uses gamification principles (Betts, Bal, & Betts, 2013); BadgeVille (http://badgeville.com/); WordPress (http://wordpress.org/), with its Achievements plug-in (WordPress Achievements, 2014) (Werbach & Johnson, 2012); and the free hosted online platform CourseSites (https://www.coursesites.com/webapps/Bb-sites-course-creation-BBLEARN/pages/index.html), which provides an integration of Mozilla Open Badges (Thomas & Berkling, 2013). Thomas & Berkling state that the multi-platform approach of using Moodle, along with a combination of online quiz-taking tools and another platform for gamification aspects, proved to be very difficult for the students. These authors also provide a comparison of using different software platforms to support course gamification. After comparing Moodle, Sakai (http://www.sakaiproject.org), and CourseSites, the authors chose and recommended CourseSites.

• Software for supporting gamification implemented as standalone applications. The authors of the corresponding papers report the development of tools to support some aspects of gamification in educational contexts (Hakulinen & Auvinen, 2014, Berkling & Thomas, 2013, Todor & Pitica, 2013, and Landers & Callan, 2011).

Reported Results. Figure 7 shows the paper distribution by the type of the results from the reported case studies’ evaluation, grouped in the following categories: positive, positive first impression but not properly evaluated, mixed or suggestive, negative, and not evaluated yet or results not accessible.
The majority of the papers report encouraging results from the experiments, including significantly higher engagement of students in forums, projects, and other learning activities (e.g., Anderson, Huttenlocher, Kleinberg, & Leskovec, 2014; Caton & Greenhill, 2013; Akpolat & Slany, 2014); increased attendance, participation, and material downloads (Barata, Gama, Jorge, & Gonçalves, 2013); positive effect on the quantity of students’ contributions/answers without a corresponding reduction in their quality (Denny, 2013); increased percentage of passing students and participation in voluntary activities and challenging assignments (Losup & Eperna, 2014); and minimizing the gap between the lowest and the top graders (Barata, Gama, Jorge, & Gonçalves, 2013). Hakulinen & Auvinen (2014) conclude that achievement badges can be used to affect the behavior of students even when the badges have no impact on the grading. The papers of this group also report that students considered the gamified instances to be more motivating, interesting, and easier to learn as compared to other courses (Mak, 2013; Barata, Gama, Jorge, & Gonçalves, 2013; de Byl & Hooper, 2013; Mitchell, Danino, & May, 2013; Leong & Yanjie, 2011).

Most of the mixed/suggestive evaluations point missed critical motivational elements in the application of gamification (Morrison & DiSalvo, 2014), sensitivity of the outcomes to small changes in the implementation, a requirement for an ongoing monetary and time investment (O’Donovan, Gain, & Marais, 2013), and the need of strong teaching staff able to design effective assignments, grade students’ work relatively quickly, and interact with students closely (Leong & Yanjie, 2011). Abramovich, Schunn, and Higashi (2013) advise that educational badge designers must consider the ability and motivations of learners when choosing what badges to include in their curricula. Berkling and Thomas (2013), however, report a somewhat negative experience: “Students did not seem to be ready for autonomy, mastery was not perceived to be relevant, and the purpose of starting project work as well as good preparation for the exam seemed unattainable to the students.” The authors suggest that gamification elements be used without being named explicitly, and that the change from the traditional style classroom to the new learning environment be introduced very slowly. In the same vein, Michigan University’s Prof. Lampe is concerned that course gamification could be “whitewashed” by merely masking the terms, for example, by calling assignments as quests and scores as experience points, without contributing to the student’s learning goals (Mak, 2013).

Conclusion

The goal of this study was to review the directions and tendencies of the conducted research on the application of gamification to education and, more specifically, to shed light on the context of application and game elements used. Concerning the limitations of the review, as we stated, the selection criteria included only papers that clearly studied the effects of implementation of game elements in educational contexts. Similarly to (Hamari, Koivisto, & Sarsa, 2014), we excluded research on topics conceptually or theoretically close to gamification (such as intrinsic motivations) or with similar measured outcomes, and papers discussing similar topics but with different terms. Thus, this review provides a fresh, in-depth look on the empirical research being done particularly on the topic of gamification in education.

The study revealed that there are many publications on the use of gamification in education but the majority describe only some game mechanisms and dynamics and re-iterate their possible use in educational context, while true empirical research on the effectiveness of incorporating game elements in learning environments is still scarce. In addition, most of the empirical studies do not include a proper evaluation, which makes it difficult to conduct a meta-analysis of the results of these studies and speculate on general reasons for their successes or negative results. While the mapping study identifies some emerging tendencies in utilizing certain configurations of game mechanics and gamification design principles, their effect in learning context remains to be demonstrated in practice.

Although proper evaluation is mostly missing, the majority of the authors of the reviewed papers share the opinion that gamification has the potential to improve learning if it is well designed and used correctly. Therefore, more substantial empirical research is needed to investigate, in particular, the motivating effects of using single game elements in specific educational contexts and for particular types of learners. This would inform instructors who are interested in gamifying their courses and help them in deciding what game elements to use in their specific context.

The study also shows that the early adopters of gamification are mostly computer science/IT educators. Our speculative explanation is that utilizing gamification assumes a certain type of environment that supports incorporating and visualizing the selected game mechanisms and dynamics. We believe that the effective classroom adoption of gamification implies both certain technological infrastructure coupled with an appropriate instructional
framework. Today’s course management systems, however, still offer restricted support for gamifying courses. Since the general population of instructors lacks the necessary skills and time for creating, adapting, and/or maintaining an appropriate supportive technological infrastructure, the early application of gamification to learning emerged mainly in CS/IT disciplines. The lack of proper technological support is one of the major obstacles for applying game elements to education. Thus, the development of software tools that can efficiently support gamification in various educational contexts would contribute to a larger-scale adoption as well as to research on the feasibility and efficacy of the gamification of education.

Last but not least, finding and sharing of new ways of applying gamification to learning contexts that are not limited to extrinsic rewards like achievements and badges and that are more meaningful to the students is very important for increasing the application of this emerging technology in education. While the concept of gamification may look simple, the analyzed work demonstrates that gamifying learning effectively is not.

Acknowledgments

This work was partially supported by the project “AComIn: Advanced Computing for Innovation” grant 316087, funded by the European Commission in FP7 Capacity (2012–2016).

References


Berkling, K., & Thomas, C. (2013). Gamification of a software engineering course and a detailed analysis of the factors that led to its failure. In M. E. Auer & D. Guralnick (Eds.), Proceedings of International Conference on Interactive Collaborative Learning (pp. 525–530). doi:10.1109/ICL.2013.6644642


## Appendix 1

### Overview of the studied publications on applying gamification to education

<table>
<thead>
<tr>
<th>Paper</th>
<th>Type/Target of application</th>
<th>Game mechanics</th>
<th>Gamification design principles</th>
<th>Evaluation</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abramovich, Schunn, &amp; Higashi, 2013</td>
<td>Blended learning course</td>
<td>Badges</td>
<td>Status</td>
<td>Mixed/proposing</td>
<td>Added to the CS2N intelligent tutoring system</td>
</tr>
<tr>
<td>Akpolat &amp; Slany, 2014</td>
<td>Traditional course</td>
<td>Badges, leaderboard</td>
<td>Status, social engagement</td>
<td>Positive</td>
<td>Course without online support</td>
</tr>
<tr>
<td>Anderson, Huttenlocher, Kleinberg, &amp; Leskovec, 2014</td>
<td>MOOC</td>
<td>Badges</td>
<td>Social engagement</td>
<td>Positive</td>
<td>Added to a discussion forum in Coursera</td>
</tr>
<tr>
<td>Barata, Gama, Jorge, &amp; Gonçalves, 2013</td>
<td>Blended learning course</td>
<td>Points, badges, levels, leaderboard</td>
<td>Status, choice, onboarding, social engagement</td>
<td>Mixed/proposing</td>
<td>Data collected manually and then processed with a program</td>
</tr>
<tr>
<td>Bartel &amp; Hagel, 2014</td>
<td>Blended learning course</td>
<td>Points, badges, levels, leaderboard</td>
<td>Feedback, status</td>
<td>Not evaluated</td>
<td>Gamification application developed</td>
</tr>
<tr>
<td>Berkling &amp; Thomas, 2013</td>
<td>Blended learning course</td>
<td>Points, levels</td>
<td>Goals, status, choice, freedom to fail, social engagement</td>
<td>Negative</td>
<td>Gamification application developed</td>
</tr>
<tr>
<td>Betts, Bal, &amp; Betts, 2013</td>
<td>Online course</td>
<td>Points, levels</td>
<td>Unlocking content, social engagement</td>
<td>Positive</td>
<td>Using Curatr</td>
</tr>
<tr>
<td>Burkey, Anastasio, &amp; Suresh, 2013</td>
<td>Traditional course</td>
<td>Points, levels, leaderboard</td>
<td>Feedback, status, storyline, social engagement</td>
<td>Positive</td>
<td>Course without online support</td>
</tr>
<tr>
<td>Caton &amp; Greenhill, 2013</td>
<td>Traditional course</td>
<td>Badges, leaderboard</td>
<td>Status, choice, social engagement</td>
<td>Positive</td>
<td>Course without online support</td>
</tr>
<tr>
<td>de Byl &amp; Hooper, 2013</td>
<td>Traditional course</td>
<td>Points, leaderboard</td>
<td>Goals, feedback, status, choice, storyline, freedom to fail</td>
<td>Positive</td>
<td>Course without online support</td>
</tr>
<tr>
<td>de-Marcos, Dominguez, Saenz-de-Navarrete, &amp; Pagés, June 2014</td>
<td>Blended learning course</td>
<td>Badges, levels, leaderboard</td>
<td>Status</td>
<td>Mixed/proposing</td>
<td>Plug-in for Blackboard</td>
</tr>
<tr>
<td>De Schutter &amp; Abeele, 2014</td>
<td>Blended learning course</td>
<td>Points, levels, leaderboard, avatars</td>
<td>Goals, feedback, status, choice, storyline, freedom to fail</td>
<td>Mixed/proposing</td>
<td>Gamification application developed</td>
</tr>
<tr>
<td>Denny, 2013</td>
<td>Online course</td>
<td>Points, badges</td>
<td>Social engagement</td>
<td>Positive</td>
<td>Added to PeerWise</td>
</tr>
<tr>
<td>Domínguez, et al., 2013</td>
<td>Blended learning course</td>
<td>Badges</td>
<td>Social engagement</td>
<td>Positive</td>
<td>Added to QuizBox</td>
</tr>
<tr>
<td>Giannetto, Chao, &amp; Fontana, 2013</td>
<td>Gamification platform</td>
<td>Points, badges, levels</td>
<td>Choice, freedom to fail, social engagement</td>
<td>Not evaluated</td>
<td>Course without online support</td>
</tr>
<tr>
<td>Gibbons, 2013</td>
<td>Blended learning course</td>
<td>Points, badges, levels, virtual currency</td>
<td>Feedback, visual progress</td>
<td>Positive</td>
<td>Added to WeBWorK</td>
</tr>
<tr>
<td>Goehle, 2013</td>
<td>Blended learning course</td>
<td>Leaderboard</td>
<td>Goals, adaptation, feedback, status, freedom to fail, time restriction</td>
<td>Positive</td>
<td>Using Diagnosys</td>
</tr>
<tr>
<td>Authors</td>
<td>Format</td>
<td>Elements</td>
<td>Status</td>
<td>Evaluation</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>-------------------------</td>
<td>----------------------------</td>
<td>-------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Haaranen, Ihantola, Hakulinen, &amp; Korhonen, 2014</td>
<td>Blended learning course</td>
<td>Badges, Goals, choice, freedom to fail</td>
<td>Mixed/proposing</td>
<td>Added to A+</td>
<td></td>
</tr>
<tr>
<td>Hakulinen &amp; Auvinen, 2014</td>
<td>Blended learning course</td>
<td>Points, badges, leaderboard</td>
<td>Status, freedom to fail, social engagement</td>
<td>Positive Added to TRAKLA2 learning environment. Added to a MOOC</td>
<td></td>
</tr>
<tr>
<td>Hentenyck &amp; Coffrin, 2014</td>
<td>MOOC</td>
<td>Leaderboard</td>
<td>Status, choice, freedom to fail</td>
<td>Positive Gamification application developed</td>
<td></td>
</tr>
<tr>
<td>Holman, Aguilar, &amp; Fishman, 2013</td>
<td>Gamification platform</td>
<td>Points, badges, levels, progress bar, leaderboard</td>
<td>Status, unlocking content, choice, freedom to fail, onboarding, social engagement</td>
<td>Positive Course without online support</td>
<td></td>
</tr>
<tr>
<td>Iosup &amp; Epema, 2014</td>
<td>Traditional course</td>
<td>Points, badges, levels, leaderboard</td>
<td>Status, choice, freedom to fail, social engagement</td>
<td>Not evaluated Gamification application developed</td>
<td></td>
</tr>
<tr>
<td>Hakulinen &amp; Auvinen, 2014</td>
<td>Blended learning course</td>
<td>Levels</td>
<td>Goals, feedback, status, choice, freedom to fail, social engagement</td>
<td>Positive Gamification application developed</td>
<td></td>
</tr>
<tr>
<td>Leong &amp; Yanjie, 2011</td>
<td>Blended learning course</td>
<td>Points, badges, levels, leaderboard</td>
<td>Goals, feedback, status, choice, freedom to fail, social engagement</td>
<td>Positive Gamification application developed</td>
<td></td>
</tr>
<tr>
<td>Li, Grossman, &amp; Fitzmaurice, 2014</td>
<td>E-learning site</td>
<td>Levels, leaderboard</td>
<td>Status, social engagement Feedback, choice, new identities, social engagement</td>
<td>Positive Gamification application developed Course without online support</td>
<td></td>
</tr>
<tr>
<td>Mak, 2013</td>
<td>Traditional course</td>
<td>Points, leaderboard</td>
<td>Status, choice, new identities, social engagement</td>
<td>Positive Course without online support</td>
<td></td>
</tr>
<tr>
<td>Mitchell, Danino, &amp; May, 2013</td>
<td>Traditional course</td>
<td>Points, leaderboard</td>
<td>Status, choice, new identities, social engagement</td>
<td>Positive Gamification application developed</td>
<td></td>
</tr>
<tr>
<td>Morrison &amp; DiSalvo, 2014</td>
<td>E-learning site</td>
<td>Points, badges, levels, progress bar</td>
<td>Status, social engagement Feedback, choice, freedom to fail</td>
<td>Positive Course without online support</td>
<td></td>
</tr>
<tr>
<td>O’Donovan, Gain, &amp; Marais, 2013</td>
<td>Blended learning course</td>
<td>Points, badges, levels, progress bar, leaderboard, virtual currency</td>
<td>Feedback, status, freedom to fail, time restriction, storyline</td>
<td>Positive Added to Sakai</td>
<td></td>
</tr>
<tr>
<td>Pirker, Riffnaller-Schiefer, &amp; Gütl, 2014</td>
<td>Blended learning course</td>
<td>Badges, leaderboard</td>
<td>Feedback, status, freedom to fail</td>
<td>Positive first impression Added to Moodle</td>
<td></td>
</tr>
<tr>
<td>Thomas &amp; Berkling, 2013</td>
<td>Blended learning course</td>
<td>Points, levels, leaderboard</td>
<td>Status, social engagement Feedback, status, choice, freedom to fail, new identities, social engagement</td>
<td>Not evaluated Gamification application developed</td>
<td></td>
</tr>
<tr>
<td>Todor &amp; Pitica, 2013</td>
<td>Online platform</td>
<td>Points, badges, levels, leaderboard</td>
<td>Status</td>
<td>Positive first impression Using CourseSites</td>
<td></td>
</tr>
<tr>
<td>Werbach &amp; Johnson, 2012</td>
<td>Blended learning course</td>
<td>Points, badges, leaderboard</td>
<td>Status</td>
<td>Not evaluated Using BadgeVille/WordPress with Achievements plug-in</td>
<td></td>
</tr>
</tbody>
</table>
Perceived Effectiveness of Using the Life-Like Multimedia Materials Tool

Hung-Hsu Tsai1*, Yen-Shou Lai2, Shih-Che Lo3 and Pao-Ta Yu4

1Department of Information Management, National Formosa University, Huwei, Yulin, Taiwan // 2Shin Kong Elementary School, Guikeng, Yulin, Taiwan // 3Department of Industrial Management, National Taiwan University of Science and Technology, Taiwan // 4Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan // thh@nfu.edu.tw // laiysmail@gmail.com // selo@mail.ntust.edu.tw // csipty@cs.ccu.edu.tw

*Corresponding author

(Submitted February 25, 2014; Revised September 1, 2014; Accepted November 3, 2014)

ABSTRACT

The paper investigates the impact of learning effects on the study of life-like multimedia materials (LMMs) which are produced by the LMM tool. Teachers can readily utilize the tool to make the LMMs for their instructions in traditional classrooms. When students study the LMMs out of class, they can realistically recall teacher’s teaching situations and/or the instructional procedures in traditional classrooms. The materials powerfully possess high media richness because of their crucial features such as displaying teacher’s face, hearing teacher’s sound, and presenting teaching materials in the largest region of the screen. A quasi-experiment was also involved to examine the impact of learning effects while using the LMM tool in class and studying the LMMs out of class. In this study, an exploratory method was conducted with 87 students. They were divided into the experimental and the control groups. The results reflect that the experimental group achieved a significant increase in the perceptional effectiveness of the multimedia-based cognitive process with the LMMs in comparison to the control group. Finally, the paper offers discussions for the impact on learning effects when students studied the LMMs out of class.

Keywords

Classroom learning, Multimedia learning, Media richness theory, Teaching strategy, Learning strategy

Introduction

A popular issue for the contemporary education in Taiwan is to integrate information technologies into classroom learning so as to improve students’ learning effects. Consequently, a key point of the issue is that teachers can readily make multimedia materials for classroom learning. As a result, students can easily access multimedia materials and then review them out of class. The way helps students quickly gain abundant multimedia material resources teachers made so as to promote their learning achievements (Zhan, Xu, & Ye, 2011).

Nowadays, several presentation software tools can produce multimedia materials which often include slides, texts, videos, graphics, sounds, etc. Currently, Microsoft PowerPoint is one of the most popular presentation software, which is widely adopted in the traditional classroom for instructions (Susskind, 2008). While using PowerPoint lectures, teachers spend less time in writing or maintaining lecture contents in contrast to the way lecturing with handwriting on a blackboard or whiteboard. Accordingly, teachers can readily highlight key points of teaching materials. However, many articles indicated no significant increase in students’ learning performance while using PowerPoint lectures (Apperson, Laws, & Scepansky, 2008; Susskind, 2008). The reason is that each slide generally comprises 6 lines for texts and/or a few pictures (Apperson, Laws, & Scepansky, 2008). Lai, Tsai, and Yu (2011) pointed out that the PowerPoint presentations do not entirely support the effect of simultaneously containing both verbal and imagery representations during instruction. Accordingly, it is necessary to create a new kind of learning multimedia materials which at least consist of teacher’s handwriting, action on PPT slides, and voice. As a result, students can repeatedly review the kind of multimedia materials out of class so as to get more understandings for teaching materials they still do not figure out in class. The reason is that students can gradually recall teachers’ explanations for their misunderstanding contents via watching the multimedia materials repeatedly.

With recent advances in multimedia recording technologies such as digital video (DV) camera and screen capture software, they have been widely applied in making multimedia materials for classroom learning. Teachers can easily exploit the DV camera in class to create video materials recording their lecturing with a blackboard, whiteboard, or screen. However, to display the kind of materials mentioned above is not clear sufficiently when students study them out of class (Lai, Yu, & Tsai, 2011; Simpson, 2006). A main reason is that the screen layout of the kind of materials...
is not devised for the deployment of the display with respect to several resources such as teacher’s face, the presentation material, drawing attention to key points, etc. This leads to that students cannot easily comprehend teacher’s presentation for their teaching materials. A solution to overcome the above shortcomings adopts screen capturing software which can record teachers’ all actions on the computer desktop and their voices (Lai, Yu, & Tsai, 2011). Nowadays, several popular products are used for screen capturing such as PowerCam, Articulate Presenter, and Camtasia Studio (Lee & Chao, 2007). However, these capturing software cannot simultaneously record teachers’ action on the screen, teacher’s face, and drawing attention to key points for lecture materials. Another kind of the capturing software is applied for video conferencing such as X-learn, Co-life, and JoinNet. They are developed to avoid the disadvantage of exploiting the screen-capturing products mentioned above (Chao & Lin, 2009). The kind of software can capture the display contents on screen and, meanwhile, record the teacher’s voice and face. A critical limitation of the kind of software is to produce low-resolution video due to a need of performing high compression ratio for video via installing software in local computers. Another shortcoming of the products is that they require a high bandwidth network environment. In addition, it requires company-dependent video player software to play the kind of multimedia materials. Unfortunately, these players are not only expensive but also platform-dependent due to special compression protocols and transmission techniques being employed for the video.

Consequently, the paper presents an LMM tool which helps teachers to easily produce the high-quality multimedia materials (the LMMs) for their lecturing in traditional classrooms. The LMMs can help learners effectively and repeatedly to review teacher’s instructional processes. While studying teaching materials via watching a simple display, for example, PowerPoint-made materials, this kind of learning situations just provides what is happening. In contrast to studying PowerPoint-made materials mentioned above, learners can gain more benefits via studying the LMMs. The learning situation of reviewing the LMMs not only offers what is happening but also explains why the content is happening. The LMMs are also helpful for students who can clearly recall the teaching actions in the traditional classroom, including all teaching actions on the computer’s desktop while lecturing. Moreover, the materials the LMM tool produces are helpful to students who can review teacher’s explanations for teaching materials by simultaneously offering teacher’s voices, handwritings, and annotations. Additionally, the screen layout of the LMMs can be flexibly designed to attract students’ attention. An example of the LMMs consists of three main data resources: teacher’s face, computer desktop, and marquee area that can be used to display notes to remind students. An important feature of the tool is that the region exhibiting computer desktop is the largest among four regions comprising a screen. Another feature of the tool is to offer a scrolling marquee and a displaying note for texts which can be used in the design of the screen layout. Teachers can exploit the feature to provide students with explicit or implicit annotations while lecturing so as to promote students’ memorization, thinking, and clarification (Hwang, Wang, & Sharples, 2007). Students can effectively recall teaching activities via repeatedly reviewing the LMMs out of class. Furthermore, the LMM tool offers the on-line broadcasting function. Therefore, students being outside classroom can use the function to synchronously receive teacher’s lecturing in classroom in a way of distance learning.

The remainder of this paper is organized as follows. Section 2 briefly reviews related multimedia learning and media richness theory. Section 3 describes the LMM tool. Section 4 shows the experimental results. Discussions and conclusions are drawn in Section 5.

**Literature review**

**Multimedia-based cognitive process**

With the rapid growth of computer technology, it is an explosion in the availability of presenting materials with different formats including on-screen texts, words, pictures, animations, audio and video (Zhang, Zhou, Briggs, & Nunamaker, 2006). Note that, in general, these formats of materials mentioned above can be divided into visual and verbal representations. The learning via studying the materials with several formats stated above is called multimedia learning (Mayer, 2001). Multimedia learning theory works an integration of cognitive load theory, dual-coding theory, and working memory model. Cognitive process refers to holding a mental representation in working memory over a time interval (Feldon, 2007). During multimedia learning, multimedia-based cognitive process lets the learner hold verbal and visual mental representations while reading multimedia materials (Reed, 2006).
The instructional multimedia material is an input of the cognitive process while learning (Astleitner & Wiesner, 2004). It is processed through triggering attention and working memory. More specifically, learners’ attention can be contracted on the inputs (Wiebe & Annetta, 2008). Additionally, the retention of inputs can be emphasized in learners’ working memory if learners rehearse on the inputs. Moreover, based on multimedia learning theory, visual and auditory information can be simultaneously processed in learners’ working memory. This way can increase learners’ attention and motivation (Astleitner & Wiesner, 2004). Therefore, presenting the LMMs in the traditional classroom became an emergently important issue. For example, the study of Yu et al. showed that using multimedia devices can effectively improve instructor’s teaching ability and student’s learning performance in the classroom learning environment (Yu, Lai, Tsai, & Chang, 2010).

Media richness theory

Media richness is regarded as the ability to facilitate the shared meaning and understanding within a time interval (Rau, Gao, & Wu, 2008). Media richness theory means that richer communication media in terms of representations for uncertain issues are generally more effective and adequate than insufficient or lean media. The information which is represented with high richness media representations can be effectively communicated and conveyed when the information is highly uncertain and equivocal (Sun & Cheng, 2007).

Media richness supports communication activities in education. The richness of each media can be calculated using four criteria (Daft, Lengel, & Trevino, 1987).

- **Capacity for immediate feedback:** The speed and quality about the medium facilitate convergence on a common interpretation.
- **Capacity to transmit multiple cues:** An array of cues, including physical presence, voice inflections, body gestures, words, numbers, and graphic symbols, facilitates conveyance of interpretation and information.
- **Language variety:** It indicates the level of concept convection such as numbers and formulas which provide greater precisions. In contrast, natural language conveys a broader set of concepts and ideas.
- **Capacity of the medium to have a personal focus:** This refers either to the conveyance of emotions and feelings, or to the ability of the medium to be tailored to the specific needs and perspectives of the receiver.

The LMM tool

The structure of the LMM tool

Figure 1 exhibits the structure of the LMM tool, which consists of four main components: Signal Composition Model (SCM), Screen Deployment Model (SDM), Video Compression Model (VCM), and Live Broadcasting Model (LBM). First, the SCM efficiently combines audio and video devices into a single signal. Second, the SDM offers a user-friendly interface. Teachers can employ the SDM to create a screen layout. According to the screen layout, teachers can devise the deployment of displaying their teaching contents and the supplementary materials. Figures 2(a) and 2(b) exhibit the design of a screen layout and an example using the screen layout, respectively. Third, the VCM plays a role to compress the audio and video signals to form multimedia data and then to produce the corresponding streaming version of the multimedia data. Finally, the LBM serves as an Internet protocol television (IPTV), which provides students with multimedia materials through Internet.

Table 1 presents a comparison among the existing recording tools considered in this paper. The merits of using the LMM tool for learning can be mentioned as follows. First, users can flexibly design a new screen layout as shown in Figure 2. Second, in Figure 2(a), two areas, marquee and supplement information, of the screen can offer explicit and implicit annotations, respectively. Texts belong to the kind of explicit annotations. In contrast, the sort of implicit annotations includes graphics, images, highlights, etc. (Hwang, Wang, & Sharples, 2007). The purpose of offering explicit and implicit annotations, as shown in Figure 2, is to promote student’s memorization, thinking, and clarification (Ovsiannikov, Arbib, & Mcneill, 1999). Third, the LMMs with high resolution can be clearly displayed for students. Finally, the compression format of the LMMs is open-form, which can be read by several popular media players.
The learning environment using the LMM tool

Figure 3 displays a learning environment which is composed of five components. The first component, the LMM tool, marked by ① is used to produce the LMMs. Second, a PC or notebook (NB) marked by ② is exploited to present teacher’s instructional materials. Third, a DV camera is marked by ③, which is employed to capture teacher’s face. Subsequently, an additional DV camera marked by ④ can be utilized to capture students’ learning.
activities. Finally, a live-video broadcasting device marked by \( \mathbb{C} \) is an optional component while building a classroom-learning environment. The component is required if it is needed to display teachers’ lecturing in two or more classrooms at the same time. When a teacher lectures in a classroom, it can almost synchronously deliver the corresponding LMMs of his/her lecturing to other classrooms.

**Figure 3.** The LMM tool learning environment can be exploited in a large classroom while lecturing.

**Experiment**

**Participants**

Students in a university of science and technology in the middle of Taiwan contributed to the experiment. Eighty-seven sophomore and junior students (48 males and 39 females) with a mean age of 20.1 years old in the two classes participated in the experiment. Students in the two classes were randomly assigned to the experimental group and the control group. The experimental group (44 students) was lectured using the LMM tool (as shown in Figure 4) and the control group (43 students) was lectured using PowerPoint presentation.

**Figure 4.** The LMM tool
Learning procedure

The learning procedure consists of three phases. First, before lecturing, the teacher prepares the teaching contents presented with PowerPoint. Then, the teacher can add the hint items and annotations in the two areas, marquee and supplement information, as shown in Figure 2(a), in order to clarify the contents in PowerPoint slides the teacher currently explains. The way of using annotations can powerfully convey more meanings and explanations for the contents in PowerPoint slides than without using annotations. Second, during lecturing, the experimental group receives teaching contents using the LMM tool (as shown in Figure 5). Meanwhile, the teacher can utilize the LMM tool to record teacher’s lecturing to form a multimedia material which consists of instructional actions such as teacher’s voices, commentaries, cues, handwritings, narrations, and classroom situations. Accordingly, the multimedia material is so-called the instructional life-like multimedia material. The LMM tool can conveniently offer teachers to produce the LMMs while teaching in a traditional classroom. Subsequently, the LMMs can be uploaded to the learning management system (LMS). The students in the experiment can repeatedly study the LMMs on the LMS out of class. Consequently, the LMM tool also provides a way of asynchronous learning. Additionally, when students cannot show up in the traditional classroom, they can receive teacher’s lecturing through the IPTV of the LMM tool to perform a fashion of synchronous learning. Third, after lecturing, students in the experimental group just get the LMMs for reviewing, and vice versa for students in the control group merely have PowerPoint slides.

![Figure 5. The LMM tool applied in the traditional classroom](image)

All students received a six-week instruction from the same teacher based on the same teaching contents (PowerPoint slides). That is, the six-week learning activity was conducted in the experiment. Table 2 presents a comparison for two groups conducted in the experiment for the six-week learning activity. Figure 6 illustrates the experimental design of the paper.

<table>
<thead>
<tr>
<th>Table 2. A comparison for two groups conducted in the experiment for a six-week learning activity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before lecturing</strong></td>
</tr>
<tr>
<td>Devise screen layout</td>
</tr>
<tr>
<td>Display teaching contents in classroom</td>
</tr>
<tr>
<td>Display multi-sources in a screen</td>
</tr>
<tr>
<td><strong>During lecture</strong></td>
</tr>
<tr>
<td>Recall teachers’ Q&amp;A for students</td>
</tr>
<tr>
<td>Review teaching contents</td>
</tr>
</tbody>
</table>
Sophomore and junior students (n = 87)

Experimental group (n = 44)  Control group (n = 43)

Pre-questionnaire for learning motivation, learning satisfaction, and learning flexibility

50 min

The LMM tool displays PowerPoint slices and produces the LMMs. Review the LMMs during out-of-class.

The PowerPoint software displays PowerPoint slides. Review PowerPoint slides during out-of-class.

6 weeks
150 min per week

Post-questionnaire for learning motivation, learning satisfaction, and learning flexibility

50 min

**Figure 6. Experiment design**

**Instruments**

The study builds a questionnaire related to the perception effectiveness of the multimedia-based cognitive process first using the LMM tool in lecturing in the traditional classroom then reviewing the LMMs. The questionnaire is written in Chinese. The students voluntarily answer the questionnaire. The questionnaire has 11 items, including three sub-questions: learning motivation, learning satisfaction, and learning flexibility. The items of the learning motivation were reedited according to Pintrich and DeGroot (1990) such as “the materials can help me to understand the difficult learning contents.” The items of the learning satisfaction were reedited based on Alavi and Leidner (2001) such as “I can study the instructional materials, clearly hear and see the solving-problem procedures.” The items of the learning flexibility were reedited according to Arbaugh (2000) such as “I can repeatedly study the material at any time and in anywhere out of class.”

The questionnaire adopts five-point Likert scale. Each item has five options, from 1 “strongly disagree” to five “strongly agree.” The higher score means the higher perception of effectiveness for the case of exploiting the LMM tool. The validity of the items is gained by two experts who major in information management and psychology. The Cronbach α is calculated for the internal consistency reliability. The total scale α coefficient was 0.83. It means the reliability of the scale is acceptable.

**Results**

The analysis of independent t-test was conducted to assess the effects of the experimental group received the LMMs and the control group received PowerPoint materials on the 11 survey items. Therefore, the “LMMs” in all items in Table 3 is replaced with “PowerPoint materials” for the control group. Table 3 exhibits the significant results which
indicate that students in the experimental group was effective for the teacher who uses the LMM tool in lecturing, \( t(85) = 2.509, p < .05 \). Exploiting the LMM tool in lecturing may help students pay attention in class, \( t(85) = 2.646, p < .01 \). Students in the experimental group can easily recall the learning activities out of class, \( t(85) = 2.804, p < .001 \). Students in the experimental group can clearly hear and see the problem-solving process, \( t(85) = 2.851, p < .001 \). They also felt that the annotations are helpful to promote their memorization, thinking, and clarification, \( t(85) = 1.965, p < .05 \). The students in the experimental group deemed that reviewing the LMMs can reduce their learning anxiety, \( t(85) = 2.641, p < .01 \). More specifically, low-experience students may not understand the meaning of the materials teacher taught in class. This implies that the kind of students may be anxious in learning the materials teacher taught in class. Subsequently, they can repeatedly watch the LMMs for the problem-solving procedures out of class. Therefore, they can figure out teacher’s explanations for solving problems. The experimental group gave 3.37 points to the 11 survey items on average, which shows that students in the experimental group have higher perceptual effectiveness of the multimedia-based cognitive process with the LMMs.

Students can study the LMMs in a way of asynchronous learning. That is, they can watch the LMMs to recall teacher’s instructional activities out of class so as to promote their understanding for the contents of the LMMs. Accordingly, these experimental results reflected that the kind of LMMs made by the LMM tool was helpful for students in the experimental group. In contrast, students in the control group felt that PowerPoint presentation usually displayed static contents such as simple words, pictures, or some unrelated sounds provided in PowerPoint. They usually cannot deeply think what implicit meaning is for these learning contents.

Table 3. Means (M), standard deviations (SD) and t test of comparison results

<table>
<thead>
<tr>
<th>Item (5, strongly agree; 1, strongly disagree)</th>
<th>EG(N = 44) M (SD)</th>
<th>CG(N = 43) M (SD)</th>
<th>t test</th>
</tr>
</thead>
<tbody>
<tr>
<td>I thought I can pay attention in class while the teacher presents the materials with the LMM tool.</td>
<td>3.50 (0.79)</td>
<td>3.14 (0.52)</td>
<td>2.509&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>I thought it was effective for teacher to use the LMM tool while presenting the teaching materials.</td>
<td>3.70 (0.67)</td>
<td>3.28 (0.83)</td>
<td>2.646&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>When I studied the materials through the LMS platform, I can easily recall the learning activities.</td>
<td>3.59 (0.82)</td>
<td>3.09 (0.84)</td>
<td>2.804&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>The materials on the LMS platform can help me to understand the difficult learning contents.</td>
<td>3.66 (0.75)</td>
<td>3.21 (0.80)</td>
<td>2.708&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>I was more confident for the exams because I can repeatedly study the material at any time and in anywhere out of class.</td>
<td>3.45 (0.70)</td>
<td>3.09 (0.81)</td>
<td>2.231&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>I thought it was effective to clarify the materials displayed in “desktop” area by offering annotation presentation.</td>
<td>3.77 (0.68)</td>
<td>3.21 (0.89)</td>
<td>3.332&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>The LMM tool displayed annotations which were helpful for me to promote my memorization, thinking, and clarification.</td>
<td>3.65 (0.81)</td>
<td>3.30 (0.89)</td>
<td>1.965&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>It was helpful for me to repeatedly study the materials for self-pacing learning.</td>
<td>3.61 (0.65)</td>
<td>3.14 (0.86)</td>
<td>2.894&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>I can study the instructional materials, clearly hear and see the solving-problem procedure.</td>
<td>3.55 (0.66)</td>
<td>3.09 (0.81)</td>
<td>2.851&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>I thought it can reduce the learning anxiety by reviewing the LMMs.</td>
<td>3.43 (0.66)</td>
<td>3.09 (0.53)</td>
<td>2.641&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>I thought the materials benefit me to make notes by studying the LMMs.</td>
<td>3.36 (0.72)</td>
<td>2.93 (0.88)</td>
<td>2.513&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>p < .05. <sup>b</sup>p < .01. <sup>c</sup>p < .001.

Discussion

It is useful for students to study the LMMs during learning. The following findings, based on the experimental results, are given to reflect that using the LMM tool can benefit teachers and students. First, students can easily recall the learning activities in the traditional classroom while studying the LMMs. The sort of the learning materials consists of teachers’ facial expressions, voices, handwritings, annotations, and teaching actions on the desktop of
teachers’ computers during instruction. According to multimedia learning theory (Mayer, 2001), students can process the information quickly because they can simultaneously review visual and auditory messages. Moreover, based on rich media theory (Daft, Lengel, & Trevino, 1987), the LMMs have a large capacity to transmit multimedia cues and to have a personal focus. Moreover, the LMM tool provides teachers with a flexible way to offer annotations so as to enhance students’ memorization, thinking, and clarification. Teachers can provide students with the explicit and implicit annotations for teaching units via displaying texts, video, images, and graphics in two regions, marquee and supplement information, of the screen, as shown in Figure 2(a).

Second, the number of areas comprising a screen layout can be flexibly increased or decreased by teachers according to their teaching requirements. That is, there are different deployments of using these regions composed of a screen such as reducing the number of regions on a screen layout or changing the positions of regions of a screen layout. Moreover, these regions inside a screen can be enlarged or shrunk. Accordingly, teachers can adequately adjust the positions and/or sizes of regions on a screen layout so as to draw students’ attention.

Finally, students can repeatedly watch the LMMs using the pausing/playing functions and the moving forward/backward functions. Students can effectively recall teachers’ explanations for their misunderstanding materials using above functions. According to media richness theory, these functions also help students to gain effective and adequate media. Thus, students can foster their self-paced learning. For the learning of mathematics induction, students can repeatedly study the LMMs to watch teacher’s solving-problem process so as to acquire more clear explanations for the process. Based on discussions mentioned above, the learning manner of using the LMM tool in classroom during instruction and then reviewing the LMMs out of class can definitely promote students’ learning effects.

Conclusion and future work

Conclusion

This paper has presented the LMM tool which provides teachers with several user-friendly functions to create the LMMs. Herewith, teachers can readily make the LMMs in the traditional classroom. Subsequently, teachers can upload the LMMs to the LMS. This is helpful for the students in a way of asynchronous learning to study the LMMs on the LMS out of class so as to recall instructional activities in class. Furthermore, the tool can help students who cannot present in the traditional classroom in a way of distance learning. More specifically, the tool can be applied to perform synchronous learning since teachers can exploit the LMM tool to produce the LMMs during lecturing and then simultaneously broadcast the LMMs in almost real time. Recently, mobile learning is a popular issue (Wang, Shen, Novak & Pan, 2009). Consequently, learners can also study the LMMs at any time and in anywhere using portable devices to realize mobile learning.

Future work

Future work includes five issues stated as follows.

- A study of an impact on synchronous learning using the IPTV of the LMM tool: The kind of synchronous learning can help students who cannot present in the classroom. It can broadcast the LMMs in almost real time and the LMMs also have a higher resolution than those existing screen capturing software produce.

- To apply the LMMs tool in the flipped learning through a MOOCs platform: Before lecturing, teachers can assign the homework studying the LMMs which can be downloaded from the MOOCs platform (Kim, Kim, Khera, & Getman, 2014; Kennedy, 2014). The contents of the LMMs will be further explained or discussed in the class by teachers during instruction. Meanwhile, the LMM tool can produce the LMMs consisting of the learning activities concerning teacher’s answers for students’ questions.

- To design a blended learning environment (Hastie, Hung, & Chen, 2010): The LMM tool produces the LMMs for some teaching units of the practice training courses such as computer language programming, software operation, writing mathematical proofs or equations, etc.
• To increase interactivity during lecturing: Teachers can select video clips of the LMMs, which mainly contain the students’ questions and then discuss these questions with students who never learn these teaching unit. This way will promote interactivity during lecturing.

• To improve teaching practices at classrooms: Teachers can watch the LMMs repeatedly to find out the shortcomings of their presentation and instruction for teaching lectures.

Acknowledgements

We would like to acknowledge the National Science Council (Ministry of Science and Technology) of Taiwan for supporting this research under Contract Numbers NSC 99-2511-S-150-002, NSC101-2511-S-150-003, and MOST 103-2511-S-150-002.

References


The Intelligent Robot Contents for Children with Speech-Language Disorder

Hawon Lee* and Eunja Hyun
Department of Child Psychology & Education, Sungkyunkwan University, Seoul, Korea // hawonlee@hanmail.net // hyunej@chol.com

*Corresponding author

(Submitted March 10, 2014; Revised September 29, 2014; Accepted January 22, 2015)

ABSTRACT
The purpose of this study is the development of language-intervention content, “Special Friend, iRobiQ.” The play-robot content serves various roles: The promotion of language interaction for children with speech-language disorders; an interlocutor friend (Talking Friend, iRobiQ); the transmission of data for diagnosis and evaluation between parents and experts; and a monitor for the supervisor of the speech-language disorder children outside of a therapy setting (Helping Friend, iRobiQ). This content design used a script approach, which provides entertaining educational elements for speech-language disorder children, as a player and learner. It has greetings and a birthday celebration script (cake, gift, song), with the theme of the practical language goals necessary for speech-language disorder children. When examining the effectiveness of its adaptability for use in the field, four certified speech-language therapists used the content with four autism/MR (Mental Retardation) children who were 4–5 years old over eight sessions in October 2012. The children learned to initiate conversations with the robot with the emotional exchange of expressions, as reported positively. These results suggest the future development of speech-language therapist assistant robots.

Keywords
Intelligent robot, Speech-language disorder, Young children, Script approach, Therapist assistant robot

Introduction
There have been recent developments in the field of education using new types of teaching strategies, purporting to help learners raise the efficiency of their learning. Hyun, Kim, and Jang (2008) suggested that education focuses more on the relationship between the education tool and the learner, rather than the teacher and the learner, as it has been in the past. One reason for this shift in interest is the ability of learners to learn independently, using their own media through the computer and Internet.

The basic research for e-learning and content development and their effects are being accomplished in various ways (Huang, Liang, & Chiu, 2013; Iskander, 2008; Lan, 2013; Njenga, & Fourie, 2010; Ramayah, Ahmad, & Hong, 2012). The special education field is concentrating its efforts on the relationship between media and children with disabilities in order to optimize the therapeutic effect (Jeon, 2014; Lee, 2013). In the language therapy field, video and audio content are used as teaching media (Buggey, 1995; Van Balkom, Verhoeven, & Van Weerdenburg, 2010; Beukelman, & Mirenda, 2012), and communication-assisting technology, termed an Augmentative and Alternative Communication System (AAC), assists children with language disorders who cannot communicate verbally. However, these media are used in a regulated interface, and are limited to one-way interaction (Han, Jo, Park, & Kim, 2005). In addition, virtual peers and avatars can be tools in speech education, but limited turn-taking and additional people who input language other than the therapist are inconveniences (Cassell & Ryokai, 2001; Ryokai, Vaucelle & Cassell, 2003).

Hence the need for solution media. Especially, it needs to have media like communication which is the essence of conversation between listener and speaker in speech-language therapy field. In this context, an intelligent robot with an evolved capability for direct interaction with learners has emerged (Hyun, Kim, & Jang, 2008). As the previous research on the robot shows, disabled children tend to interact smoothly with robots, as they are simple and predictable in their behavior and speech (Cho, Kwon, & Shin, 2009; Pierno, Mari, Lusher, & Castiello, 2008). Disabled children seem to show psychological stability with a robot who is an easy playmate (Dautenhahn & Billard, 2002), and willingly show communicational interests and development (Feil-Seifer & Mataric, 2008; Jeon, 2014; Lee, 2013). Such qualities of robot indicate its suitability as a tool in speech-language treatment within a special educational setting (Jeon 2014; Lee, 2013). Despite this, we use the Robot as an educational tool in general
Background and related works

Speech treatment using a robot

Autistic children tend to play with a humanoid robot or moving toys, but not with other children (Cohen, 1994, Dautenhahn & Billard, 2002; Feil-Seifer & Mataric, 2008). There is research showing that an autistic child using a humanoid robot can develop communication skills (Cho, Kwon, & Shin, 2009; Dautenhahn & Billard, 2002; Feil-Seifer & Mataric, 2008; Jeon, 2014; Lee, 2013; Robins, Dickerson, Stribling, & Dautenhahn 2004; Shin & Cho 2010). There is also a study of children using a robot as a mediator when communicating with others (Cho, Kwon, & Shin, 2009) with some research on how a robot can increase autistic children’s social skills. (Diehl, Schmitt, Villano, & Crowell, 2012; Shin & Cho, 2010). Some of the reasons for the effectiveness of robots in a disabled child’s speech-language treatment are their simplicity (the hypothesis of information quantity), safety, and their attributes as an instrument and a subject (Cho, Kwon, & Shin, 2009).

Humans use communicational interaction tools such as gestures, facial expressions, and verbal skills in messages that are more subtle and complicated than those used by a robot (Pierno, Mari, Lusher, & Castiello, 2008). However, a disabled child may have difficulty recognizing these complicated messages in conversational settings. As robots have predictable and simple conversational functions, it may provoke disabled children to interact with them to improve their interaction skills. Pierno et al. studied children’s reactions to a robot grabbing an object and a person grabbing an object. While a non-disabled child focused more on a human grabbing an object, a disabled child focused more on a robot grabbing an object (Pierno, Mari, Lusher, & Castiello, 2008), which was understood to mean that the disabled child preferred to imitate the robot as it had no complicated implications or signs. This corresponds to the hypothesis of information quantity, suggesting the importance of the amount of information presented in a conversation (Cho, Kwon, & Shin, 2009). Generally, humans use not only nonverbal cues such as facial expression or speech tone, but also the gestures of every body part to deliver information. This is too much information for a disabled child to perceive (Cohen, 1994). The various, complicated and sometimes delicate messages of humans, including those of a therapist, may decrease the disabled one’s desire to interact. As robots have predictable and simple conversational functions, it may encourage disabled children to interact with them to improve their interaction skills.

Next we consider the safety of robots. Disabled children would most likely think of robots as being safer than the possible threats presented in human interactions. If this hypothesis was correct, a disabled child would interact more with a robot than with humans. Dalton et al., using brain film with eye-tracking research, reported that disabled children feel threatened when people provide complicated non-verbal messages, even when these people are familiar to them (Dalton, Nacewicz, Johnstone, Schaefer, Gernsbacher, Goldsmith, & Alexander, 2005). A robot, which gives simple, delimited amounts of information, is recognized as a safe and stable being to the disabled children. Also, research has shown that the disabled ones increased the frequency with which they looked at the robot’s eyes (Mead, Wade, Johnson, St. Clair, Chen, & Mataric, 2010), indicating that they feel comfortable with robots.

A robot can provide the role of instrument and subject at the same time, according to Turkle (2005). He delineated that a robot not only works for the human, but can also work with and to the human, as it stimulates lingual capabilities as an instrument, and interacts and communicates as a subject with a disabled child. Robins et al. (2004) also suggested that a disabled child can not only use the robot as a communicator, but also as a mediator to interact with his or her parents. Language is a basic social building block, so disabled children who cannot obtain language have difficulty building social relationships. However, a humanoid robot can provide efficient media for the improvement of relationships. Therefore, that a disabled child can increase his or her social skills using a robot as a therapeutic medium is promising for the speech-language therapy field. Moreover, the autonomy and initiatives practiced by a disabled child with a robot as an assistive educational medium (Shin & Cho, 2010) minimize the intensive work of therapists and the child’s family. However, there is almost no substantial case of a treatment using a robot (Kang, Kim, & Lee, 2013), and most research has used AAC (Kim, 2003) such as a PC, communication board, or talker. While it is interesting and convenient, AAC lacks efficiency in its lack of ability to take turns, one of the most important purposes of speech therapy. Robots provide what is lacking in previous educational media, as people can relate to them more easily, it may present newer therapy-assisting methods for the 21st century.
Script approach method

Script theory began with research into computer information processing on how adults understand specific information in a periodic and spatial context (Kim, 2003). It is defined as simply analyzing the shared event knowledge between a listener and speaker in a particular context (e.g., washing hands script: Turn on the water → Wash hands with soap → Wash hands with water → Wipe hands with the towel). A disabled child can easily predict such to understand what people are saying, and can learn situational language within the scriptural context. Therefore, the script approach method is a representative treatment method in speech/language therapy. Various previous research studies have suggested the efficiency of the script approach method in speech therapy (Cherney, Halper, Holland, & Cole, 2008; Constable, 1986; Youmans, Youmans, & Hancock, 2011; Kim, & Lombardino, 1991; Tyler, & Watterson, 1991).

According to Nelson et al. (1979), the script approach method is efficient in the speech therapy of disabled children, in that the treatment method is based on contextual knowledge (Kim, 2003). Contextual knowledge is that in which a listener and a speaker can share a context immediately and simultaneously. A disabled child can easily recognize what a therapist or parents are trying to say from participating in the context together. The context of conversation gives disabled children the “ability of prediction,” which helps them communicate; the ultimate purpose of speech therapy. Such contextual knowledge gives the maximum opportunity for acquiring the language. The researchers particularly emphasize that the opportunities for turn-taking situations and proper language expressions that the script approach method provided are important for disabled children’s language development. They also suggest that as children get used to the set scripts, it is recommended that the script’s patterns or order are purposefully changed, so that they can expand in varied language expressions (Kim, 2003). The script approach method, which was connected periodically, spatially, and causally in daily life patterns, aided disabled children’s entertainment and enjoyment while learning, so that they increased their rate of language acquirement, and developed autonomy with motivation towards therapy. This current research has tried to optimize speech-language treatment using speech patterns with scriptural contexts in the iRobiQ. It has communication and appearance functions that enable it to be recognized as a friend to the children (Hyun, Lee, & Yeon, 2012) in Korea. Also, iRobiQ has shown two-way interaction and educational effects (Hyun, Kim, & Jang, 2008), and is expected to be distinguished in special education speech therapy, providing a suitable communication partner in the speech therapy field.

Therefore the objectives of this study were to develop content which is applied to the script approach method, a representative method of speech therapy, using the humanoid robot iRobiQ. In order to do that the following objectives were identified:

- To develop the speech-language treatment content using the iRobiQ
- To inspect developed contents responses by the experts

“Special Friend, iRobiQ” contents design

Development direction

Purpose and object

For this study, the iRobiQ (Figure 1) from Yujin Robot Co., Korea was used, because it is the most widespread educational robot platform in Korea; this 7 kg teacher-assisting robot is used in kindergartens. The iRobiQ has a maximum speed of 45cm/sec, it has one degree of freedom for its arms, two degrees of freedom for its head, and can produce emotional facial expressions ranging from anger, fear, disgust, sadness, happiness, and surprise to more complex expressions. There are six touch sensors in various locations on the robot; its head, arms, and wheels each have two sensors. Furthermore, there are infrared and bumper sensors, and the torso contains a 10-inch touch screen. The iRobiQ has an image recognition engine for face detection and recognition, uses marker-based AR technology, and uses English and Korean text-to-speech to recognize 80% of its 200 words.

The purpose of the content development design for the talking humanoid robot that could replace a conversational partner was the lack of a peer language partner in the speech-language therapy setting. iRobiQ provides the opportunity for general learned language skill acquisition for disabled children. The robot also assists therapists by relieving them of some of the more intensive work, as when the therapist and parent meet for counseling, because
iRobiQ can play with the disabled children, autonomically. A function of sharing the child’s speech samples for analysis between a doctor/therapist and parents and monitoring functions are also available.

Figure 1. iRobiQ

**Design**

As the name “Special Friend, iRobiQ” implies, the robot plays the role of friend for disabled child. It is able to make use of different language levels from the one-word stage (e.g., father, toy, milk) to the word-combination stage (e.g., sister’s shoes, the horse has gone). The iRobiQ has two different parts: “Talking Friend, iRobiQ” and “Helping Friend, iRobiQ,” (Figure 2) according to the specialized contents of its role. The “Talking Friend, iRobiQ” part was designed with the script approach method, having the “Greetings” and “Birthday Celebration” options as its detailed contents. We decided to focus on greetings and birthday celebration, as these are essential parts of children’s lives, including disabled children, and are a topic of importance in therapy settings. In “Helping Friend, iRobiQ,” reflecting the needs-analysis of the experts and doctors are the functions of sharing data between experts and parents, as well monitoring the children’s therapy room for supervisors, trainees, and parents.

Figure 2. Talking/Helping Friend iRobiQ

**Talking friend iRobiQ**

“Talking friend iRobiQ” has various entertaining scripts (see Table 1), stories such as greetings (e.g., entering in and out), and birthday celebrations, including a song, cake, and gifts (Figure 3).

In the “Entering In and Out” section, when a child initiates a conversation with iRobiQ, it expresses its thanks with feedback and joyful facial expressions as positive reinforcement. Such expression helps children initiate conversations, which is often lacking in disabled children. As shown in Figure 4, iRobiQ has a script suggesting the sharing emotions (joy, sadness, anger, tiredness, worries, pain) with the child. It creates a prosocial relationship between iRobiQ and the child, so that they can become more intimate communication partners with one another. This function provides emotional interaction between an iRobiQ and a disabled child, whereas the previous version used in preschool only provided attendance checking and one-way expression of the child’s emotions to the iRobiQ.
(Hyun, Kim, & Jang, 2008). The newly developed version provides authentic communication media, as in the following example:

Child: “How do you feel?”
iRobiQ: “I am so glad to see you, let’s have some fun today!”

The “Birthday Celebrating Script” section has a song, cake, and gifts, which are essential elements of a birthday party, enabling a disabled child to practice communicating with iRobiQ as though at a real birthday party. Especially, its use of picture cards optimizes HRI (Human–Robot Interaction). The picture cards have a picture-recognizing sensor, so when the child presents a song card to the robot, the robot is able to recognize it, and the birthday song is played. Moreover, the child can select the birthday gift card and suggest it as a present to the robot in the gift script section. This process provides instrumental functions of language such as gestures and finger pointing. Language is a tool of communication for conveying what people want; hence, after acquiring the concept of using tools such as gestures, cards, and pointing to get what one desires, a child can start to communicate. It is the “means-end concept” of the developmental psychology. In using such, the child comes to know that when he or she wants something, they need to suggest it, as when they use the cards as instruments.

In the motion of blowing out the candles, the robot with its volume-recognizing system asks, “Oh, the candle did not go out, blow harder!” when the child blows too weakly. Such feedback would be for a realistic situation, so the child can train their breath-control. Most autistic children do not have a normal breathing pattern, especially in the volume of their breath (Sae-Nu-Ri Handicapped Parent Association, 2010). Using this function, the child is provided with good practice for improving their breathing. The content also has a scissor-rock-paper game, which is often played by preschoolers (Figure 5), with feedback for winning and losing. Children are able to learn how to speak in various unpredictable situations rather than with the same response, as they experience altered scripts in the advanced version over the fixed-script programs.
There is also the “King–Servant” play. This activity prevents children with language disorders from being caught into only answering. The disabled child mostly replies very compliantly, without autonomous questions or orders, in his or her cognitive deficiency. They need to practice active speech, including questioning, requesting, ordering and claiming, not only replying. In this activity, the child can command iRobiQ (e.g., “Raise your right hand!”), and the robot reacts to the commands. Children with language disorders are used to following orders, and this activity lets them experience how to be self-reliant and active; it is seen as enjoyable at the same time. iRobiQ’s responses are also random with a shifting script, enabling the child to say what is wrong and how to fix it. The robot also has a program that invites another person into the conversation (e.g., “My friend, you should give that cake to the teacher”). This invites an adult into the communication, so that they can experience a more dynamic conversation program together.

“Gift script” is divided into two cases of ready and not ready to give. A gift (Figure 6) is suggested through picture cards, and iRobiQ can spell out the names of the gifts through its picture-recognizing sensor (e.g., Wow! It is a toy. I like this). When the gift is not ready to be given, iRobiQ shows disappointment and asks for an apology from the child.

<table>
<thead>
<tr>
<th>Contents</th>
<th>Screen image</th>
<th>Story</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entering in &amp; out</td>
<td><img src="image" alt="Screen image" /></td>
<td>When the child enters in, he/she says hi to iRobiQ. After greetings, the child and iRobiQ may build friendship by sharing their feelings. When the child leaves, iRobiQ asks the child where the child is going, with whom, and the expected return point, trying to continue the friendship.</td>
</tr>
</tbody>
</table>
Birthday Song

Special friend iRobiQ invites his friends to his birthday party. iRobiQ sings the Birthday song with his friend (child) and builds their friendship.

Birthday Cake

Special friend iRobiQ has good manners. He not only cuts cakes and shares with his friend, but also gives the cake to the teacher. In addition, When the candles do not go out, he encourages the child to blow harder.

Birthday Gift

Special friend iRobiQ is a friend of kids, so he acts like a kid. One of his behaviors is that he has a lot of interest in birthday gifts. He gives thanks to the child who has brought the gift. He is disappointed with anyone who has not prepared a gift for him, and he asks the child to prepare a gift for him next time.

The given target language in the therapy setting is analyzed by the semantic relationship and pragmatic function. iRobiQ has a variety of utterances with the same meaning in the same scenario. For example, Take care, bye-bye, and I’m missing you are all farewell greetings, with the same meaning. As such, the disabled child can listen to a variety of sentences with the same meaning. This encourages them not to respond from memory, but with responses that are more flexible (Table 2). This is made to be as close to real world situations as possible, as if the child is communicating with another person.

Table 2. The semantic and pragmatic analysis of conversation

<table>
<thead>
<tr>
<th>Robot utterance (Random)</th>
<th>Target language</th>
<th>Language structure</th>
<th>Semantic relationship</th>
<th>Pragmatic function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oh? Watch out, there is a candle, who wants to blow it?</td>
<td>I want to do it.</td>
<td>Agent + Action</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oh, the candle is melting. Who wants to blow it?</td>
<td>I want to do it together.</td>
<td>Condition + Action</td>
<td>Responding to the Question</td>
<td></td>
</tr>
<tr>
<td>Huh? Here is a candle. Who wants to blow it hard?</td>
<td>I want to do it with my teacher.</td>
<td>Comitative + Action</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ouch, it is hot. There is a candle, who wants to blow it?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Okay, but we are doing it together, are you ready?

Really? But let’s do it together.       Okay.   Communication devices (Affirmation)  Responding to the Question

Yes.  

Hmm.. Okay, but how about doing it together?

One, Two, Three       One Two Three       Automatic Speech  Repetition

Hoo, Wow            Hoo Wow            Accompanying Sounds  Repetition

Every script is managed by the therapist with a remote controller, which means that the therapist can control the script order and progression speed according to the child’s language level. This means that a shifting script is possible, which is an essential part of education. The shifting script is made so that disabled children can respond properly to unpredictable situations and so become used to various real-life situations, a novel development from previous education media with its predictable, uniform content. The ability to shift the script is necessary for children who have previously learned through a set script, as the shifting script acts as a mediator that generalizes content they have learned into real-life situations.

Helping friend iRobiQ

Helping Friend iRobiQ has two functions (Figure 7). The first is the recording and sending data function, where it can send and receive a child’s data (language and activity level, therapy scene, behaviors) to parents and experts. The second function is monitoring, which allows the parents, trainees, and supervisor to watch over the child as may be necessary. They can observe therapy sessions through a smart-phone with the ROBOTOK app. This functionality is necessary when the author, who is a speech therapist, is performing speech therapy with disorder children in the field. It is also necessary for when parents and trainees want to see the therapy to share the child’s condition with experts, not just linguistically share by making verbal promises or a record, but also suggest precise recommendations and provide assessment through a real audio-visual recording of a child with a disorder. This function could prove necessary in the field of special education with R-Learning.

Methodology

This part is field application for availability for developed content. This was survey/interview research and the methods were as follows:
Participants

The participants of this study were four speech-language disorder children and their therapists. Three autistic girls were at the speech-language therapy center, and one boy at the preschoolers with disability integration education center (Kyounggi-do Province, Korea) had MR (Mental Retardation). The children’s average age was 56 months (4; 8), and their language stage was 2–3 word combinations. All therapists hold first rank speech-language therapy licenses, are affiliated with a therapy center, and have more than 16 years of clinical experience.

Research instrument

The research instrument was a questionnaire obtaining information concerning the suitability of developed content. Participants in the testing of the pilot questionnaire are 3 speech-language therapists. The questions in the test are related to knowledge and availability of R-Learning in speech-language therapy settings. The questionnaire was composed of checklist and open end questions. The questionnaire was divided into 3 parts “talking friend iRobiQ,” “helping friend iRobiQ” and “ETC.”

The responses at the field application of content are to the questions that are as in Table 3. In the 37 questions, two are yes or no type in a two-point scale to see the preliminary knowledge of speech therapy using a robot. For [talking friend iRobiQ] content, 29 questions are asked to determine suitability (two questions), interest (eight), availability (two), treatment effect (two), interaction (seven), and function (eight) in five scales (strongly agree, agree, neutral, disagree, strongly disagree). For [helping friend iRobiQ], six questions are used to determine the need for content on the same five-point scale; also, any opinions can be written.

Procedure

The authors have carried out surveys with the experts who used this content. Data collection was finished after the treatment of speech-language disorder children by experts. Three speech-language therapists used the content with three autistic children, each for about two weeks, October 9th–21st, 2012 (eight sessions). A special education expert used this content in group activities with an MR child and ten non-disabled children for the period October 10th–17th, 2012. All data were given to authors directly.

<table>
<thead>
<tr>
<th>Table 3. Category for questions in survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
</tr>
<tr>
<td>Talking Friend iRobiQ</td>
</tr>
<tr>
<td>Perception on Robot therapy</td>
</tr>
<tr>
<td>Suitability of topics</td>
</tr>
<tr>
<td>Interest</td>
</tr>
<tr>
<td>Availability</td>
</tr>
<tr>
<td>Treatment effects</td>
</tr>
<tr>
<td>Interaction</td>
</tr>
<tr>
<td>Function</td>
</tr>
<tr>
<td>Helping Friend, iRobiQ</td>
</tr>
<tr>
<td>Necessities</td>
</tr>
<tr>
<td>ETC.</td>
</tr>
<tr>
<td>Suggestion</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
Results

The results from the experts working with autistic children were as follows in Table 4. Most responses were positive in the survey; however, two of the three experts responded that they were not knowledgeable about “robot therapy,” and believed that it is not the best method of treatment for speech-language disorder children, and that human therapists may be better at treating language disorders. Those less inclined to use the robot also claimed that the treatment was wholly a human responsibility, and that a robot may not be able to fulfill the various roles of humans.

In Talking friend iRobiQ, two of the three experts gave positive responses for the suitability of the content. The reason reported that the subject of content reflects daily life of children. However, two of three experts saw the availability of content as negative. The reason was that human beings are more suitable than robots as a tool for therapy. The interest in content scored 8 out of 15 points (five questions x three experts), which was the highest point of all.

The claim that familiarity and maintenance factors of the robot in the interest category led to a friendly social response also had two out of three positive responses. All the surveyed experts responded that there are effective language treatments when using iRobiQ, in that the autistic children said “(calling) iRobiQ,” “(patting iRobiQ’s head) I love you” and “Let’s play together,” which are expressions used in interactions with peers. All respondents to the survey were positive, except an expert who was a neutral respondent for the robot’s functions and interactions. There are “convenient and using remote controller” parts in interaction category. For the neutral responses in interaction questions, there have been slow technical responses for robots, so that interaction may not have been comfortable. For functional questions, the reason reported is the human factor that the robot was not able to give a natural response to children for speech that was outside of its script.

For Helping friend iRobiQ, its data transfer function had two of three positive responses, its recording function had two neutral, and its monitoring capabilities had three neutral responses. The reason for this response to data transfer is convenience, of recording and monitoring is that robot is not the only tool with such functionality. They reported that smartphones are also capable of recording and monitoring. Finally, the suggestion from the experts was the development of more content other than the birthday celebration script, such as hospital play, market play, and school play scripts to help the child adjust to a variety of situations in their daily life.

Table 4. Survey results

<table>
<thead>
<tr>
<th>Category</th>
<th>No. Questions</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talking Friend, iRobiQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception of robot therapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Have you ever heard of robot therapy?</td>
<td>Yes No 3</td>
<td></td>
</tr>
<tr>
<td>2. What do you think of using a robot in therapy?</td>
<td>Positive Negative 3</td>
<td></td>
</tr>
<tr>
<td>Suitability of topics</td>
<td>3-4 Suitability</td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>5-6 Maintenance</td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>7-8 Motivation</td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>9-10 Familiarity</td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>11-12 Fun</td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>13-14 Spontaneity</td>
<td></td>
</tr>
<tr>
<td>Availability</td>
<td>15-16 Availability</td>
<td></td>
</tr>
</tbody>
</table>
Treatment effects | 17-18 Effect | 3 | 3
Interaction | 19-20 Convenient | 1 | 2 | 3
Interaction | 21-22 Remote controller | 3 | 3
Function | 23-24 Size | 2 | 1 | 3
Function | 25-26 Loudness | 3 | 3
Function | 27-28 Image | 2 | 1 | 3
Function | 29-30 Humanoid factor | 1 | 1 | 1 | 3

Helping Friend, iRobiQ

| Necessities | 31-32 Monitoring | 3 | 3
Necessities | 33-34 Recording | 2 | 1 | 3
Necessities | 35-36 Sending data | 1 | 2 | 3

ETC

| Suggestion | 37
Total | 37

The interview focused on the MR child’s language expression, its type, and changes in their relationship with their peer group while using the iRobiQ. As a result, the MR child used intimate expressions when speaking to the iRobiQ such as “Are you hungry?” “Sing for me,” and “Where are you, iRobi?”. The expert also reported that the MR child had increased their relationship language with their friends, even though the MR child had shown almost no expression to the peer previously, and this content can be a communication medium for mutual reactions, as the disabled preschooler invited normal preschoolers. The expert reported that “Helping Friend iRobiQ” plays very necessary functions in treatment and caring in the education environment, because both parents and experts can reliably check on the disabled preschoolers’ ability to communicate with other normal preschoolers, and their behavior for diagnosis outside of the education center and home environments. To compare the results of children with mental retardation and children with autism, both recognize iRobiQ as a peer, so that their social speech interaction increased. However, children with mental retardation produced learned language in interactions with normal children better than did children with autism. The source of this difference is also related to their diagnosis, but it is evidence to show that language-disorder children improve their language using the Robot.

**Conclusion**

This research is on the speech-therapy assisting robot content of “Special Friend iRobiQ,” in its plans and development for children with speech-language disorders. While the role of language is communication with others, current treatment methods mostly can only provide the experience of lingual interaction with a therapist, or with a computer as a one-way medium. Disabled children find it difficult to interact with people, as humans use complicated language cues. Therefore, by using a robot with special education media in the form of a friend, we have come to know that a robot can have a significant effect in assisting speech treatment, and may provide a different role from that of a computer.

Considering the responses of content availability and robot perception, the use of a robot as an assistant in a therapy setting and the background pertaining to the therapy is currently lacking in Korea. It is the same result as Kang, Kim, & Lee (2013). In the interest question of content, the intimacy and maintenance had positive responses that are the same as found by Han, Jo, Park & Kim (2005), Kim, Berkovits, Bernier, Leyzberg, Shic & Paul, Scassellati (2013). Hyun, Lee, & Yeon (2012) also reported that iRobi is recognized as a friend by children as a humanoid robot.

In addition, the questions tell how a robot interacts mutually, unlike computers, so that durability increases are the same as found by Pae (2013). According to Pae (2013), computers also provide interesting factors for speech-
language disorder children, but they do not have language interactions with a computer as a therapeutic medium, only using it as a tool for language practice. The cause of this maintenance is from the variety of stimuli of robots for children, which shows similar results as Han, Jo, Park & Kim (2005), Kanda, Hirano, Eaton & Ishiguro (2004), Kim, Berkovits, Bernier, Leyzberg, Shic, Paul, & Scassellati (2012).

Treatment effects garnered positive responses from all experts; this is the same result as Lee (2013), who studied the syntactic abilities of autism through this medium, the increase of speech length (Jeon, 2014), and the increase of speech interaction (François, Powell, & Dautenhahn, 2009; Goodrich, Colton, Brinton, Fujiki, Alan Atherton, Robinson, Ricks, Maxfield, & Acerson, 2012).

In Cassell & Ryokai (2001) and Ryokai, Vaucelle & Cassell (2003) who carried out speech education using virtual peers, narratives were improved. iRobiQ is both tangible and can move, so it can interact like a real friend (Hyun, Lee & Yeon, 2012) and is easy to generalize to normal children, but a virtual peer cannot. Furthermore, when the speech is not understood, the virtual peer was silent and stopped taking its turn, whereas iRobiQ used the context of script for interaction to achieve a target language without silence. It shows language modes with more natural interactions, like human beings.

In functional questions, the humanoid factor was not the only positive response, but in intimacy questions, its humanoid qualities garnered positive responses. The reason for these two seemingly conflicting responses, according to the written opinions, is that humanoid factor was assessed through appearance in intimacy and through speech in function. This means that interacting with a robot does not produce various pragmatic skills as interacting with a human would.

This is internationally the first robot content with a script approach as a therapist assistant robot has been tested, with more functions developed for special educational purposes from the original iRobiQ version, as is used in the normal preschooler education setting in Korea. The robot is available for children to emotionally interact with and so acquire language, with a humanoid robot in a predictable script flow, making for optimized HRI. Furthermore, this is suitable for assisting experts help disabled children and their families.

The content of this paper is about the first R-Learning prototype for assisting disabled children with speech/language disorders in Korea. Therefore, in order to apply this content to the speech therapy field, we need long-term, in-depth experiments with the development of a variety of content. In spite of this, our current work entails the validation of significant R-Learning therapy content in Korea, by which we expect to inform its future refinement, and its possible extension to a larger set of disabilities and types of content.

References


Sae-Nu-Ri Handicapped Parent Association (2010). A Parent’s guide to autistic child. Retrieved October 10, 2012, from http://cafe.daum.net/92man/59pG/120?docid=4196734662&q=%C0%DA%C6%F3%BE%C6%B5%BF%20%C8%A3%C8%ED&re=1


Exploring Application, Attitudes and Integration of Video Games: MinecraftEdu in Middle School

José-Manuel Sáez-López1*, John Miller2, Esteban Vázquez-Cano1 and María-Concepción Domínguez-Garrido1

1Spanish National University of Distance Education (UNED), Spain // 2Chalone Peaks Middle School, CA, USA // jmsaezlopez@edu.uned.es // jwmiller@kcusd.org // evazquez@edu.uned.es // cdominguez@edu.uned.es

*Corresponding author

(Submitted March 15, 2014; Revised November 5, 2014; Accepted November 16, 2014)

ABSTRACT

The aim of this study is to assess the use of MinecraftEdu in classroom practice analyzing the outcomes and attitudes of all members of the educational community through a quasi-experimental approach. The research presents three dimensions oriented to assessing the use of this application in a didactic unit “History and Architecture” compared through statistical inference ($t$-student) to a control group that develops the same unit with slides and traditional expositional methods. The second dimension values the attitudes of teachers, students and parents regarding the implementation of video games in formal education using descriptive analysis and nonparametric statistical inference through the Jonckheere-Terpstra test and the Kruskal-Wallis test, which allows each group ranks to be compared. The third dimension analyzes interactions in a virtual learning environment related to the implementation of MinecraftEdu. Although there are no significant improvements regarding academic outcomes and some parents hold negative attitudes, it is noteworthy that the majority of the sample considered that MinecraftEdu is fun, enhances creativity, develops discovery and is a good application for creating and exploring immersive historical environments.

Keywords

Computer-mediated communication, elementary education, game-based learning, improving classroom teaching, interactive learning environments, teaching/learning strategies

Introduction

Since the 80s, video game use has risen to the point where 60% of children between 8 and 18 years old now play them (Rideout, Foerh, & Roberts, 2010). Pew Internet and American Life Project showed that recreational use of video games is widespread, with 97% of young people and 53% of adults using them (Lenhart, Jones, & Macgill, 2008; Lenhart, Kahne, Middaugh, Macgill, Evans, & Vitak, 2008).

Video games are popular mainly because they are fun. Teenagers’ intrinsic motivation towards games contrasts with their often noted lack of interest in curricular contents (Prensky, 2003). Motivation could be combined with contents in school (Eseryel, Law, Ifenthaler, Ge, & Miller, 2014); thereby, video games may also have advantages from a pedagogical perspective. Educational research provides findings that help to determine whether it is advisable to adopt goals and encourage learning activities that are meaningful and motivating for students.

Several theorists claim that there is insufficient scientific evidence regarding the relationship between gaming and learning. “There is not enough research to determine the relationship between video games and learning” (Blunt, 2007, p. 2). There is limited evidence regarding how educational games can be used to solve the problems inherent in the structure of traditional K–12 schooling and academia (Young, Slota, Cutler, Jalette, Mullin, Lai, Simeoni, Tran, & Yukhymenko, 2012).

Some authors ensure that there is no theoretical basis in this field. “I challenge anyone to show me a literature review of empirical studies about game-based learning. There are none. We are charging headlong into game-based learning without knowing if it works or not. We need studies” (Cannon-Bowers, 2006, p. 2).

Educational video games require a greater foundation in the evaluation processes. “Although a number of frameworks exist that are intended to guide and support the evaluation of educational software, few have been designed that consider explicitly the use of games or simulations in education” (de Freitas & Oliver, 2006, p. 262).
It is essential to “research educational video games already in use” (Young et al., 2012, p. 81). Some teachers utilize educational video games in their daily practice, therefore, analysing their current application would provide more valuable information regarding how video games influence student performance.

Taking the aforementioned research needs related to game-based learning into consideration, the motivation of the present research aims to provide information regarding the use of MinecraftEdu in educational settings, particularly in middle schools.

Several studies highlight the advantages of game-based learning as environments that promote student motivation and engagement (Blunt, 2007; Gee, 2007; Greenfield, 2010); therefore, it is important to confirm advantages related to this approach in educational settings though educational research. Some institutions, such as the Sweden educational system, are considering including Minecraft as an essential tool across the curriculum, even as a mandatory class (http://www.edudemic.com/this-swedish-school-now-has-a-mandatory-minecraft-class/). Thereby, administrators, policy makers, teachers, parents and students need to understand real possibilities related to game-based learning in general and with MinecraftEdu in particular.

**Theoretical framework**

Klopfer, Osterweil, and Salen (2009, p. 21) define digital-learning games as: “Those that target the acquisition of knowledge as its own end and foster habits of mind and understanding that are generally useful or useful within an academic context.”

Serious games are defined as immersive virtual environments explicitly trying to educate (Shute, Ventura, Bauer, & Zapata-Rivera, 2009). The most important features of game-based learning are related to the fact that they are educational and they allow interaction in the virtual environment. In these environments, players are part of the learning environment, as decisions directly affect the course of the game (Prensky, 2001). The general trends in research indicate an increasing popularity among students using game-based learning that is integrated into the objectives of the curriculum (Aldrich, 2004; Blunt, 2007; Young et al., 2012), detailing statistically significant improvements. Games are able to promote higher-order thinking and social skills.

Much had been written about the educational potential of video games. Several studies present positive evidence regarding the use of game-based learning in educational contexts (Barab, Dodge, Ingram-Goble, Pettyjohn, Peppler, Volk & Solomou, 2010; Blunt, 2007; Chen, Shih, & Ma, 2014; Eseryel et al., 2014; Hickey, Ingram-Goble, & Jameson, 2009; Shaffer, 2007; Squire, 2006; Steinkuehler, 2006; Young et al., 2012), detailing statistically significant improvements. Games are able to promote higher-order thinking and social skills (Dondlinger, 2007; Steinkuehler & Duncan, 2008). In this context, positive evidence in several studies recommends designing games for educational purposes, reinforcing the concept of game-based learning.

From this evidence the questions are: Do teachers really apply game-based learning in education? What do students, teachers and parents think about this approach? In spite of the previously mentioned positive evidence, many educators are not open to the idea of using video games in their classrooms (Mayo, 2009). Another important question would be: “Do video games enhance academic achievement?” (Young et al., 2012, p. 84).

Prensky (2001) stresses that game-based learning provide feedback and enhance the development of activities related to real life and foster skills related to problem solving. With a proper design suitable for teaching, these resources can be applied to activities in which students solve problems and develop content. “Some educators see games as a useful and perhaps even necessary learning environment suitable for learners of all ages” (Blunt, 2007, p. 2).

**Active approach of video games**

Learning is more effective when it is active and problem-based and gives immediate feedback. In a context focused simply on acquiring information for later playback, responsibility and authority are external to the students, so this practice undermines the learning process (Gresalfi, Martin, Hand, & Greeno, 2009).
Educational video games foster the fact that students are actually part of the learning environment, rather than being a passive recipient listening to someone with more experience. “One of the most powerful opportunities offered by games is that players are not just observers but are often protagonists who make decisions that affect the game world” (Barab et al., 2010, p. 527).

Gee (2004) shows in his study that educational video games are a learning tool that allows students to be placed within the learning environment and contribute actively in the educational process. In the real world, constructivist learning experienced by players in an educational video game offers one of the few available truly three-dimensional learning experiences (DeKanter, 2005). Gee (2003) notes that players experience the game in a different way to reality, because in the real world they cannot test and test all around them.

More generally, game-based learning offer new technologies and methodologies for creating a deeply immersive and highly interactive curriculum. Studies detailing experiences with games include serious discussions that describe the approach related to educational benefits regarding gamification.

There are several video games that create a 3D immersive environment that recreates a period of history, so players experience interesting interactions with this kind of game. There are several games, such as “Civilization, Age of Empires and Rome Total War, that provide the opportunity to recreate historical events. Narratives embedded in historical content allow history games to offer unique affordances for reenacting, replaying, and gaining first person experiences within the realms of history and social studies” (Young et al., 2012, p. 78).

History-based video games, properly implemented from a pedagogical perspective, are motivating and engaging for students (Devlin-Scherer & Sardone, 2010; Lee & Probert, 2010; Watson, Mong, & Harris, 2011). However, we have to be careful regarding historical misconceptions and inaccuracies that may be fostered by video games (Charsky & Mims, 2008), and take into account that adding text or historical information to gaming is not enough to foster learning (Akkerman, Admiraal, & Huizenga, 2009; Moshirnia & Israel, 2010).

Method

Research design

There are models and methodological approaches in educational research, and the proposed research model applied is Design-Based Research (DRB) (Anderson & Shattuck, 2012), which is a strategy that allows a systematic and interactive process focused on learning and research as subjects to innovate in educational contexts. This naturalistic approach enables understanding of learning processes through informed exploration, enactment, evaluation within a local context, and the development of design principles (Anderson & Shattuck, 2012).

The DRB approach improves educational practice and research processes. “DBR offers a “best practice” stance that has proved useful in complex learning environments, where formative evaluation plays a significant role, and this methodology incorporates both evaluation and empirical analyses and provides multiple entry points for various scholarly endeavors” (Dede, Ketelhut, Whitehouse, Breit, & McCloskey, 2009, p. 6).

The main objective is to analyze the educational benefits of MinecraftEdu in middle school (6th to 8th grade). The specific objectives are:

- Check students’ outcome improvement with MinecraftEdu.
- Assess students’ outcomes with respect to learning, motivation, fun and engagement when they use video games in the history classroom.
- Analyze interactions regarding the use of games in virtual learning environments.
- Assess attitudes of the school community regarding the implementation of MinecraftEdu in history.

This research presents the following dimensions (Table 1):
Table 1. Dimensions, indicators and survey instruments (Sáez López, Leo, & Miyata, 2013, p. 6)

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Indicators</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension 1: Evidence of learning with MinecraftEdu</td>
<td>Student motivation, Digital literacy, Active approach, Process evaluation, Academic results</td>
<td>Test results academic unit (P1mp), t-student, Control G and Exp G.</td>
</tr>
<tr>
<td>Dimension 2: Attitudes of parents, teachers and students regarding the use of MinecraftEdu</td>
<td>Educational effectiveness, Content development, Fun, Creativity, Discovery in the virtual world</td>
<td>Questionnaire education mixed (APMA)</td>
</tr>
<tr>
<td>Dimension 3: Analysis VLE interactions regarding use of MinecraftEdu</td>
<td>Collaborative work, Safe handling class group, Student engagement, Using resources and communities, Interaction and communication</td>
<td>Message analysis – Edmodo VLE posts (HyperResearch)</td>
</tr>
</tbody>
</table>

Basic principles

Research intervention analyzes the pedagogical implementation of cross-curricular thematic approaches through interdisciplinary approaches. Principles of instruction (Gagne, Briggs, & Wager, 1992) are elemental in this process: gain the attention of the students, inform students of the objectives, stimulate recall of prior learning, present the content, provide learning guidance, elicit performance (practice), provide feedback, assess performance, enhance retention and transfer to the job.

Significant prior learning is important from the perspective of other classic authors being taken into account in this pedagogical design and collaborative learning through critical thinking, discovery learning (Ausubel, 1978; Bruner, 1966) and Project-Based Learning (Jonassen, 1977). Social interactions in learning environments are essential from the perspective of constructivism and sociocultural theory (Vygotsky, 1978). Interactions between the social and cultural context are important, developing educational activity and situated learning (Brown, Collins, & Duguid, 1989; Wenger & Snyder, 2000) and enabling active participation in learning communities with an intercultural component in this case.

Participants

In the first dimension, the research sample included students that are from 11 to 14 years old from several schools in the USA and Spain; they participated voluntarily. The experimental group included 131 students that worked with MinecraftEdu in classrooms, 41 students from a school in the province of Albacete (Spain) and 90 students from a high school in California, USA. Moreover, there were another 50 students (acting as a control group) from a school in the province of Cuenca (Spain). The experimental group was 61% of girls and 39% of students. Contingency analysis is not detailed because there are no significant differences regarding gender, country or school.

Dimension 2 analyses the attitudes of the school community across 205 participants (62.9% female and 37.1% male) who participated voluntarily. In this dimension, the sample consisted of 10.7% teachers, 25.4% parents and 63.9% students, who belong to the experimental group (Figure 1). Regarding countries, 50.7% of the sample was from Spain and 49.3% was from the USA. The experimental group of 131 students took part in the Edmodo platform, and interactions are analysed in Dimension 3 (Figures 7 and 8).
Procedure and intervention

This research analyses the pedagogical use of the application MinecraftEdu (http://minecraftedu.com/), which is the educational version of the popular Minecraft. Several teachers have designed and developed units and educational projects to work on in this program (http://minecraftedu.com/wiki/index.php?title=Main_Page).

MinecraftEdu is a collaborative effort of a team of educators and programmers in the United States and Finland in collaboration with Mojang AB in Sweden. It is intended that the application is affordable and accessible to schools worldwide.

Narrative video games in education can become active curricula that promote dynamic interaction between players and the storyline, between action and understanding (Barab et al., 2010). However, MinecraftEdu is an open virtual world in which there is not any plot or story; it leaves full freedom for exploring everything. You can explore, create, discover and experiment in this immersive environment in collaboration with classmates and tutored by the teacher, who also has an avatar in this world (Figure 2).

Through Local Area Connection (LAN) teachers and students connect and enter this world in class. The teacher designs the unit and creates a map and may raise allocations that students should develop within this environment. The possibilities for interaction, exploration and discovery are numerous. This Mod is designed to give full control to the teacher in this virtual world.
One effective use of this application is to analyze the impact of educational contexts. It also discusses the attitudes of the school community to this approach.

It applies a unit entitled “History and Architecture,” which contains learning about ancient civilizations and buildings in these periods, including the Chichen Itza Pyramid, the Roman Colosseum, the Pantheon in Rome (Figure 3) and medieval buildings (Figure 4).

![Figure 3. MinecraftEdu History and Architecture unit (Sáez López, J. M, 2015b)](image1)

Through a quasi-experimental approach and data triangulation, the research presents three dimensions that seek to respond to the research objectives. In the first dimension, the mean of a control group and an experimental group is compared from results of a test after students worked on the mentioned unit (Test 1 MinecraftEdu Project, P1mp). In the second dimension, attitudes of the school community regarding game-based learning are measured. In the third dimension we analyze the interactions of students and teachers from Spain and the United States on the Edmodo platform.

![Figure 4. Medieval world based on a real English city (Sáez López, J. M, 2015a)](image2)

**Instruments**

Information is collected using a test (P1mp), a mixed questionnaire (APMA) and by analyzing messages on the Edmodo platform. These instruments present content validity through 14 judges in Spanish National University of Distance Education (UNED). Through data triangulation there is sufficient evidence to uphold the validity, which minimizes error variance (Goetz & LeCompte, 1988). Data triangulation (Cohen, Manion, & Morrison, 2000) was developed from quantitative test information (P1mp), the questionnaire, and the contributions in the analysis of the messages and open questions.

To assess the use of this application teachers developed a didactic unit “History and Architecture” in a control group with traditional expositional methods, using slides. Moreover, the experimental group learns all the content from this unit through the immersive environment created on MinecraftEdu. In order to assess the knowledge acquired by the
control group and the experimental group all students take a test (P1mp) after developing the “History and Architecture” unit.

The application of a student $t$-test enables analysis of the significant differences in academic performance. Test results (P1mp) are compared through statistical inference ($t$-student) comparing means from these independent groups.

Moreover, an AMPA mixed questionnaire also discusses the views and attitudes of the school community (parents, teachers and students) regarding pedagogical use of MinecraftEdu. These results derive from a descriptive analysis, with nonparametric statistical inference through the Kruskal-Wallis test and the Jonckheere-Terpstra test with a 0.01 level of significance (Table 4) analyzing the rankings obtained in these tests.

Applying data instruments provided with different approaches enables data triangulation, which reinforces the research validity (same result from different instruments and tests). The Cronbach’s alpha reliability in the APMA questionnaire is 0.793.

Moreover, in dimension 3, the research shows an analysis and classification of messages on Edmodo using the HyperResearch application in order to appreciate students’ interactions and discussions, from a creative perspective, of the use of MinecraftEdu in the mentioned didactic unit (Figures 7 and 8).

Design-Based Research (DRB) allows a systematic and interactive process focused on learning and research as subjects to innovate in educational contexts, understanding of learning processes. Data triangulation using different dimensions and instruments confirms and validate obtained results.

**Results and discussion**

**Dimension 1: Evidence of learning with MinecraftEdu**

The control group students work on the contents of the unit “History and Architecture” through an expositional approach with slide shows. The unit was developed by the experimental group using the MinecraftEdu application (http://ticjm.blogspot.com.es/2013/03/minecraft-edu-primary-school-project.html). Once both groups had finished their unit and taken the P1mp test, the results were as follows:

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>Median</th>
<th>$SD$</th>
<th>$SE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>131</td>
<td>8.98</td>
<td>1.113</td>
<td>.097</td>
</tr>
<tr>
<td>Control Group</td>
<td>50</td>
<td>8.78</td>
<td>1.375</td>
<td>.194</td>
</tr>
</tbody>
</table>

Given the sample size and the Kolmogorov-Smirnov test, normality is assumed. There is also equality of variances due to the significant value of 0.07 in the Levene test. Therefore, the requirements for implementing the test are confirmed. The $t$-test gives a value of 0.996 (significance of 0.32), so the difference is not significant. Although the experimental group obtained a higher mean, there is no significant difference in scores between the control group and the experimental group.

<table>
<thead>
<tr>
<th></th>
<th>Levene’s test for equality of variances</th>
<th>$t$-test for equality of means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$</td>
<td>Sig. 0.01</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>7.457</td>
<td>.007</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>.907</td>
<td>74.809</td>
</tr>
</tbody>
</table>
Dimension 2: Attitudes of parents, teachers and students regarding the use of MinecraftEdu in education

Dimension 2 presents an analysis of attitudes of the school community regarding the use of MinecraftEdu in classrooms. Below are detailed descriptive data (Table 4 and Figure 5), Kruskal-Wallis test data and Jonckheere-Terpstra test data (Table 4 and Table 5), which are significant for items 1, 3, 4 and 7.

The results of questionnaire 1 (APMA) indicate that most of the subjects in the sample thought that MinecraftEdu is fun (98.5 %): It enables discovery of new things (96.6%), it encourages learning about historical contents (97.1%), it enables rich interactions using virtual environments (96.6%), and it enhances creativity (96.1%) and learning (83.4%). Using game-based learning (71.7%) and exploiting the time in the classroom (63.9%) get lower results.

Attitudes of teachers, students and parents are very positive in general according to the results from the current descriptive analysis.

Table 4. Attitudes of parents, teachers and students toward working with the MinecraftEdu program (Descriptive analysis, Kruskal-Wallis and Jonckheere-Terpstra test)

<table>
<thead>
<tr>
<th>Attitudes of parents, teachers and students with regard to the application of MinecraftEdu (Item)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Kruskal-W. sign (0.01)</th>
<th>Jonckheere-T. sign (0.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. When you work with MinecraftEdu you learn in _______________________________________________</td>
<td>5.9</td>
<td>10.7</td>
<td>44.9</td>
<td>38.5</td>
<td>0.00*</td>
<td>0.00*</td>
</tr>
<tr>
<td>2. The historical contents of buildings with MinecraftEdu are suitable and interesting ______</td>
<td>1</td>
<td>2</td>
<td>28.3</td>
<td>68.8</td>
<td>0.28</td>
<td>0.15</td>
</tr>
<tr>
<td>3. Learning with this game is fun ___________________________________________________________________________</td>
<td>0.5</td>
<td>1</td>
<td>22.4</td>
<td>76.1</td>
<td>0.00*</td>
<td>0.00*</td>
</tr>
<tr>
<td>4. Working with MinecraftEdu exploits the time in ________________________________________________</td>
<td>10.2</td>
<td>25.9</td>
<td>27.8</td>
<td>36.1</td>
<td>0.00*</td>
<td>0.00*</td>
</tr>
</tbody>
</table>
5. Building in this environment develops creativity
6. In this virtual world we discover many new things
7. It is appropriate to use game-based learning in
8. Interacting with groups from other countries regarding creations in MinecraftEdu is positive

Note. 1 = Totally disagree; 2 = Disagree; 3 = Agree; 4 = Totally agree.

Figure 6. Dimension 2: Descriptive analysis of the attitudes of parents, teachers and students

When analyzing different groups, there are significant differences in some items between students and parents. The Kruskal-Wallis test and the Jonckheere-Terpstra test do not reflect significant differences in several items (Item 2, 5, 6 and 8), so parents, students and teachers have the same positive opinion on all these items.

Moreover, there are differences or discrepancies between these groups in items 1, 3, 4 and 7. When analyzing the means and Kruskal-Wallis rankings (Table 5), data show that students feel that MinecraftEdu is good for learning (Item 1), takes advantage of time in class (Item 4) and that it is appropriate to use game-based learning (Item 7), while teachers and especially parents show lower results in this regard due to low values in these groups. Although parents and teachers think that this application is fun (Item 3), students in the sample show significantly higher values in this regard.

Table 5. Kruskal-Wallis rankings

<table>
<thead>
<tr>
<th>Item</th>
<th>Community</th>
<th>Rank promedio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. When you work with MinecraftEdu you learn in class</td>
<td>Students</td>
<td>121.93</td>
</tr>
<tr>
<td></td>
<td>Parents</td>
<td>58.37</td>
</tr>
<tr>
<td></td>
<td>Teachers</td>
<td>95.80</td>
</tr>
<tr>
<td>2. The historical contents of buildings with MinecraftEdu are suitable and interesting</td>
<td>Students</td>
<td>99.85</td>
</tr>
<tr>
<td></td>
<td>Parents</td>
<td>105.05</td>
</tr>
<tr>
<td></td>
<td>Teachers</td>
<td>116.91</td>
</tr>
<tr>
<td>3. Learning with this game is fun</td>
<td>Students</td>
<td>115.37</td>
</tr>
<tr>
<td></td>
<td>Parents</td>
<td>80.88</td>
</tr>
<tr>
<td></td>
<td>Teachers</td>
<td>81.59</td>
</tr>
<tr>
<td>4. Working with MinecraftEdu exploits the time in class</td>
<td>Students</td>
<td>130.17</td>
</tr>
<tr>
<td></td>
<td>Parents</td>
<td>46.50</td>
</tr>
<tr>
<td></td>
<td>Teachers</td>
<td>74.77</td>
</tr>
<tr>
<td>5. Building in this environment develops creativity</td>
<td>Students</td>
<td>105.85</td>
</tr>
<tr>
<td></td>
<td>Parents</td>
<td>93.84</td>
</tr>
<tr>
<td></td>
<td>Teachers</td>
<td>107.68</td>
</tr>
<tr>
<td>6. In this virtual world we discover many new things</td>
<td>Students</td>
<td>108.76</td>
</tr>
</tbody>
</table>
7. It is appropriate to use game-based learning in school

<table>
<thead>
<tr>
<th></th>
<th>Parents</th>
<th>Teachers</th>
<th>Students</th>
<th>Parents</th>
<th>Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents</td>
<td>89.87</td>
<td>99.77</td>
<td>129.92</td>
<td>35.42</td>
<td>102.43</td>
</tr>
<tr>
<td>Teachers</td>
<td>89.77</td>
<td>99.77</td>
<td>129.92</td>
<td>35.42</td>
<td>102.43</td>
</tr>
</tbody>
</table>

8. Interacting with groups from other countries regarding creations in MinecraftEdu is positive

Mixed open-question questionnaire

When asked 9.-AB, Do you think MinecraftEdu should be used in school?, this gives a number of responses of interest analyzed by the program HyperResearch V 1.25. Participants responded by providing diverse opinions openly because the instructions in this open question invite to reasoned answers (Figure 7).

The main frequencies obtained in this section highlight the fact that the MinecraftEdu app is fun and can be used as a support tool in the classroom (79). Some individuals believe that it should be fully integrated (12) and that it is innovative (34). On the other hand, there are participants who believe that you lose time in the classroom applying it (24) and it should be applied outside the classroom (36). The most negative responses to this application are given by parents, with 79.1% of the frequencies relating to waste of time and 75% of the frequencies stating that MinecraftEdu should be applied outside the classroom.

Dimension 3: Analysis of VLE interactions regarding use of MinecraftEdu

In this dimension interactions are detailed in a group in the Edmodo platform (http://www.edmodo.com/?language=es) called Minecraft create and discover (Figures 7 and 8). There are students and teachers from Spain and the USA in the group, interacting with messages in English. These interactions have several advantages and benefits in the learning process (Sáez López, Leo, & Miyata, 2013).

The procedure used in this dimension is to analyze interactions in the Edmodo group called Minecraft create and discover from January 8, 2013 to March 27, 2013. We distinguish and classify the interactions of groups according to their purpose, quantifying and classifying messages using the application HyperResearch V 1.25. In Table 6 the messages are classified and quantified according to their purpose.
There is a very enriching interaction between students and teachers addressing topics of interest in the unit with the opportunity to discover and display creations (Figure 9):

Teacher (19/03/2013): Hello, this is the Pantheon. We learned about this today.
Student 1 (20/03/2013): The Pantheon is very old. The Romans built it and it is like a dome that is thicker...
Student 2 (20/03/2013): The Pantheon of Minecraft is very cool because it was built by the Romans...
Student 3 (20/03/2013): The Pantheon is one of the first domes ever built...
Student 4: (22/03/2013): This monument is very nice. I love it. The roof is made of a stone called pumice stone, it is volcanic...

Once presented analyzed results, it is possible to discuss and compare several outcomes by other authors related to impact of game-based learning approach in educational settings, and how games for learning can engage players and support learning and skills.
Regarding academic results, some studies assure that game-based learning improves significantly students’ outcomes (Cameron and Dwyer, 2005; Davidovitch, Parush and Shrub, 2008; Miller and Hegelheimer, 2006; Orvis, Horn, and Belanich, 2008; Yaman, Nerdel & Bayrhuber, 2008). Nevertheless, the present study describes in dimension 1 that there are not significant improvements in academic results after using MinecraftEdu. In this sense, the present research agree to Kirriemuir and McFarlane (2004) findings, which highlight that there are few examples of entertainment games being explicitly used in the classroom, due to difficulties in matching the entertainment to curricular outcomes.

Moreover, some researches note positive attitudes and advantages related to motivation and interaction when applying game-based learning in educational contexts (Barab, Dodge, Ingram-Goble, Pettyjohn, Peppler, Volk & Solomou, 2010; Blunt, 2007; Chen, Shih, & Ma, 2014; Eseryel et al., 2014; Hickey, Ingram-Goble, & Jameson, 2009; Russell, & Newton, 2008; Shaffer, 2007; Squire, 2006; Steinkuehler, 2006; Young et al., 2012). In the present study we agree that these authors given the information collected in dimension 2 and dimension 3.

Conclusions

The aim of this study is to analyze the pedagogical benefit of using the MinecraftEdu application in an educational context. We estimated the results obtained in relation to the evidence of learning, as well as interactions and attitudes of the school community to the pedagogical integration of this application.

From data triangulation the conclusions are as follows:

- We did not notice significant improvements in academic results after using MinecraftEdu from the tests applied and the resulting value of the Student t-test (Under the section Dimension 1: Evidence of learning with MinecraftEdu, Table 2 and 3)

- Most of the participants thought that MinecraftEdu enhances creativity (96.1%), improves learning (83.4%), is fun (98.5%), enables discovery (96.6%) and facilitates learning of historical content (97.1%) (Items 1, 2, 3, 5 and 6; under the section Dimension 2: Attitudes of parents, teachers and students regarding the use of MinecraftEdu in education, Figure 5).

- Interactions on Edmodo are very rich (96.6%) (Item 8, under section Dimension 3: Analysis of VLE interactions regarding use of MinecraftEdu, Table 6). In this context, students and teachers interact about topics of interest in the unit with the opportunity to comment on activities and display creations (Figures 7 and 8).

- 71.7% of the participants thought that applying game-based learning in class is appropriate for the learning process. There are discrepancies between students and parents in items 4 and 7. When checking means and Kruskal-Wallis rankings, it is noteworthy that students feel that MinecraftEdu takes advantage of time in class (Item 4) and that it is appropriate to apply game-based learning in education (Item 7), while teachers and especially parents show significantly lower results in these items (Kruskal-Wallis and Jonckheere-Terpstra test). These significant values highlighted clear differences of opinion between parents, who believe that this approach will waste time in the classroom, and students, who find it appropriate (Under the section Dimension 2: Attitudes of parents, teachers and students regarding the use of MinecraftEdu in education, Tables 4 and 5).

- There are also differences of opinion among students and parents about the possibilities of MinecraftEdu for learning in class (Items 1 & 3, Figure 3).

There were some important questions mentioned in the theoretical framework: What do students, teachers and parents think about this approach? Do video games enhance academic achievement? The present research gives some particular answers based on the intervention and collected data.

From the theoretical foundation it is shown that there are numerous serious and important studies that are widely considered to be very beneficial to the use of game-based learning (Barab et al., 2010; Blunt, 2007; Hickey, Ingram-Goble, & Jameson, 2009; Shaffer, 2007; Squire, 2006; Steinkuehler, 2006; Shute et al., 2009; Young et al., 2012).

When applying the unit “History and Architecture” there are no significant improvements regarding the academic results when applying MinecraftEdu in the classroom in this study (Under the section Dimension 1: Evidence of
learning with MinecraftEdu, Table 3). Moshirnia and Israel (2010) found similar results: no significant difference between the knowledge gained in the traditional expositional classroom using PowerPoint and the application of the unit with video games.

Some parents retain clearly negative attitudes and opinions towards the use of game-based learning in general and the use of MinecraftEdu in particular. Some of them also highlight that it is a waste of time in the classroom. They express (Figure 6) that it could be a good tool outside the classroom or occasionally in some subjects.

Nevertheless, there is a consensus and a majority agreement by the entire school community that recognizes the pedagogical benefits of MinecraftEdu (Table 4, Figure 5) due to several advantages that practically all participants highlighted: MinecraftEdu is an appropriate application for creating immersive activities with historical buildings and content; this approach enhances creativity, facilitates learning through discovery, is fun and provides interactive advantages using virtual learning environments.

Teachers’ attitudes are positive although moderate (Table 6). Students are in full agreement with this approach, mainly because of the fun and dynamic classes that allow them to be active protagonists who discover and develop contents and creativity in an immersive world.

References


Personal Learning Environments Acceptance Model: The Role of Need for Cognition, e-Learning Satisfaction and Students’ Perceptions

Salvador del Barrio-García1, José L. Arquero2 and Esteban Romero-Frías3*

1Universidad de Granada, Departamento de Comercialización e Investigación de Mercados, Granada, Spain //
2Universidad de Sevilla, Departamento de Contabilidad, Sevilla, Spain // 3Universidad de Granada, Departamento de Economía Financiera y Contabilidad, Granada, Spain // dbarrio@ugr.es // arquero@us.es // erf@ugr.es

*Corresponding author

(Submitted April 8, 2014; Revised September 30, 2014; Accepted November 11, 2014)

ABSTRACT

As long as students use Web 2.0 tools extensively for social purposes, there is an opportunity to improve students’ engagement in Higher Education by using these tools for academic purposes under a Personal Learning Environment approach (PLE 2.0). The success of these attempts depends upon the reactions and acceptance of users towards e-learning using Web 2.0. This paper aims to analyse the factors (e-learning satisfaction and students’ perceptions, among others) that determine the intention of use of a PLE 2.0 initiative. The study in addition analyses the moderating role of the Need for Cognition (NFC) in the model. The results indicate that the model proposed has a high explanatory power of the intention to use a PLE 2.0 and gives support to the moderating role of NFC. The study discusses how this analysis can help to improve course designs by teachers.

Keywords
Learning experience, Personal learning environment, Learning satisfaction, Need for cognition, Open education

Introduction

Information and Communication Technologies (ICTs), particularly Internet and mobile technologies, have been widely adopted by young generations for social purposes in Western countries, for instance, the USA (Pew Research Center, 2010) and Spain (AIMC, 2013). The so-called Web 2.0 or Social Web services play a paramount role in this adoption since they have surpassed the technical and economic barriers to create, share and distribute digital contents through a broad variety of devices (from smartphones to tablets and video-consoles).

The educational sector has reacted to these socio-technical changes by experimenting in the application of ICTs in education (Lee, 2010), resulting in an increased adoption of e-Learning platforms and, less frequently, Web 2.0 services. These services are claimed to be effective in connecting people and resources, facilitating interaction, fostering collaboration and active participation and aiding opportunities for critical thinking, among others (Romero-Frías & Arquero, 2013; Ajjaj & Hartshorne, 2008; Mason, 2006; Selwyn, 2007). The open and distributed nature of most of these services expedites environments for informal and emergent learning. Nevertheless, the complex scenario created requires the teachers’ creativity and flexibility to incorporate these novelties into formal settings.

Platforms specifically designed for e-Learning, such as Moodle or Blackboard, are more focused on institutional course design or instructor needs; whereas Personal Learning Environment (PLE) is an approach to integrate a consciously different sort of practices and resources (i.e., commonly used Web 2.0 services) to solve personal learning needs. It represents a more flexible approach focused on students’ needs (Attwell, 2007). From this idea, an educational experiment was designed to help students to develop their PLEs by using a set of tools and services that cover basic functions in their learning process. However, in order to evaluate the potential success of any educational design based on novel technologies, it is necessary to understand the users’ attitudes and their level of acceptance of this kind of technology for learning (Teo, 2010).

The Technology Acceptance Model (TAM, Davis, Bagozzi & Warshaw, 1989) has been widely used in education to evaluate Learning Management Systems (LMS) - technological systems that are generally based on closed environments specifically designed for learning. However, our PLE approach established on general use 2.0 tools represents a significant difference since it is designed to help students to develop autonomous and sustainable learning resources based on open interactions and personal needs. To our knowledge, there is no analysis about technology acceptance in this sort of open environments.
Therefore, this study aims to develop an extended TAM model for a PLE experience, integrating variables that could improve the predictive power of the model in this kind of experiences -learning satisfaction (Del Barrio, Romero-Frias & Arquero, 2013) and the perceived impact of the experience on key dimensions of the students’ learning process- and take intrinsic human factors into account (Sanchez-Franco, 2010), such as the students’ need for cognition (NFC; Cacioppo & Petty, 1982). We refer to this theoretical model design to open digital environments as Learning 2.0 system acceptance. By doing so, we intend to understand the factors that determine PLE acceptance in order to improve the design of this sort of proposals.

Theoretical background

According to the previous introduction, our theoretical background is focused on: (1) Personal Learning Environments and the Web 2.0 technologies, (2) the TAM model in education and (3) our extension of the model.

Educational literature: Personal learning environments and the Web 2.0 technologies

The pedagogical approach adopted in this educational experience is based on social constructivism (Brown & Adler, 2008; Sturm et al., 2009) and operationalised through Personal learning environments (PLE) (Attwell, 2007; Häkkinen & Hämäläinen, 2012; Rajagopal et al., 2012; Tu et al., 2012; Wilson et al., 2009). PLE represents a more flexible approach to use digital technology in education because it focuses on the students’ personal needs instead of institutional course designs or instructor needs, as generally occurs in Learning Management Systems (LMS), such as Moodle or Blackboard.

PLE is not established on a specific platform but on a set of functions that can be achieved through different tools according to user preferences. The development of a PLE using Web 2.0 is an extension of the social use that a majority of students already apply. Using these tools for learning allows students to experience the real world online and develop a set of services and competences that could be useful for personal and professional purposes and, remarkably, for lifelong learning (Romero-Frias & Arquero, 2013). Many studies have reported a positive impact of using 2.0 services in education (i.e., Richardson, 2009; Solomon & Schrum, 2007; Redecker et al., 2010), for example, in developing essential skills such as selecting relevant information, critically interpreting and analysing the socio-cultural context, working collaboratively and sharing knowledge.

This evidence indicates that an appropriate combination of 2.0 technologies and educational designs can provide a positive impact in developing key competences in education. However, which variables determine the acceptance of technology is still a question that has not been addressed in open and social contexts, such as those afforded by Web 2.0 services in a PLE approach. The next section refers to this question.

Technology acceptance model and education

Numerous studies have analysed the use and acceptance of technology in education since the publication in 1989 of the seminal work by Davis, Bagozzi and Warshaw, who proposed a Technology Acceptance Model where the variables Perceived Usefulness (PU) and Perceived Ease of Use (PEU) were used to predict ICTs acceptance.

PU captures the extent to which a potential adopter views the target technology as offering better value over alternative methods of carrying out the same task (Liu et al., 2009). PEU refers to the degree to which a potential adopter conceives the usage of the target technology to be relatively free of effort (Davis, Bagozzi & Warshaw, 1989).

As Park (2009) indicates, knowing the factors that influence students’ intentions and beliefs about e-learning could help academic managers to design better scenarios that favour the adoption of this type of learning approach (Grandon, Alshare & Kwan, 2005). In a context of massive open e-learning proposals (i.e., MOOCs), even policymakers could benefit from this research in order to avoid one-size-fits-all policies. The application of the TAM
in education has been shown to have good predictive validity (Sanchez-Franco, 2010) and there is a consistent body of research applying TAM, particularly to learning management systems (i.e., Martins & Kellermanns, 2004; Pituch & Lee, 2006; Liu, Liao & Pratt, 2009; Arteaga & Duarte, 2010; Sanchez-Franco 2010). There is no research, as far as we know, based on open general-use Web 2.0 tools integrated within the concept of PLE, although this approach is very extended in formal and informal learning.

Some papers using TAM in educational settings included additional variables to increase their predictive power: functionality, interactivity and response (Pituch and Lee, 2006); technical support (Ngai, Poon & Chang, 2007; Arteaga & Duarte, 2010); computer self-efficacy (Arteaga & Duarte, 2010); media richness (Liu, Liao & Pratt, 2009); flow (Liu, Liao & Pratt, 2009; Sanchez-Franco, 2010). Also, Lee et al. (2003) combined a TAM with Social Network Analysis to show that the students’ initial expectation affected the perceptions of and use of the system, but also that one student’s attitude change was influenced by other students’ changes.

**Proposed extended TAM**

Given the high satisfaction attained by users using Web 2.0 for social purposes, user satisfaction within a learning 2.0 context could help improving the TAM. Satisfaction is a key variable in explaining the usage of ICT (Doll & Torkzadeh, 1988; Hayashi, et al., 2004; Lin, Wu & Tsai, 2005), considered as a mediator in the processing of online information (Casaló, Flavián, & Guinalíu, 2008; Castañeda, Rodriguez & Luque, 2009).

In education the learning satisfaction concept can be defined as a student’s overall positive assessment of his or her learning experience (Keller, 1983). PU and PEU are found to be antecedents of learning satisfaction (Hui et al., 2008; Martin-Michieliot & Mendelsohn, 2000; Sun et al., 2008) and there is also a significant relationship between learning satisfaction and the intention to use e-Learning (Roca, Chiub & Martinez, 2006; Liaw & Huang, 2011). All these studies, supported by the recent study by Del Barrio, Romero-Frias and Arquero (2013) suggest that the use of learning satisfaction (eSAT) as a mediator in the TAM is more predictive than as an external variable.

Furthermore, Students’ perceptions about the attributes and quality of the course could be considered to have an impact on the PU of the whole system (Lee, 2006). Martins and Kellermanns (2004) suggest that when students perceive that using the system will have implications for their performance in a class, they will be likely to perceive the system as being useful. They will then show greater acceptance. It can thus be hypothesised that the awareness of the capabilities of the e-learning tool will be positively related to the students’ PU. Liaw, Huang and Chen (2007) highlighted the role of students’ perceptions on key aspects of the e-learning environment, such as effectiveness (defined as the impact on skills development) or active learning, on the attitudes towards its use. Similarly, the results by Selim (2007) point out that the improvements in collaborative and active aspects of e-learning are critical factors for success. Arquero and Romero-Frias (2013) reported a perceived positive impact of using Web 2.0 tools for educational purposes on content learning, skills development, and collaborative and active aspects of the process.

The literature on Psychology and Consumer Behaviour has revealed the convenience of considering intrinsic factors in explaining human reactions to any system. In this case, the Need for Cognition (NFC) is particularly relevant as a moderating variable, as the PLE approach depends on personal needs. NFC, as a concept, was introduced in 1955 by Cohen, Scotland and Wolfe, and describes the need to structure relevant situations in meaningful and integrated ways. Cacioppo and Petty (1982) used NFC in an investigation of differences among individuals in their tendency to engage in and enjoy thinking.

NFC has been extensively studied in a variety of social contexts (Cacioppo et al., 1996). On the one hand, High NFC individuals (cognisers) are characterised by their tendency to seek, acquire, think about, and reflect back on information to make sense of stimuli and events. These processes lead to the generating of a more stable cognitive change. On the other hand, low NFC individuals (cognitive misers) are more likely to rely on others, secondary stimuli or social comparison processes to make sense of information (Cacioppo et al., 1996, Evans, Kirby & Fabrigar, 2003). This leads to a more temporary, unstable or unpredictable cognitive change. Zhang and Buda (1999) pointed out that, in some way, cognitive efforts by students depend on their NFC.

Given that educational researchers are interested in how students learn and process information (Evans, Kirby & Fabrigar, 2003), some studies have used NFC in relation to learning systems (Evans, Kirby & Fabrigar, 2003; Chen
& Wu, 2012; Turner & Croucher, 2013; Kai-Wen, 2011). However, to our knowledge, no previous research has analysed the moderating role of NFC in the acceptance of Personal Learning Environments based on Web 2.0 services. Given that these services provide, on the one hand, autonomy to look for new information and, on the other hand, a social environment to interact and communicate with others, we consider that NFC is a relevant variable to understand how the technology acceptance works in this context.

According to the classical literature about TAM and the aforementioned additional variables, we propose a PLE 2.0 Acceptance Model (PLE 2.0 AM) (Figure 1) that integrates:

- eSAT as a mediating variable between perceived usability (PU & PEU) and Attitude Towards Use (ATU) and Behavioural Intention to Use of the learning 2.0 system (BIU),
- perceptions of students as external variables, and
- NFC as a moderating variable of the whole extended model proposed.

Methodology

Educational experience

As previously mentioned the educational design is based on the idea of the personal learning environment (PLE 2.0), whose aim is to help students to develop academic and professional uses of services (such as blogs, wikis, Social networks, etc.) that are generally employed for social purposes. By doing so, they could gain autonomy in their learning process and improve their competences for lifelong learning.

The services that integrate the PLE 2.0 experience included:

- A private Facebook group. Created to communicate and coordinate activities in the course in order to build a learning community where informal learning could emerge besides the contents of the course.
• A Twitter hashtag. Agreed on to share information and foster informal relations between the students and the teacher.
• A personal blog. Created to publish posts with the students’ critical opinions about the different topics in the course (30% of the final grade).
• Descuadrando.com (wiki platform). An open encyclopaedia about business where students created academic and professional style entries (20% grade).

Students could use their own profiles if they already had a presence in the different types of tools.
In addition to the former activities, the students had to do a final exam (50% grade).

Sample and data collection

The sample is composed of 203 students enrolled in a course on International Accounting (Business and Administration Degree) at the University of Granada (Spain). The composition of the sample, by gender is 31% male and 69% female. The students’ age range is from 20 to 43 years old (mean 23).

The data were gathered through a web-based questionnaire at the end of the course. Students were asked to provide sincere answers and confidentiality was assured. The inexistence of correct-incorrect responses was also highlighted, as was the fact that the data would only be used for research purposes.

Measures

Similarly to previous studies, TAM variables are measured by adapting to widely-used scales in educational settings. Thus, the Perceived Usefulness (PU) and Perceived Ease of Use (PEU) scales (3 and 4 items, respectively) were adapted from the Koufaris, Kambil and LaBarbera (2002) reduced versions of the original scales from Venkatesh and Davis (1996). The Attitude Towards Using (ATU) scale (3 items) was adapted from Chen, Gillenson and Sherrell (2002). The classic scale (4 items) by Zeithaml, Berry and Parasuraman (1996) was used to measure Behavioural Intention of Use (BIU). e-Learning Satisfaction (eSAT) was measured by using 2 items scale proposed by Szymanski and Hise (2000). This measure had been used extensively in previous studies (Szymanski & Henard, 2001; Evanschitzky et al., 2004; Jayawardhena, 2004). Finally, the Need for Cognition (NFC) was measured using the 5 items scale proposed by Del Barrio (2000); a reduced version of the larger original scales proposed by Cacioppo and Petty (1982) and Cacioppo, Petty and Kao (1984).

All previous variables were measured on a seven-point Likert scale, ranging from (1) strongly disagree to (7) strongly agree.

The students’ experience perceptions were measured through a questionnaire derived from the Assessment of University Teaching Activities Questionnaire (AUTAQ), developed by the Learning & Teaching Institute (ICE) of the University of Seville in order to evaluate key aspects of educational innovations. Although previously used, the first published version is found in Villar (1999). As Villar and Alegre (2006) highlight, the AUTAQ was designed to appraise students’ perceptions of their classroom environment and its design was guided by relationship, personal growth, and curriculum change dimensions for conceptualising university quality assurance. These questionnaires has been used to measure the impact perceived by students in evaluations of educational innovations (i.e., Arquero, Jiménez & Joyce, 2004; Lobo, Escobar & Arquero, 2009) and particularly in educational experiences using Web 2.0 tools (Arquero & Romero-Frias, 2013).

Results

The first step was to classify the respondents by their NFC score (mean of the items comprising the scale; Cronbach’s alpha = .727). Using the distribution of the NFC score, the students were assigned to three groups, excluding the central interval group for comparison purposes. The descriptive information of the resulting groups is shown in Table 1.
Table 1. Students groups according to NFC

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>t-test sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-NFC</td>
<td>82</td>
<td>2.7951</td>
<td>0.39249</td>
<td>.000</td>
</tr>
<tr>
<td>High-NFC</td>
<td>73</td>
<td>4.1644</td>
<td>0.36567</td>
<td></td>
</tr>
</tbody>
</table>

In general terms, the experience was positively valued by students in all the aspects included in Table 2, irrespective of their NFC level. The impact is particularly high in the improvement in the collaborative dimensions of learning (solving doubts with other students, sharing interests and ideas, etc.). There is a significant difference between groups ($t = -2.28, p < .05$) indicating that high NFC students perceived a greater improvement in communication skills than low NFC students.

Table 2. Perceived impact of students’ experience

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-NFC</td>
<td>3.53</td>
<td>0.927</td>
</tr>
<tr>
<td>High-NFC</td>
<td>3.67</td>
<td>0.958</td>
</tr>
<tr>
<td>Collaborative learning development</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-NFC</td>
<td>4.22</td>
<td>0.683</td>
</tr>
<tr>
<td>High-NFC</td>
<td>4.29</td>
<td>0.536</td>
</tr>
<tr>
<td>Relational collaborative learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-NFC</td>
<td>3.54</td>
<td>0.849</td>
</tr>
<tr>
<td>High-NFC</td>
<td>3.74</td>
<td>0.843</td>
</tr>
<tr>
<td>Communication skills</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-NFC</td>
<td>3.49</td>
<td>0.826</td>
</tr>
<tr>
<td>High-NFC</td>
<td>3.79</td>
<td>0.816</td>
</tr>
<tr>
<td>Content Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-NFC</td>
<td>3.54</td>
<td>0.849</td>
</tr>
<tr>
<td>High-NFC</td>
<td>3.74</td>
<td>0.843</td>
</tr>
</tbody>
</table>

Note. All the means are significantly higher than 3 (indifference) at 1% ($t$-test).

Students were also asked about the perceived ease of use of specific tools (ranging from 1, not easy at all to 5, very easy). The rating was Facebook group (4.6), blog (3.82) and wiki (3.27), with no significant differences between groups. When asked about the potential usefulness of these tools for their professional careers and lifelong learning (1, not useful at all – 5, very useful), the results were very consistent: Facebook group (3.69), blog (3.68) and wiki (3.65).

In order to test the theoretical model proposed, Partial Least Squares (PLS) path modelling was used (Chin, 1998a) by implementing a multigroup analysis (high-NFC versus low-NFC). PLS has been used extensively as a data analysis method in the literature. It is particularly indicated for “predictive” purposes and theory development (Anderson & Gerbin, 1998) and when the sample size is reduced (Chin, 1998a).

PLS was the most appropriate analytic method given the characteristics of the study: (1) the model in Figure 1 is formed of 10 latent variables measured through 27 observed variables, and (2) the sample is made up of 73 and 82 students, in the high and low NFC groups, respectively. The software package used was SmartPLS 2.0 (Ringle, Wende & Will, 2005).

Table 3 shows the adequate psychometric properties of the scales. After bootstrapping, all the loadings were significant ($p < .05$) and the values of composite reliability (CR) and average variance extracted (AVE) were above the acceptable cut-off level (0.8 and 0.5, respectively). Finally, discriminant validity was tested for each group, by implementing the method suggested by Fornell and Larcker (1981).

Table 3. Analysis of the psychometric properties of the scales (outer model)

<table>
<thead>
<tr>
<th>Latent variables</th>
<th>Indicators</th>
<th>Loadings</th>
<th>CR</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relational collaborative learning</td>
<td>REL_COL1</td>
<td>.898</td>
<td>.908</td>
<td>.907</td>
</tr>
<tr>
<td></td>
<td>REL_COL2</td>
<td>.923</td>
<td>.950</td>
<td></td>
</tr>
<tr>
<td>Collaborative learning development</td>
<td>COL_DEV1</td>
<td>.803</td>
<td>.847</td>
<td>.797</td>
</tr>
<tr>
<td></td>
<td>COL_DEV2</td>
<td>.807</td>
<td>.785</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COL_DEV3</td>
<td>.642</td>
<td>.547</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2 shows the results of estimating the structural model (inner model) for both groups, after applying bootstrapping. To test for group differences we applied PLS-MGA approach (Henseler, Ringle & Sinkovics, 2009), as this method accounts for the distribution-free assumption of the data (see Table 4).
Table 4. Multigroup invariance analysis results (PLS-MGA)

<table>
<thead>
<tr>
<th>Source of causation</th>
<th>High-NFC</th>
<th>Low-NFC</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>REL_COL→PU</td>
<td>.055</td>
<td>.032</td>
<td>.267</td>
</tr>
<tr>
<td>COL_DEV→PU</td>
<td>.081</td>
<td>.166</td>
<td>.274</td>
</tr>
<tr>
<td>CON-LEARN→PU</td>
<td>.225</td>
<td>.339</td>
<td>.252</td>
</tr>
<tr>
<td>ACT_LEARN→PU</td>
<td>.213</td>
<td>.095</td>
<td>.268</td>
</tr>
<tr>
<td>COM_SK→PU</td>
<td>.429</td>
<td>.181</td>
<td>.036</td>
</tr>
<tr>
<td>PEU→PU</td>
<td>.114</td>
<td>.216</td>
<td>.217</td>
</tr>
<tr>
<td>PEU→ATU</td>
<td>.515</td>
<td>.456</td>
<td>.334</td>
</tr>
<tr>
<td>PU→eSAT</td>
<td>.555</td>
<td>.206</td>
<td>.001</td>
</tr>
<tr>
<td>PU→ATU</td>
<td>.168</td>
<td>.335</td>
<td>.155</td>
</tr>
<tr>
<td>PEU→BIU</td>
<td>.308</td>
<td>.397</td>
<td>.305</td>
</tr>
<tr>
<td>ATU→eSAT</td>
<td>.310</td>
<td>.696</td>
<td>.001</td>
</tr>
<tr>
<td>eSAT→BIU</td>
<td>.484</td>
<td>.275</td>
<td>.115</td>
</tr>
<tr>
<td>ATU→BIU</td>
<td>.111</td>
<td>.231</td>
<td>.783</td>
</tr>
</tbody>
</table>

Regarding the impact of students’ learning experiences on PU, only perceived impact of content learning ($\beta_{CON\_LEARN\rightarrow PU\_HNFC} = .225; \beta_{CON\_LEARN\rightarrow PU\_LNFC} = .339$) and improvement in communication skills ($\beta_{COM\_SK\rightarrow PU\_HNFC} = .429; \beta_{COM\_SK\rightarrow PU\_LNFC} = .181$) present a significant and positive influence on PU for both groups. However, it is worth noting that the perceived improvement in communication skills is significantly greater for High NFC students ($p = .035$). No difference exists between both groups regarding content learning.

The effect of relational collaborative learning on PU is not significant for any of the groups ($\beta_{REL\_COL\rightarrow PU\_HNFC} = .055; \beta_{REL\_COL\rightarrow PU\_LNFC} = .032$).

Collaborative learning development only has a significant positive effect on PU for Low NFC students ($\beta_{COL\_DEV\rightarrow PU\_HNFC} = .081; \beta_{COL\_DEV\rightarrow PU\_LNFC} = .166$), whereas Active learning only has a significant positive effect for High NFC ($\beta_{ACT\_LEARN\rightarrow UP\_HNFC} = .213; \beta_{ACT\_LEARN\rightarrow UP\_LNFC} = .095$).

To sum up, we highlight that the main antecedent of Perceived Usefulness of the Learning 2.0 methodology is content learning for Low NFC students and the communication experience for High NFC students.

Regarding the mediating role of eSAT, the results in Table 4 are indicative of the significant moderating role of NFC between BIU and the PU variables.

The satisfaction of students with the Learning 2.0 model proposed is positively influenced by their perceptions about the usefulness of the system ($\beta_{PU\rightarrow eSAT\_HNFC} = .555; \beta_{PU\rightarrow eSAT\_LNFC} = .206$). This relation is remarkably higher for the High NFC students ($p < .05$). On the contrary, the effect of Attitude towards Use (ATU) of PLE 2.0 on eSAT was significantly higher ($p > .05$) for Low NFC Students ($\beta_{ATU\rightarrow eSAT\_HNFC} = .310; \beta_{ATU\rightarrow eSAT\_LNFC} = .696$). Secondly, the Attitude towards Use (ATU) of PLE 2.0 is more dependent on PEU than on PU for both groups. The effect of PEU on ATU is more intense for High NFC students than for Low NFC students ($\beta_{PEU\rightarrow ATU\_HNFC} = .515; \beta_{PEU\rightarrow ATU\_LNFC} = .456$). Regarding the relation PU→ATU, the stronger effect corresponds to the Low NFC group ($\beta_{PU\rightarrow ATU\_HNFC} = .168; \beta_{PU\rightarrow ATU\_LNFC} = .335$). However, these differences are not significant ($p > .05$).

Finally, the intention of use (BIU) of the PLE 2.0 proposal is significant and determined positively by PU, eSAT and ATU. The most important variable explaining BIU is eSAT ($\beta_{eSAT\rightarrow BIU\_HNFC} = .484; \beta_{eSAT\rightarrow BIU\_LNFC} = .308; \beta_{ATU\rightarrow BIU\_HNFC} = .111$) for High NFC students and PU ($\beta_{PU\rightarrow BIU\_HNFC} = .379; \beta_{ATU\rightarrow BIU\_LNFC} = .275; \beta_{PU\rightarrow BIU\_LNFC} = .231$) for Low NFC students, although the differences are not significant ($p > .05$).

Table 5 presents the standardised total effects (direct and indirect) on BIU. For High-NFC students, the most relevant variable explaining BIU is PU (.620) followed by eSAT (.484) and PeU presents the lowest effect. For Low-NFC students the most relevant variable explaining BIU is PU (.577), followed by ATU (.422), PEU (.317), and, finally, eSAT (.274).
Table 5. Standardised total effects on BIU

<table>
<thead>
<tr>
<th></th>
<th>High-NFC</th>
<th>Low-NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PU→BIU</td>
<td>.620</td>
<td>.577</td>
</tr>
<tr>
<td>PEU→BIU</td>
<td>.205</td>
<td>.317</td>
</tr>
<tr>
<td>eSAT→BIU</td>
<td>.484</td>
<td>.274</td>
</tr>
<tr>
<td>ATU→BIU</td>
<td>.260</td>
<td>.422</td>
</tr>
</tbody>
</table>

To conclude the analysis of the results, the explicative and predictive power of the model was tested using the percentage of variance explained of the dependent constructs ($R^2$) and predictive relevance (Stone-Geisser-Q2) through a blindfolding procedure (Chin, 1998b). According to Hair et al., (2011) the percentages of variance explained of the five dependent variables in the model and the predictive relevance were adequate (Figure 2).

**Discussion, implications and future research**

The aim of this study was to develop an extended TAM model for an open PLE experience. The results suggest that the theoretical model proposed in this study has adequate predictive power to understand the future intention of use of a PLE based on Web 2.0 tools. The pedagogical approach adopted is innovative because it is not based on a closed virtual learning environment (such as those of most LMS) but on a personal and open approach that is intended to adapt the available technological tools to the particular needs and goals of each student for autonomous and lifelong learning (Attwell, 2007; Wilson et al., 2009).

The causal relationships between the constructs postulated by the structural model are well supported. The mediating role of Satisfaction between Perceived Usefulness and Attitude towards the system was confirmed in an educational context, supporting the evidence found in other settings (i.e., Szymanski & Hise, 2000; Casaló, Flavián & Guinaliu, 2008). The effect of the perceived impact of the experience in some key dimensions of learning (Collaborative learning, Content Learning, Active Learning, Communication skills) on PU is significant and positive in most cases. There are two aspects that impact on PU more significantly for both groups: improvement in content learning and development of communication skills; however, the most valued impact was collaborative learning development. These results suggest that further exploration is needed concerning the factors that students take into account to form their perceptions of the PU for these experiences. A better knowledge of the impact of these factors on the intention to use a PLE 2.0 can help teachers to underline the importance of some characteristics in the first weeks of the course in order to improve educational performance.

Regarding the impact of intrinsic human factors in the acceptance of this experience, the results support the moderating role of students’ Need for Cognition. For High-NFC students, the acceptance depends mainly on the perceived usefulness and satisfaction towards the system; whereas, for Low-NFC students, the Perceived Ease of Use plays a more relevant role. Furthermore, the impact of eSAT is also moderated by NFC. This has a more relevant role for High-NFC students.

The moderating effect of NFC has several implications. The intention to use IT-related educational innovations could be affected by specific students’ characteristics, having an impact in the institutionalisation and the transferability of these innovations. This highlights the need for further research to understand which characteristics play a relevant role in acceptance and which actions could be implemented to avoid undesired effects.

Students’ NFC profiles could be measured at the beginning of the course, in order to plan actions focusing on the relevant variables. For groups composed of High NFC students (for whom learning satisfaction and usefulness are more relevant), it is possible to propose more complex activities and tools as long as they generate satisfaction and are perceived as useful. For low NFC students training activities and support could be relevant to improve their engagement.

The relevant role of the perceived ease of use (key for Low NFC students, but also important for High NFC students) supports the idea that the adoption of technology in education is facilitated when the tools proposed were already used by students for other purposes and were perceived as accessible. Although the use of tools which are not specifically designed for education in this PLE 2.0 experience requires more creativity and diminishes the instructor’s control, it is noteworthy that some of these tools were already used on a daily basis (78.4% of students...
use Facebook every day) and are considered by students to have a high potential (3.7 out of 5) for their future professional careers and lifelong learning. In contrast, LMS usage is restricted to formal education settings (Mott, 2010). Therefore, through the PLE 2.0 approach we are contributing to the development of a more sustainable learning environment for students. This is in line with the objectives set by the European Union (2006).

The results are considered exploratory. Students participating in this study were enrolled in an elective subject at the last courses of a business degree. Further research should be done with larger samples including different degrees, compulsory subjects and entry level students.

Future research should look at objective performance measures and not only at students’ perceptions in order to discover if there is an improvement after using technology depending on the NFC level. Also, following some studies (i.e., Gu, Zhu & Guo 2013), we would like to compare differences between teachers and students in using these learning methodologies. Finally, the application of Technology Acceptance Models to Massive Open Online Courses (MOOC) could contribute critically to the adoption of these initiatives worldwide, as long as MOOCs are a reference in terms of open education becoming part of the PLE of students and professionals.

References


Flipping a College Calculus Course: A Case Study

Alpaslan Sahin1*, Baki Cavlazoglu2 and Yunus E. Zeytuncu2
1Department of Teaching, Learning and Culture, Texas A&M University, College Station, Texas, USA
// 2Department of Mathematics & Statistics, University of Michigan-Dearborn, Dearborn, Michigan, USA
// sahinalpaslan38@gmail.com // baki42@gmail.com // yezeytuncu@gmail.com
*Corresponding author

(Submitted April 24, 2014; Revised December 2, 2014; Accepted December 4, 2014)

ABSTRACT
As online videos have become more easily available and more attractive to the new generation of students, and as new student-learning approaches tend to have more technology integration, the flipped classroom model has become very popular. The purpose of this study was to understand college students’ views on flipped courses and investigate how the flipping affects their achievement in mathematics. We also studied how college students prepared for flipped classroom sections. Finally, college students’ views were analyzed to see what they think about flipping in terms of benefits and preparation. Participants were 96 college students consisting of mostly freshmen & sophomores. We utilized descriptive statistics and paired t-test to analyze the data. Descriptive statistics revealed that participants preferred watching flip class videos (44%) over reading the sections from the textbook (17%) for preparation. Dependent t-test results showed that there is a statistically significant difference between students’ average quiz scores from non-flipped sections and flipped sections. Students achieved significantly higher quiz scores in flipped sections than non-flipped ones. Overall, most of the students (83%) stated that flipped-taught lessons prepared them better.

Keywords
Flipped classroom, College calculus, Blended learning, Technology, Online video

Introduction
Today, we live in a technology and media driven environment with rapid changes in technology tools (Partnership for 21st Century Skills, 2011). With these rapid technological changes, teaching has become a more complex and challenging profession (Moore, 2005; Good & Brophy, 2008). The art of teaching in today’s classrooms has also changed dramatically because students’ learning approaches are different than those of the previous generation (Brunsell & Horejsi, 2013). In addition, various forms of technology have been integrated into learning and teaching environments. Accordingly, teachers need to have a large repertoire of skills and the ability to apply them in different situations for success (Moore, 2005). A teacher should be able to “design, implement, and assess learning experiences to engage students and improve learning; enrich professional practices; and provide positive models for students, colleagues, and the community” (International Society for Technology in Education [ISTE], 2008, p. 1) and also use his or her “knowledge of subject matter, teaching and learning, and technology to facilitate experiences that advance student learning, creativity, and innovation in both face-to-face and virtual environments” (ISTE, 2008, p. 1) for effective teaching in today’s learning environments.

As technology has increasingly been used in teaching and learning to meet 21st century learning goals (Crippen & Archambault, 2012), integrating appropriate learning opportunities, including the use of technology-enhanced teaching strategies to improve students’ learning, has become indispensable (ISTE, 2008). To accomplish this, teachers should learn necessary technological developments and use them appropriately in their teaching. If related technology is integrated into teaching appropriately, students develop more positive attitudes towards the learning subject and become more successful (Moore, 2005).

Effective teachers have to master a variety of teaching techniques and strategies to establish and maintain student interest (Moore, 2005). In today’s learning environments, there are many different teaching strategies that have the advantage of technology incorporation. For example, blended learning is one of the most common technology-supported teaching strategies. In this approach, students first learn the content and instruction through online delivery, with the students controlling the time, place, path, and/or pace (Staker & Horn, 2012). Second, as Staker and Horn (2012, p. 3) pointed out, “students have the opportunity to be involved in teacher supervision and/or instruction in a specific location away from home.” There are several types of blended learning models (e.g., Staker & Horn, 2012). Among these models, the flipped classroom model is the most commonly used; however, research on
its implementation is limited. Therefore, examining the flipped classroom model in different classrooms is indispensable. This study aims to investigate the effect of a flipped classroom in college calculus students' preparation and achievement for both flipped- and non-flipped classes.

**Theoretical framework**

A theoretical framework for this study is derived from the literature on (a) a general description of blended learning and the flipped classroom model, (b) benefits of flipped classrooms for students and teachers, (c) other advantages of flipped classrooms, and (d) flipped classrooms in different contexts, levels, and subjects.

**Blended learning and the flipped classroom model**

The flipped classroom model is one of the most commonly used blended learning models. In this model, students have the opportunity to learn new content by online delivery and instruction from a remote location (Staker & Horn, 2012), and in the classroom students do their homework or assignments instead of listening to lectures, and therefore there is more scaffolding, student-teacher interaction, and opportunities for student collaboration in the face-to-face class meetings. The content can be delivered via reading materials, videos, online presentations, or any combination of different delivery methods. However, since available videos appear more appealing to the new generation of students, many educators prefer using online videos instead of reading materials when preparing students out-of-classroom for in-class active learning (Herreid, 2013).

Literature on flipped classrooms document four blended learning models (Staker & Horn, 2012); (1) rotation model, (2) flex model, (3) self-blended model, and (4) enriched-virtual model. The rotation model has four sub-categories as well, including (a) the station-rotation model, (b) the lab-rotation model, (c) the flipped classroom model, and (d) the individual-rotation model (Staker & Horn, 2012). Of these four, the flipped-classroom model is the most frequently used one, in which “students rotate on a fixed schedule between face-to-face teacher-guided practice (or projects) on campus during the standard school day and online delivery of content and instruction of the same subject from a remote location (often home) after school” (Staker & Horn, 2012, p. 10) and/or before school (Herreid, 2013) for a specific course or subject (e.g., math). This variedly used approach has many advantages for 21st century generation of students, as well as for teachers, administrators, and parents.

**Benefits of the flipped classroom for students and teachers**

Current studies report that the flipped classroom model has many benefits for students and teachers. For instance, one study reported that students are able to make progress at their own pace (Fulton, 2012). Also, doing homework in class provides great opportunity for teachers to see students’ difficulties and learning styles. This lets teachers easily modify lessons according to students’ needs within this model. Additionally, classroom time can be used more efficiently and creatively with a flipped classroom because students have the opportunity to familiarize themselves with the content before class meetings.

Fulton (2012) found that use of the flipped model increased student achievement, interest, and engagement as well as provided many opportunities for using technology effectively for 21st century learning. Pearson Partners on Flipped Learning (2013) reported a study on using the flipped classroom model for at-risk students in a high school in Michigan. In the studied school, student-passing rates were low in many subjects, including English language, mathematics, science, and social studies. To solve the problem, school administration decided to use the flipped classroom model. The administration provided teacher support, collaborative learning opportunities, and instructional technology access prior to actual flipped classroom implementation. As the flipped classroom model was used throughout the first semester, the student-passing rates, test scores, and college interest dramatically increased; the students’ engagement improved; and discipline problems decreased (Pearson Partners on Flipped Learning, 2013).

In another recent study, two hundred science teachers implemented the flipped classroom model into their STEM classrooms and found it had a big impact on student learning (Herreid, 2013). In their implementation, they first
provided reading materials and videos about subjects prior to classroom meetings and asked for a pre-assignment about the subjects. This approach helped teachers better adjust the flow of their lessons to address student needs (Fulton, 2012; Herreid, 2013). Herreid (2013) reported that teachers emphasized several benefits of the flipped classroom model, such as finding more time to spend with students on research, engaging students with scientific equipment in the classroom, providing alternative learning opportunities (e.g., students who miss classes are able to watch missed lectures on their own), and creating a conducive atmosphere for students to actively engage and enjoy the learning process.

On the other hand, the teachers also pointed out possible negative qualities of the flipped classroom. For example, teachers stated that some students who are new to this approach could be resistant because of their background in traditional learning approaches. Moreover, the materials for the flipped classroom should be in good quality and they should match with students’ level. Furthermore, preparing a good quality video can be very time consuming for teachers and some teachers can be resistant because of their lack of experience with the necessary technology. Therefore, using flipped classroom model may require some additional preparation time compared to the traditional classroom teaching model for both teachers and students.

Other benefits of flipped classroom

In addition to many prevalent benefits of the flipped classroom, there are other indirect advantages as well. Brunsell and Horejsi (2013) detected other benefits of the flipped classroom model including increases in parent involvement, administrative support, and teacher job satisfaction in addition to higher student success and engagement. By observing online lectures, parents were able to see what was going on in the class and the quality of content. In addition, Brunsell and Horejsi stated that the availability of flipped classroom videos allowed administrators to see how courses are taught in terms of content and quality in their schools. Additionally, the videos helped the administrators establish trust and accountability with their teachers. The study demonstrated that teachers who tried the flipped classroom model indicated improved job satisfaction. Seventy percent of their students increased their standardized test scores, and 80% of the students developed better attitudes when they were taught via flipped classroom model (Brunsell & Horejsi, 2013).

The flipped classroom in different contexts, levels, and subjects

The flipped classroom approach has been implemented in different ways within varied contexts. For example, Ruddick (2012) conducted a project-based flipped classroom study in which students in a college preparatory chemistry course used video lectures and PowerPoint materials at home to prepare for class and used class time for problem-solving activities. In her study, Ruddick compared the final exam scores and overall course grades of the standard lecture-based approach and the flipped classroom model. She found that students from flipped classrooms outperformed those who were taught through traditional lectures. Moreover, students from the flipped classroom showed positive attitudes towards learning chemistry and became more interested in chemistry. For a freshman undergraduate calculus course, Kay and Kletskin (2012) developed a series of 59 problem-based podcasts in which they covered some important concepts including operations with functions, solving equations, linear functions, exponential and logarithmic functions, and trigonometric functions. They posted all podcasts to the course website for a 21-day time span and tracked the number of students’ video podcast visits. Their results illustrated that approximately 67% of the undergraduate students visited the video podcasts and they stated that this method was useful, easy to follow, and the video podcasts helped their understanding of new concepts.

The dramatic increase in the use of technology means that teaching is no longer as limited to traditional lectures settings, but in some subjects (e.g., mathematics, physics, calculus, etc.) the traditional lecture habit continues despite the increasing number of preparation materials available for a flipped classroom model. Also, the number of studies on the use of the flipped classroom model has varied from middle school level through college level. There are many current studies on flipped classrooms in middle and high school settings (e.g., Fulton, 2012; Herreid, 2013; Pearson Partners on Flipped Learning, 2013; Ruddick, 2012) and for college level subjects such as chemistry and physics (e.g., Brunsell & Horejsi, 2013; Zappe et al., 2009). Yet until very recently there has been a lack of research on the
flipped classroom in college level calculus teaching (Kay & Kletskin, 2012). This lack is notable because calculus is one of the core subjects that is necessary for the understanding of many other subjects (e.g., physics, chemistry, engineering, etc.) that are especially important for those planning to proceed in STEM-related fields (National Research Council, 2011). Therefore, a close attention to the effectiveness of the flipped classroom at the undergraduate level of calculus teaching is vital.

These are the research questions we sought answers to:
- How do college students prepare for a calculus class in a flipped classroom?
- How do college students’ achievements differ by instruction types (traditional versus flipped classroom)?
- What are college students’ perceptions about the flipped classroom model?

Method

Participants

The sample was comprised of 96 college students (due to a non-strict attendance policy, the actual number of students varied by day) who took Math 152-Engineering Mathematics II course during the Spring 2013 semester in a southeast Texas college. Student gender demographics consisted of 79 male and 17 female students. Most participants were freshman (60) and sophomore (34) with only one student from junior and senior grades. The course consisted of ten subjects/sections throughout the semester, and students were provided with flipped classroom model opportunities in three sections out of ten. In other words, three sections were flipped and seven sections were traditionally taught during the semester, and flipped and non-flipped (traditional) sections were compared.

Instrument

We developed two different short surveys; one to investigate how sections/topics were taught with the flipped classroom model, the other to determine how the flipped classroom model affected college calculus students’ preparation, understanding, and their performance. The first survey was given three times at the start of each flipped classroom section and aimed to determine how the students prepared for their classes, what types of preparation they did before they came to class, and how the flipped classroom videos for that specific section helped them. The second survey was given at the end of the semester to examine participants’ overall thoughts about the use of the flipped classroom model in their classes. More specifically, the second survey was designed to determine how much and to what extent the flipped class videos helped the students. Additionally, a pop quiz, designed to assess the content of the section taught with videos, was given in class meetings.

Procedure

Calculus is one of the standard subjects for many college students. In any regular calculus course, students are assumed to be familiar with certain mathematical skills from high school. However, in practice, a lack of preparation was observed quite frequently. Consequently, many calculus instructors use class time at different junctures in order to review these background materials. In an ideal setup, students are expected to fill in this gap by doing the recommended preparation before lectures. The most common and suggested form of preparation appears as reading relevant sections from a textbook before class. Unfortunately, this is not a common practice anymore among new generation students. As observed in our surveys, about half of the students in our study did not do any preparation before class. Only 17% of students read parts of the textbook. Similar studies in the literature (Hoeft, 2012) have also concluded similar lack of reading before class.

For this particular calculus course, we decided to use the flipped classroom model to improve students’ achievement through changes in the way they prepare before attending class. Not surprisingly, many students found watching a 10 minute video easier than reading the textbook. With this in mind, we prepared three videos for students watch before coming to class at different times during the semester. In the videos we included some basic background material and
a short introduction to the lecture. We informed students in advance about the videos and tried to provide them ample time to watch the videos. Instead of campus-based video venues, we used YouTube as a platform. This was easier for students to access from many different devices and locations. We administered surveys and pop quizzes after flipped classroom implementations to gauge students’ involvement.

Analysis

We did descriptive analyses to answer the first question. For the second question, we ran a dependent t-test to see how students’ quiz scores differed in a class taught with and without the flipped classroom model. We provided descriptive statistics to disclose college students’ perceptions about the flipped classroom and how it helped their preparation for their classes.

Results

For the first question, descriptive statistics revealed the types of preparation students did before coming to the sections that utilized flipped classroom. Table 1 summarizes the frequency of students’ preparation preferences. From these statistics, it is clear that students preferred videos for preparation more often than reading related preparations. In each survey, student answers showed that they preferred watching flipped classroom videos to reading textbooks for class preparation except in Time-2, where 39% of participants indicated that they did not do any preparation for the class (see Table 1). In each survey, students were told to check all that apply for survey question No. 2, so the numbers of answers added up to more than the actual number of survey takers. In Time-1, 76 students (the change in the number is due to optional attendance) completed the survey, and of those, 41% watched the flipped class video whereas only 11% preferred reading the course textbook for preparation. In Time-2, out of 62 participants, 24% watched the flipped classroom video whereas only 14% read the textbook for the class preparation. Likewise, student responses in Time-3 revealed similar patterns of preparation in which a sizeable percentage of the class (38%) watched the video prepared for the class. Again, only 15% of students chose reading the textbook section instead of watching the video. Overall, participants preferred watching flipped classroom videos (35%) over reading the sections of a related textbook (13%) for preparation.

<table>
<thead>
<tr>
<th>Times</th>
<th>Reading the textbook</th>
<th>Watching the flipped video</th>
<th>Reading other texts</th>
<th>Watching other videos</th>
<th>Didn’t do any prep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-1 (76)</td>
<td>11 (11%)</td>
<td>41 (41%)</td>
<td>23 (23%)</td>
<td>14 (14%)</td>
<td>11 (11%)</td>
</tr>
<tr>
<td>Time-2 (62)</td>
<td>10 (14%)</td>
<td>18 (24%)</td>
<td>9 (12%)</td>
<td>8 (11%)</td>
<td>29 (39%)</td>
</tr>
<tr>
<td>Time-3 (67)</td>
<td>13 (15%)</td>
<td>33 (38%)</td>
<td>14 (16%)</td>
<td>10 (11%)</td>
<td>18 (20%)</td>
</tr>
<tr>
<td>Overall</td>
<td>34 (13%)</td>
<td>92 (35%)</td>
<td>46 (18%)</td>
<td>32 (12%)</td>
<td>58 (22%)</td>
</tr>
</tbody>
</table>

Note. 1 denotes the number of survey takers.

To answer the second research question, we ran a dependent t-test to test our null hypothesis in which we hypothesized that students’ quiz score means were the same regardless of teaching styles (non-flipped versus flipped). After running the t-test, we rejected the null hypothesis and accepted the alternative hypothesis stating that students’ quiz scores were different. Our analysis showed that students’ average quiz scores from flipped classroom sections were significantly higher (t(94) = 3.502, p = .001) than students’ quiz scores from non-flipped sections. Students achieved significantly higher quiz scores in flipped classroom sections (M = 8.32, SD = 1.36) than non-flipped-taught sections (M = 7.54, SD = 1.69). We also calculated Cohen’s d effect size and came up with d as -0.51.

To answer the third question, we analyzed the second survey questions separately and provided frequencies for each part of the question to have a better understanding of what students thought about flipped classroom and how they benefitted from those classes. For question No. 3 of Survey 2, the majority of the participating students (83%) stated that flipped classroom sections prepared them better for their classes throughout the semester (see Table 2).
Table 2. Flipped classroom and students’ preparation

<table>
<thead>
<tr>
<th>The video lessons helped me prepare more for the sections throughout the semester.</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly agree</td>
<td>19</td>
<td>26%</td>
</tr>
<tr>
<td>Agree</td>
<td>41</td>
<td>57%</td>
</tr>
<tr>
<td>Neither agree nor disagree</td>
<td>11</td>
<td>15%</td>
</tr>
<tr>
<td>Disagree</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

Students’ responses to the question No. 4 showed that 85% of students stated that the video lessons of flipped classroom helped them perform better during the class (see Table 3). Only 15% of students were not sure about whether flipped classroom was helpful or not. No students indicated any negative thoughts about flipped classroom video lessons.

Table 3. Students’ overall thoughts about how flipped classroom helped their performance

<table>
<thead>
<tr>
<th>Overall, the video lessons helped me perform better.</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly agree</td>
<td>15</td>
<td>21%</td>
</tr>
<tr>
<td>Agree</td>
<td>46</td>
<td>64%</td>
</tr>
<tr>
<td>Neither agree nor disagree</td>
<td>11</td>
<td>15%</td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

To question No. 6 in Survey 2, the majority of students (81%) indicated that flipped classroom videos helped them feel more confident. The rest of the participants were neutral about the benefits of flipped classroom in terms of providing confidence (see Table 4).

Table 4. Flipped classroom and students’ confidence

<table>
<thead>
<tr>
<th>The video lessons helped me feel more confident during those sections.</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly agree</td>
<td>18</td>
<td>25%</td>
</tr>
<tr>
<td>Agree</td>
<td>40</td>
<td>56%</td>
</tr>
<tr>
<td>Neither agree nor disagree</td>
<td>14</td>
<td>19%</td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

In the Survey 2, question No. 5 was about how the students usually prepared for the class during sections when there were no flipped classroom videos. Students’ answers to question No. 5 were similar to the question in Survey 1 regarding how they prepared for the classroom when it was flipped. Seventeen percent of students indicated that they read the section from the textbook. Another 17% of students reported that they also consulted other textbooks. Finding videos related to the section was the most popular choice among the participating college calculus students (25%). Forty-one percent of students reported that they did not do any preparation for the class (see Table 5).

Table 5. Students’ preparation types for classes

<table>
<thead>
<tr>
<th>How do you usually prepare for a class when there is no video lesson?</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>I read the section from the textbook</td>
<td>12</td>
<td>17%</td>
</tr>
<tr>
<td>I read other supplemental textbooks (including online resources)</td>
<td>12</td>
<td>17%</td>
</tr>
<tr>
<td>I try to find videos about the section</td>
<td>18</td>
<td>25%</td>
</tr>
<tr>
<td>I do no preparation</td>
<td>29</td>
<td>41%</td>
</tr>
</tbody>
</table>

Table 6. Students’ preference about course types

<table>
<thead>
<tr>
<th>I’d prefer a course with video lessons over a course without video lessons.</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly agree</td>
<td>37</td>
<td>51%</td>
</tr>
<tr>
<td>Agree</td>
<td>23</td>
<td>32%</td>
</tr>
<tr>
<td>Neither agree nor disagree</td>
<td>11</td>
<td>15%</td>
</tr>
<tr>
<td>Disagree</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>
Students’ responses to question No. 7 of Survey 2 showed that 83% of students preferred a course with video lessons to a course without videos (see Table 6).

**Discussion and conclusion**

The flipped classroom model has become very popular in today’s classroom, especially in the United States. In this study, we aimed to investigate how integration of the flipped classroom model into a college calculus class affected students’ class preparation, learning, and achievement. Concurring with other studies on flipped classrooms (Brunsell & Horejsi, 2013; Herreid, 2013; Kay & Kletskin, 2012; Pearson Partners on Flipped Learning, 2013; Ruddick, 2012; Zappe et al., 2009), we found that students in the flipped classroom preferred watching flipped classroom videos more often than reading sections from the textbook. This study also revealed that the flipped classroom model resulted in an increase in student achievement in calculus. Our findings also indicated several additional benefits of the flipped classroom model, such as changes in preparation habits before attending class, improved levels of understanding, and higher levels of self-efficacy during the lectures. Moreover, we found that students were already using videos to study and prepare even when they did not use flipped classroom material in their calculus course. In this study, the students preferred flipped class videos over other course materials for the class preparation. This is congruent with previous research (Herreid, 2013; Kay & Kletskin, 2012; Zappe et al., 2009). Last but not least, we learned that students were eager to choose a course with videos over a course without videos within the flipped classroom model.

Students’ responses showed they much preferred and benefited from the flipped classroom model in which flipped class videos as part of their calculus class were available. It was interesting to see that almost three times as many students in this particular college calculus class preferred watching flipped classroom videos (35%) than reading the section of the textbook (13%), even though reading calculus text is still the most common means of studying calculus among college students. One way to explain why students prefer watching videos to other preparation methods is due to its flexibility. Students have an option to watch videos whenever they want without feeling any pressures of time or instructor observation, as well as no worry about being embarrassed in front of the class (Brunsell & Horejsi, 2013; Zappe et al., 2009). Videos are also convenient and useful for students who miss some classes (Pearson Partners on Flipped Learning, 2013). Additionally, students may have more time and opportunities to interact with their teachers and peers (Herreid, 2013; Ruddick, 2012). Thus, it is clear that watching videos as a means of preparation in flipped classroom model is more convenient, useful, easy to access, and enjoyable for students.

The second research question revealed that students’ calculus achievement increased when they were taught within the flipped classroom model. Students’ quiz results in flipped classroom sections indicated that flipped classroom experiences with video preparation resulted in better achievement in calculus compared to traditionally taught sections. This finding is parallel with the previous research, which found similar results in other subject areas, including chemistry, physics, and engineering (e.g., Brunsell & Horejsi, 2013; Fulton, 2012; Herreid, 2013; Pearson Partners on Flipped Learning, 2013; Ruddick, 2012). This is because within the flipped classroom model students were better able to prepare for the class meetings and had more opportunities to interact with the instructor and peers than during traditional lectures.

In addition, students in flipped classroom sections showed higher preparation efforts. The results revealed that 22% of students in flipped classrooms did not do any preparation, while 41% of students in traditional sections did not do any preparation. This result documented that students in a flipped classroom have higher motivation to prepare for class. Also, findings from the third research question supported this assertion because students stated that flipped classrooms prepared them better for the class and helped them learn better. Therefore, students with higher motivation for preparation had higher scores in flipped classroom sections than traditional sections.

The student responses showed that the flipped classroom model prepared them better, helped them learn better, and enabled them to feel more confident in learning calculus. Again, these findings can be seen to stem from the fact that students in the flipped classroom model have more freedom and flexibility to choose their preparation methods for the class (Fulton, 2012). They do not have the anxiety of missing or not understanding the material (Ruddick, 2012) since they have opportunities to re-watch the videos when they miss something (Brunsell & Horejsi, 2013; Zappe et al., 2009) and ask questions in face-to-face class meetings with their teachers or peers. They can watch videos before and after class as many times as they want until they learn the materials, without being exposed to any external
pressure or stress (Ruddick, 2012). Students may have felt more confident and ready for the class and therefore became more successful with the flipped classroom teaching model (Pearson Partners on Flipped Learning, 2013).

The answer to question No. 5 (about when the flipped classroom model is not used) showed that almost half of the students did not do any preparation for their classes. Other research supports this finding. Small (2005) found that students do not go to college to learn but to socialize, and Arum and Roksa (2011) reported that students do not show significant improvement in the key measures of critical thinking, complex reasoning and writing by the end of sophomore year because students do not study, write, or read much within traditional teaching methods. But interestingly, the same student group in a non-flipped class found and watched Internet videos related to the subject even though they were not required to do so. Therefore, it would be reasonable to say that using the flipped classroom model with a video preparation method may result in better preparation, less anxiety, and increased student achievement when teaching college calculus to new generation students.

Limitations of the study

This research study has several limitations. First, the sample of study could have been more representative of all college level students in terms of grade level and gender distribution: The sample mainly consisted of freshman and sophomore college students, and the gender distribution was 82% male and 18% female. Second, students were taught with flipped classroom only 3 times out of 10 sections, and their success in the 3 flipped and 7 traditional sections were analyzed to see whether there was a statistical difference. Intervention of more sections or an improved study design (i.e., a true experimental design with separate control and intervention groups) could be more beneficial in reflecting the effect of flipped classrooms on student achievements in college calculus learning. Third, the students did not have identification numbers to track each individual student’s changes throughout the semester. Some advanced statistical methods (i.e., Repeated Measures ANOVA) could allow seeing more details about how each participant’s achievement, preparation type, and other attitudes change through the semester. Lastly, in addition to using the survey method, acquiring additional qualitative data could give greater insight into participants’ thoughts and opinions about their experiences. For example, personal interviews could be vital in getting more information regarding participants’ preferences of preparation methods, their thoughts on how flipped classroom helped their performance, and the benefits of the flipped classroom in college calculus teaching.

Implications

As observed in this study and similar ones in the literature, the flipped classroom model is becoming an effective and common teaching model in different subject areas. New generation students find the flipped classroom model useful, enjoyable, and appropriate in learning different content subjects, including calculus. Many students indicate that watching videos for class preparation is easier and more enjoyable than reading the related section from the textbook in a flipped classroom. Thus, we strongly recommend that teachers incorporate videos in their teaching. Accordingly, teachers also enjoy teaching when students show up in class with better preparation. Students seek these resources on online platforms even when the teachers do not provide them directly. Therefore it is essential for teachers to get familiar with these resources and produce their own online tools so that students can navigate the course material more effectively.

Also, the flipped classroom model proves to be effective in courses that require a certain amount of background knowledge. The online resources and videos can be effectively used to remind students of topics that they are expected to recall from previous courses. This can be especially useful in a calculus course where students can be reminded of required information before class. This will save ample time for the actual topic of the lecture. Therefore, integrating videos into courses would provide students with an invaluable learning opportunity that has fewer time and social pressure limitations.

References


Appendix

Survey 1

1. How much did you prepare for the most recent section?
2. What did you do to prepare for this class? (Check that all apply)
   a. I read the section from the textbook
   b. I watched the video made for the section
   c. I read some other texts about this section
   d. I watched some other videos about this section
   e. I couldn’t do any preparation for the class.

Survey 2

In this survey, we would like to learn about your ideas on the online video lessons that were made for flipped classrooms in this semester.
Thank you for taking the time to complete this survey.

1. Did you watch the videos before the class?
   a. I watched all of them
   b. I watched 4 of them
   c. I watched 3 of them
   d. I watched 2 of them
   e. I watched 1 of them
   f. I did not watch any of them before the class.

2. Did you watch the videos after the class?
   a. I watched all of them
   b. I watched 4 of them
   c. I watched 3 of them
   d. I watched 2 of them
   e. I watched 1 of them
   f. I did not watch any of them before the class.

3. The video lessons helped me prepare more for the sections throughout the semester.
   a. Strongly agree
   b. Agree
   c. Neither agree nor disagree
   d. Disagree
   e. Strongly disagree

4. Overall, the video lessons helped me perform better.
   a. Strongly agree
   b. Agree
   c. Neither agree nor disagree
   d. Disagree
   e. Strongly disagree

5. How do you usually prepare for a class when there is no video lesson?
   a. I read the section from the textbook
   b. I read other supplemental textbooks (including online resources)
   c. I try to find videos about the section
   d. I do no preparation
6. The video lessons helped me feel more confident during those sections.
   a. Strongly agree
   b. Agree
   c. Neither agree nor disagree
   d. Disagree
   e. Strongly disagree

7. I’d prefer a course with video lessons over a course without video lessons
   a. Strongly agree
   b. Agree
   c. Neither agree nor disagree
   d. Disagree
   e. Strongly disagree
Effects of Cueing by a Pedagogical Agent in an Instructional Animation: A Cognitive Load Approach

Hsin I. Yung1* and Fred Paas2

1Graduate Institute of Technological & Vocational Education, National Taipei University of Technology, Taiwan // 2Institute of Psychology, Erasmus University Rotterdam, Rotterdam, The Netherlands // hcy103@mail.ntut.edu.tw // paas@fsw.eur.nl

*Corresponding author

(Submitted May 19, 2014; Revised November 27, 2014; Accepted December 12, 2014)

ABSTRACT

This study investigated the effects of a pedagogical agent that cued relevant information in a story-based instructional animation on the cardiovascular system. Based on cognitive load theory, it was expected that the experimental condition with the pedagogical agent would facilitate students to distinguish between relevant and irrelevant information, and therefore lead to higher learning outcomes and lower cognitive load than the control condition without a pedagogical agent. Based on 133 seventh-grade students as participants, the results confirmed the hypotheses, indicating that the pedagogical agent had a more positive effect on learning outcomes and resulted in a more favorable relationship between learning outcomes and cognitive load. The results are discussed in terms of theoretical and practical implications for designing instructional animation.

Keywords

Pedagogical agent, Cueing, Cognitive load

Introduction

In recent years, numerous studies have explored the effects of instructional animations on learning (for reviews see, Höfller & Leutner, 2007; Tversky, Morrison, & Bétrancourt, 2002). Animation is assumed to be particularly effective in the learning of complex dynamic systems, such as can be found in the domains of science and biology. Animations provide immersive external representations and are assumed to be effective for portraying visual changes of concepts, presenting implicit knowledge, and facilitating comprehension and learning (Scase & Rogers, 1996; Tversky et al., 2002). However, animations can only be effective for learning if they are designed in such a way that they engage learners in processing its relevant parts and understanding the relations between those parts. Many animations contain both relevant and irrelevant parts, which may lead to learners spending part of their limited cognitive resources on processing irrelevant parts, leaving less cognitive capacity to process the relevant parts (Ayres & Paas, 2007a, 2007b).

Tversky et al., (2002) have argued that animations may fail to improve learning because they are too complex or too fast to be accurately perceived. In addition, learners’ attention is automatically allocated to perceptually salient elements in an animation, for example, elements with a sudden onset/offset, elements that move or elements that change in color. Because these elements are not necessarily the most relevant elements in the animation, attentional resources can be wasted (Ayres & Paas, 2007a, 2007b; Tversky et al., 2002). Under these circumstances, researchers are challenged to find instructional techniques that support learners in focusing their cognitive resources on aspects of the animation that are relevant to the learning goal.

Ayres and Paas (2007a, 2007b) have argued that animations are often less effective than static pictures, because of their intrinsic transient nature and poor design features, which impose high extraneous cognitive loads. One design feature that imposes a high extraneous load is related to the split-attention effect, which materializes when learners have to mentally integrate two physically separated sources of information, such as a diagram and text, before the learning task can be understood (e.g., Ayres & Sweller, 2005; Chandler & Sweller, 1992; Mayer & Moreno, 2003). However, the main reason for high extraneous load in animations (e.g., Ayres & Paas, 2007a, 2007b; Hegarty, 2004; Lewalter, 2003) is that learners are required to process current information, remember this information when it disappears and mentally integrate it with newly appearing information, and, when available, with information previously stored in long-term memory. Consequently, a substantial part of the limited working memory resources is focused on dealing with the demands of the presentational format, rather than on learning. Static visualizations do not have these problems, because information is continuously available, which prevents
learners from having to hold the information in their working memory. This also explains why some instructional manipulations that reduce transient information in animations and the associated demands on working memory, such as some forms of user-interactivity (for an overview see, Wouters, Paas, & van Merriënboer, 2008), segmentation (for an overview see, Spanjers, van Gog, & van Merriënboer, 2010), attention cueing (for an overview see, de Koning et al., 2009), and adding a human movement component (Paas & Sweller, 2012; Wong et al., 2009) improve the effectiveness of animations.

**Literature review**

**The pedagogical agent with cueing**

One way to provide instructional support in computer-based learning environments is by using an animated pedagogical agent (Atkinson, 2002; Baylor, 2009; Moreno, Mayer, Spires, & Lester, 2001). Pedagogical agents are human-like onscreen characters that provide hints and feedback as well as direct learners’ attention by using gesture, gaze, speech, or combinations of those modalities. In recent years, researchers have been focusing on the internal and external properties of pedagogical agents (Mayer, 2005; Moreno, 2007; Wouters et al., 2008). The internal properties of pedagogical agents relate to instructional strategies that promote learners’ cognitive engagement and reduce extraneous cognitive load, including providing verbal explanations, feedback or guidance. The external properties of pedagogical agents relate to their social and motivational features to encourage the learning process (Mayer & DaPra, 2012; Wang et al., 2008).

The research on pedagogical agents has revealed a general positive effect on affective and cognitive variables. For example Moreno et al. (2001) compared a discovery-based learning environment containing an animated pedagogical agent (i.e., Herman the Bug) to a text-based version of the same environment and found higher levels of transfer performance, motivation and interest for the pedagogical agent condition. Atkinson (2002) found positive effects on transfer performance of a pedagogical agent that was incorporated in a computer-based learning environment designed to teach learners how to solve word problems. He concluded that an animated pedagogical agent can help optimize learning from examples. Lusk and Atkinson (2007) compared different levels of embodiment of a pedagogical agent in a worked example based learning environment on how to solve multi-step proportional word problems. They showed that the learning environment with the embodied agent resulted in better transfer performance than a voice-only condition without an agent.

One way of directing learners’ attention that has been found effective for learning from animations is cueing. Cognitive load theory provides a strong framework for the effect of attention cueing in terms of learners’ limited cognitive resources (Sweller, 2010; Sweller, van Merriënboer, & Paas, 1998). Ayres and Paas (2007a) have proposed that attention cueing can be an effective method to reduce extraneous cognitive load by focusing learners’ attention to essential information in the learning task. According to Lowe and Boucheix (2011) cues can be categorized as internal or external. Whereas internal cues are embedded within the text or graphic, such as color, external cues, such as a verbal statement, are separated from the targeted information. Internal cues have greater effects to raise its visual contrast and attract learners’ attention (Boucheix & Lowe, 2010). Several studies that have been conducted in the context of cognitive load theory, have shown positive effects of cues that focus learners’ attention on the information that is relevant at a certain point in time on learning from animations (e.g., Amadieu, Mariné, & Laimay, 2011; Boucheix & Lowe, 2010; Mayer & Moreno, 2003). Cueing has been designed in a variety of visual and verbal forms including color coding (e.g., Kalyuga, Chandler, & Sweller, 1999) and picture labeling (e.g., Florax & Ploetzner, 2010) to guide learners in discriminating between relevant and irrelevant information. De Koning, Tabbers, Rikers, and Paas (2009) have classified three functions for attention cueing in learning. Firstly, a selective function to support learners in discriminating between relevant and irrelevant information. Secondly, an organizational function to support learners in indentifying structure and organization in the learning materials. Thirdly, an integrative function is to emphasize relations between information elements in the learning materials.

Previous research on the use of pedagogical agents in multimedia learning has focused on user experience (Dehn & van Mulken, 2000), motivational effects (Baylor, 2009) and embodiment effects (Mayer & DaPra, 2012; Moreno, Mayer, Spires, & Lester, 2001). Only a few studies on pedagogical agents have focused on the pedagogical agent’s support of learner’s cognitive functions in learning. In addition, previous research has only considered affective measures or learning performance tests to measure the effects of the pedagogical agent. In this study we investigated
the effects of cueing by a pedagogical agent on learning performance, cognitive load, and a composite measure of instructional efficiency.

**Research questions and hypotheses**

In this study students studied an animation about the blood circulation in the human heart with or without a pedagogical agent. It was investigated whether students’ learning from an animation with a pedagogical agent that supports the learner’s distinction between relevant and irrelevant aspects through cueing is more efficient than learning from an animation without a pedagogical agent. It was argued that the pedagogical agent with cueing would facilitate learners to identify relevant information, which would reduce extraneous cognitive load and leave more working memory resources to construct a cognitive schema (i.e., intrinsic cognitive load). Therefore, it was hypothesized that the participants in the experimental condition with the pedagogical agent would show a more favorable relationship between cognitive load during the learning phase and performance during the testing phase (i.e., higher instructional efficiency: Paas & van Merriënboer, 1993) than the participants in the condition without the pedagogical agent.

**Method**

**Participants and design**

Participants were 133 seventh grade students (67 boys, 66 girls) with an average age of 12 years from a large junior high school at Taipei, Taiwan. All participants were volunteers. Participants were randomly assigned to two experimental conditions, in such a way participants with pedagogical agent condition, participants without pedagogical agent condition. The dependent variables were comprehension test performance, cognitive load and instructional efficiency. The experimental group received animation with the pedagogical agent while the control group received the same animation without the pedagogical agent.

**Materials**

The animation of the cardiovascular system was developed by a team consisting of an animation expert, an instructional designer, and a content expert. The content of the animation on the cardiovascular system was segmented into 23 separate screens and was adapted from a middle school biology textbook, which was written into a story titled “the adventure of the red blood cell.” It was expected that the use of a story format would stimulate students’ engagement with the materials. The animation story was based on three characters: red blood cell, white blood cell and the pedagogical agent. The pedagogical agent was developed using Maya and Adobe after effect (AE) software and was designed in the form of an instructor. The pedagogical agent was capable of exhibiting gestures such as waiving and pointing words in order to focus learners’ attention to important concepts. The story line of the animation was written in collaboration with two biology teachers of the school in which the experiment took place. The animation portrayed the process of the heart expanding when it fills with blood, the circulation of the blood through the heart, and the contraction of the heart.

The story began with the red blood cell starting a journey and making friends with a white blood cell. While they traveled around the four chambers of the heart, the pedagogical agent joined the two characters, and told them important concepts of the heart and guided them to find the right path to finish their journey. The pedagogical agent, which was positioned to the right of the “adventure of the red blood cell” animation, directed the learner’s attention by deictic gestures, such as moving his hand for highlighting relevant information (see Figure 1a, 1b), and providing information, such as telling the red blood cell how to distinguish the difference between atrium and ventricle. It was expected that the visual cueing could guide learners’ visual attention to actively engage in knowledge construction and decrease the possibility that learners pay attention to irrelevant elements. During the adventure, the pedagogical agent acted like an instructor and explained fundamental concepts in the animation. The attention cueing strategy was expected to support learners to select essential information. Our main assumption was that applying the pedagogical agent with a red and a white blood cell in a story format would enhance learners’ understanding of the dynamic heart circulation, rather than just stimulate them to memorize concepts. It should be noted that the
The instructional content of the animation and the presented fundamental concepts were identical in the experimental and control condition. The only aspect that differed was the presence of the pedagogical agent in the experimental condition.

Instruments

Comprehension test

The comprehension test evaluated participants’ understanding of the heart blood circulation. It consisted of 10 multiple-choice questions and asked participants to indicate the correct answer within four options. The questions were developed based on important concepts of the heart circulation system and process of the pulmonary circulation. An example question of the comprehension test, “please indicate whether the following option has the correct path of pulmonary circulation?”

Figure 1a. Example of the pedagogical agent with explanations

Figure 1b. Example of explaining the main function of the red blood cell
Participants could receive one point for each correct answer. For each correct answer, the participants received one point, otherwise they received zero points. A maximum of 10 points could be obtained on the comprehension test. Cronbach’s alpha for the comprehension test was 0.61. All comprehension questions were evaluated by three biology instructors with 10 years of teaching experience.

Cognitive load measures

The present study tested the effects of a pedagogical agent with cueing on cognitive load. Measurement of cognitive load was consisted two types of subjective measures: perceived task difficulty and perceived amount of invested mental effort (Paas, 1992). The five-point Likert scales that were used for measuring perceived task difficulty and effort ranged from very low to very high was adapted from Paas (1992). Using self-ratings of mental effort and perceived difficulty would provide a reliable and valid indication of cognitive load. The perceived task difficulty scale had a Cronbach’s alpha of .81. The perceived mental effort scale had a Cronbach’s alpha of .80.

According to Paas and van Merriënboer (1993, 1994), the perceived amount of invested mental effort and the perceived task difficulty can be considered as indices of cognitive load. Although the individual measures of cognitive load, such as the perceived amount of invested mental effort, can be considered important to determine the power of different instructional conditions. Paas and van Merriënboer (1993) have argued that a meaningful interpretation of a certain level of cognitive load can only be given in the context of its associated performance level and vice versa. They developed a computational approach to combine measures of mental effort with measures of the associated primary task performance to compare the instructional efficiency of instructional conditions. Since then, a whole range of studies has successfully applied this method or an alternative method combining effort and performance (for an overview see, Paas et al., 2003). In this study, we used ratings of task difficulty as a measure of cognitive load and ratings of mental effort in combination with performance measures to calculate instructional efficiency.

Procedure

Participants were randomly assigned to either experimental condition (n = 64) or control condition (n = 69). The experiment was run in a group session with 35 participants in the schools’ computer classrooms. Each student worked individually at a computer. At the start of the experimental session, participants were provided with headphones and asked not to talk to the other participants. Subsequently each participant had about 5 minutes to work on the tutorial before the start of the learning phase, which had a maximum duration of 40 minutes. They were instructed to watch the animation carefully in order to comprehend the material. After the experiment, participants were asked to rate the perceived cognitive load on a rating scale. Subsequently, they received the comprehension test without having access to the animation. The whole experiment lasted approximately 45-50 minutes.

Results

Analyses of variance (ANOVA) were performed for the mean scores of each dependent variable. Table 1 shows the mean scores and corresponding standard deviations of the comprehension test scores, cognitive load, and the instructional efficiency scores for students in the experimental and control condition.

| Table 1. Means and standard deviations of comprehension performance and cognitive load for students in the experimental and control condition |
|-------------------------------------------------|-----------------|-----------------|
|                                                | Experimental group (N = 64) | Control group (N = 69) |
|                                                | M       | SD  | M    | SD |
| Comprehension                                  | 6.67    | 1.28 | 6.01 | 1.20 |
| Cognitive load                                 | 3.02    | 0.94 | 3.03 | 0.88 |
| Instructional efficiency                       | 0.24    | 1.20 | -0.24| 0.94 |

An ANOVA for comprehension test performance revealed a significant effect of the pedagogical agent, $F(1, 129) = 8.772, p = .004$, In line with our hypothesis, the pedagogical agent with cueing condition had a positive effect on
learning. However, the ANOVA for the mean ratings of task difficulty and mental effort did not reveal a significant difference between the conditions, $F(1, 129) = 0.665, p = .416$.

**Instructional efficiency**

Paas and van Merriënboer (1993) developed instructional efficiency formula by combining mental effort and learning performance. Instructional efficiency was calculated using learning performance and mental effort. By standardizing participants’ mental effort and learning performance scores to $z$-scores, this formula can reveal important information on the efficiency of instructional conditions that is not necessarily reflected by each of those measures alone. Consistent with the original method, we used ratings of mental effort in combination with performance measures to calculate instructional efficiency.

$$E = \frac{Z_{\text{Performance}} - Z_{\text{Mental Effort}}}{\sqrt{2}}$$

The result showed that there was a main effect of the pedagogical agent on instructional efficiency, $F(1, 129) = 5.53, p = .020$. It indicated that the participants in the pedagogical agent conditions exhibited a more favorable relationship between the amount of mental effort invested in the learning phase and the performance in the test phase.

**Conclusion and discussion**

This study investigated the effect of a pedagogical agent with cueing on students’ learning performance, cognitive load, and instructional efficiency. It was hypothesized that directing learners’ attention to the relevant aspects of the task by the pedagogical agent would decrease extraneous cognitive load, improve learning, and lead to a more favorable relationship between learning effort and test performance (i.e., instructional efficiency). The results indicated that cueing by the pedagogical agent had a positive effect on learning performance and instructional efficiency. No significant differences in cognitive load were found. This finding might have been caused by the fact that only the total cognitive load was measured, which makes it impossible to draw conclusions on the different types of cognitive load, i.e., intrinsic, extraneous and germane load that underlie the overall load (Paas et al., 2003). More specifically, based on the main effects found for pedagogical agent it can be assumed that the cognitive capacity that was freed by the pedagogical agent (by decreasing extraneous load) was used for activities that fostered learning (i.e., imposed a germane load) reduction in extraneous load. Unfortunately, there are currently no methods available for measuring the specific types of load that constitute total load (but see, Leppink, Paas, van der Vleuten, van Gog, & van Merriënboer, 2013).

The results regarding learning and learning efficiency indicated that the cueing of the pedagogical agent was strong enough to direct learners’ attention in such a way that they could acquire an adequate mental representation the blood circulation in the human heart. In general, learning from instructional animation is a challenge for learners because they need to extract relevant information from transient information to construct adequate mental representations. Using cueing is considered critical in the instructional design of animation to guide the learners’ attention and enhance their cognitive processing of the crucial elements in animation (De Koning et al., 2009). According to Fischer and Schwan (2010), if cueing only underlines specific parts of an animation and unced parts of the animation are still required for integrating the conceptual structure of the materials, students’ learning improvement may be limited. Because, the blood circulation in the human heart is a dynamic system, which consists of a range of factors with causal relationships, the cueing of the pedagogical agent in the current study might not have been optimal. Future research is needed to investigate which aspects of the dynamic system can best be cued.

It is clear that the positive effects of the pedagogical agent on comprehension test performance and instructional efficiency cannot be generalized based on the results of this study. In the present study a very specific story-based format of a pedagogical agent was used. Although, such a format might be attractive for the young population that was used in this study, it is not clear whether adolescents and adults would show the same effects. In addition, it is not clear whether other formats of a cueing pedagogical agent would lead to similar results in a similar population. It is also not clear whether this specific type of pedagogical agent would work in other domains. Therefore, future
research is needed to investigate the effectiveness of different story-based, cueing pedagogical agents in different populations and in different domains.

In the present study, we only focused on the attention guiding function of cueing. Therefore, it would be interesting for future studies to explore other functions of cueing, such as the organizational function, and their effects on cognitive load and learning. From a practical point of view, the study highlighted the positive effect of the pedagogical agent with cueing as an effective instructional strategy for reducing the complexity of animation. Implementing the cueing effect in the instructional design for learning from animation of the complex dynamic system may support learners in discriminating between relevant and irrelevant information in order to construct a mental representation of the system. Although the study is not unique in terms of investigating effects of cueing, it presents a new approach to the instructional design of animation in biology domain, based on the theoretical framework of cognitive load. In conclusion, using cueing with the pedagogical support can be an effective strategy to enhance learning from animation.

References


Collaborative Professional Development of Mentor Teachers and Pre-Service Teachers in Relation to Technology Integration

Shih-Hsiung Liu1*, Hsien-Chang Tsai2 and Yu-Ting Huang3

1Center for Teacher Education, National Changhua University of Education, Taiwan // 2Department of Biology, National Changhua University of Education, Taiwan // 3Department of English, Changhua National Changhua University of Education, Taiwan // shsiung@cc.ncue.edu.tw // bihft@cc.ncue.edu.tw // happyalice8211@gmail.com

*Corresponding author

(Submitted July 7, 2014; Revised October 26, 2014; Accepted December 8, 2014)

ABSTRACT
This study examines the collaborative professional development (CPD) of three pairs of mentor teachers and pre-service teachers in a junior high school. Of particular focus is the integration of technology into instruction, by using technological pedagogical and content knowledge (TPACK) to evaluate professional development. A qualitative research method based on classroom observations and focus group interviews (FGIs) is adopted. Additionally, data obtained from instructional plans, video-recorded classroom observations and FGIs are analyzed using a constant-comparison analysis method. Analysis results indicate that mentor teachers adjust their instruction methods when they receive the support of pre-service teachers specifically by moving from presenting technological content knowledge (TCK) to constructing various TCK bases. The pre-service teachers constantly apply TCK and technological pedagogical knowledge (TPK) to develop professionally, especially in TPACK-related technology integration concepts. Notably, the CPD program benefits pre-service teachers more than mentor teachers because the former actively seek more opportunities to apply TPACK than the latter, who simply exchange digital instructional materials.

Keywords
Collaborative learning, Learning communities, Teacher professional development, Technology integration

Introduction

Integrating technology into classroom instruction can increase student motivation, learning efficacy, curiosity and creativity (Carle, Jaffee & Miller, 2009; Idris & Nor, 2010; Molins-Ruano, Sevilla, Santini, Haya, Rodriguez & Sacha, 2014). However, teachers frequently use technology to perform non-instructional tasks such as grading and monitoring attendance (Gray, Thomas, & Lewis, 2010; Russell, Bebell, O’Dwyer & O’Connor, 2003). According to Govender and Govender (2014), most teachers with access to technology and computer competency skills fail to incorporate technology in their teaching. Unsuccessful experiences in technology adoption in the classroom may inhibit teacher motivation (Slaouti & Barton, 2007), explaining the need to create successful classroom experiences of technology integration.

Akbaba-Altun (2006) verified that in-service training courses lack hands-on activities, and fail to prepare teachers adequately to integrate technology. Enabling teachers to integrate technology is more complex than simply delivering instructions and technology-related skills (Ferdig, 2006). Despite possessing valuable teaching experiences or proficiency in technological skills, in-service teachers may still fail to successfully apply technology in the classroom. Related studies have demonstrated that observing successful teachers enhances the professional development of teachers, increases the likelihood that they will adopt new technologies in their classrooms, and also improve their teaching methods and contents (Anderson, Barksdale, & Hite, 2005; Powell & Napoliello, 2005). Successfully collaborating with colleagues is essential to effectively integrating technology in the classroom (Tondeur, Kershaw, Vanderlinde & van Braak, 2013).

Although extremely familiar with related technologies, pre-service teachers tend to lack sufficient pedagogical skills. Although teacher education programs include various courses that provide pre-service teachers with technology integration-related knowledge, pre-service teachers lack opportunities to apply such knowledge. Related studies have shown that the design of teacher education courses varies considerably (Lee & Lee, 2014). Such courses also fail to help teachers to integrate technology into the classroom (Goktas, Yildirim & Yildirim, 2008).
According to Russell et al. (2003), pre-service teachers express more confidence and proficiency in computer use than experienced teachers do. In-service teachers with sufficient teaching experience tend to have fixed teaching philosophies, yet lack technological skills. However, effectively integrating technology with teaching involves both technological skills and teaching experience. Ideally, the ability of in-service and pre-service teachers to learn collaboratively would allow them to effectively integrate technology with classroom instruction.

Related studies have suggested that the professional development of pre-service teachers in technology integration depends on knowledgeable mentors and sufficient technology access for practice and curriculum development (Grove, Strudler & Odell, 2004). While demonstrating the effectiveness of pre-service and in-service teachers collaborating with each other (Chen, 2012), other studies have identified the advantages of collaboration as including mutual learning and providing professional support (Goodnough, Osmond, Dibbon, Glassman & Stevens, 2009; Spilková, 2001). However, the collaborative professional development (CPD) of mentor teachers and pre-service teachers in a specific school in relation to technology integration has seldom been examined. Based on a case study involving a teaching team consisting of three mentor teachers and three pre-service teachers, this study examines the CPD of mentors and pre-service teachers in terms of their ability to integrate technology with classroom instruction.

Literature review

Technology integration by TPACK constructs

Teachers often fail to successfully apply technology in the classroom owing to factors such as lessons and objectives, classroom strategies for implementing technology, as well as hardware and software use (Liu, 2011). Mishra and Koehler (2006) expanded the concept of pedagogical content knowledge (PCK), as developed by Shulman (1986), to develop the technological, pedagogical and content knowledge (TPACK) framework, which involves a complex interaction among content knowledge (CK), pedagogical knowledge (PK), and technological knowledge (TK). The interaction of these three basic forms of knowledge gives rise to pedagogical content knowledge (PCK), which closely resembles Shulman’s idea of knowledge of pedagogy, which is applicable to the teaching of specific contents. Meanwhile, technological content knowledge (TCK): refers to knowledge of how to use technology in order to develop innovative ways of teaching content (Niess, 2005). Moreover, technological pedagogical knowledge (TPK) attempts to understand how teaching and learning can change in impact when technologies are used in specific ways. Furthermore, TPACK describes knowledge required by teachers to integrate technology in the classroom for specific content areas (Koehler & Mishra, 2009). According to Koehler and Mishra (2009), TPACK provides the foundation for effective teaching by using technology. Cox and Graham (2009) asserted that TPACK has inspired teachers, teacher educators, and educational technologists to re-evaluate their technological knowledge and classroom application of technology. Some TPACK-related studies have also attempted to improve the ability of in-service teachers to integrate technologies into teaching (Niess, 2011; Wetzel & Marshall, 2012). In sum, TPACK provides the foundation for effective technology-based instruction.

Other investigations also examined how to improve the integration of technology by pre-service teachers using TPACK (Koh, Chai, & Tsai, 2010; Koh, & Divaharan, 2011). A few studies have developed TPACK applications for pre-service teachers. Koh and Divaharan (2011) found that pre-service teachers focus mainly on TK independent of ICT training. As they subsequently incorporate technology into their classroom instruction, those teachers began to address TPK-related issues. When allowed to consider the connections between technology-related skills and subject content, pre-service teachers present their TCK in course design (Koh, & Chai, 2014). However, pedagogical planning is essential to facilitate the transformation of subject content knowledge for learners (Pamuk, 2012). The lack of teaching experience among pre-service teachers significantly impedes their ability to apply technology in the classroom.

Several TPACK surveys have identified the pedagogical approaches of in-service teachers (e.g., Archambault & Barnett, 2010). Some studies have established that the PK of in-service teachers drives their integration of ICT (Harris, Grandgenett & Hofer, 2010; Liu, 2013). Moreover, other studies have suggested that in-service teachers emphasize subject content when they integrate technology into their classroom instruction (Koh & Divaharan, 2013; Koh & Chai, 2014). In Liu (2013), in-service teachers applied a PK basis, combined PK with subject content and
technology during a professional development program, and devised TPACK-related concepts. Most in-service teachers with PK and CK that exceeded their TK can consider various instructional strategies or subject content and adopt a technology that matches their own notions regarding technological integration.

In-service teachers and pre-service teachers may have expertise in various TPACK constructs. Based on the relationships between the two, mentors and mentees can exchange technology integration-related expertise.

**Collaboration among mentor teachers and pre-service teachers**

While generally more experienced than mentees, mentor teachers possess knowledge skills desired by the mentee (Ambrosetti & Dekkers, 2010). In an internship, mentor teachers are considered as a sample model for teaching by pre-service teachers. Teacher education has long adopted the apprenticeship model in mentoring, which reflects a hierarchical relationship between mentor and pre-service teachers.

Fairbanks, Freedman and Kahn (2000) defined mentoring in teacher education as complex social interactions in which mentor teachers and pre-service teachers construct and negotiate for various professional purposes and in response to contextual factors. Mentors and pre-service teachers can develop teaching expertise collaboratively while doing so in a school environment (Nilsson & van Driel, 2010). Unlike a traditional apprenticeship, the new relationship between mentors and interns deepens their interactions.

In the collaborative mentoring approach of van Velzen, Volman, Brekelmans and White (2012), mentors and pre-service teachers cooperate by sharing knowledge expertise, discussing subject contents and reflecting on teaching practices. Pre-service teachers and mentors learn from each other, subsequently improving their ability to identify and explain their teaching practices. Meaningful and collaborative relationships between pre-service teachers and mentors facilitate a focused dialogue on teaching and learning. The willingness of mentor teachers to discuss their individual challenges is important to conversing with pre-service teachers (Eller, Lev & Feure, 2014).

In terms of professional development, in-service teachers can also benefit from collaboration. Individually and collectively, teachers adopt new concepts in their classrooms and monitor the success of their efforts by jointly reviewing their work, considering outcomes, and reflecting upon their teaching efforts (Lau, 2013). However, even when teachers schedule regular meetings, the lack of effective cooperation procedures may impede collaboration (Pawan & Ortlof, 2011; San Martín-Rodríguez, Beaulieu, D’Amour, & Ferrada-Videla, 2005). Consequently, collaboration focuses solely on information exchanges rather than instructional observation, deep discussion and reflection (Pawan & Ortlof, 2011).

School-based teacher education involves more than just collaboration among in-service teachers, or between mentor teachers and pre-service teachers, as described above. Pre-service teachers can participate in peer mentoring, in which they play the mentor role during the activity (Le Cornu, 2005; Sundli, 2007). Pre-service teachers must openly share and support each other. The questioning of the ideas of a pre-service teacher by a peer fundamentally differs from that peer having their ideas questioned by an experienced mentor teacher (Nokes, Bullough Jr., Egan, Birrell & Hansen, 2008). Peer-mentoring programs enable pre-service teachers to serve as critical friends and provide constructive feedback following lessons so they do not need to depend entirely on feedback from their mentor teacher (Le Cornu & Ewing, 2008).

Theoretically, collaboration between mentors with sufficient teaching experience and pre-service teachers with technological skills can facilitate technology integration for professional development purposes. Notably, a school-based mentoring team that comprises both mentor teachers and pre-service teachers creates extensive opportunities for mutual learning.

Based on the study purposes and the literature review, this study addresses the following research questions.

- How do pre-service teachers and mentor teachers change in TPACK-related applications while participating in the CPD program?
- What collaborative experiences motivate pre-service teachers and mentors to adopt TPACK in practice?
Research methodology

This study examines the collaborative professional development of mentor teachers and pre-service teachers in relation to technology integration. As mentioned earlier, TPACK provides the foundation for applying technology to teaching effectively. Based on TPACK constructs, this study assesses the CPD of both mentor teachers and pre-service teachers with respect to their ability to integrate technology into their classroom instruction, with a particular focus on changes in TPACK. Therefore, this study adopts qualitative research methods, including classroom observations and focus groups.

Participants

In Taiwan, roughly 6,000 pre-service teachers participate in school-based field practice annually, located in about 1,800 accredited K-12 schools nationwide by local governments for teaching practice purposes. Each pre-service teacher is partnered with at least one experienced teacher for mentoring based on the Teacher Education Act legislated by the Taiwanese government. To obtain transferability, an accredited junior high school was invited to participate in this study. After the rights and obligations of the participants were announced, three mentor-teachers (i.e., one special education teacher, a mathematics teacher and a biology teacher, with 15, 14 and 18 years of teaching experiences, respectively) and their pre-service teachers from different universities of teacher education voluntarily participated in the CPD program from October 2013 to January 2014. No participants had graduated from a technology-related university department. Three mentors occasionally implemented teaching activities with PowerPoint presentations and a video projector.

Collaborative professional development program

The CPD program was implemented via four cycles, with each one involving practical teaching with technology/classroom observations and participation in FGI. Before teaching, instructors were asked to design teaching activities that used technology. While teaching, observers had to write observation records. Meanwhile, following teaching instruction, all participants attended a FGI.

In the first and the third cycles, each mentor demonstrated teaching activities that used technology. In the second and fourth cycles, each pre-service teacher taught a lesson using technology. After each cycle was completed, the instructional plans and recorded videotapes of the instructors were collected for analysis to design interview questions for FGI. During FGI, teachers could discuss their teaching practices and identify how their work changed between cycles. The FGIs enabled participants to learn from the experiences of others and share ideas on technology integration. Analytical data obtained from classroom observation were triangulated, based on perspectives of technology integration obtained from FGI records.

Besides interacting regularly with students, parents and school activities, each mentor teacher and pre-service teacher pair were encouraged to collaboratively develop technology integration-related ideas. During their interactions, the pre-service teachers helped their mentors to design digital instructional materials; otherwise, the pair collaborated with each other. A pre-service teacher in Taiwan can usually obtain teaching guidance from their mentors before teaching practice. After teaching practice, the mentor often identifies his or her related strengths and weaknesses and also offers recommendations to pre-service teachers. In this study, the interactions were extended to facilitate collaborative professional development in the area of technology integration in order to include interactions with other mentors and pre-service teachers. Accordingly, each mentor teacher or pre-service teacher contacted other participants regarding their technology integration activities. Notably, besides the relationships with mentors, pre-service teachers can learn of TPACK-related applications from peer mentoring and the mentors of other pre-service teachers. These interactions were not recorded yet encouraged to promote self-reflection and the development of new perspectives that could then be expressed during FGI meetings.
Instruments

This study also attempted to achieve dependability in qualitative research (as demonstrated by the use of overlapping methods) and validate the analytical results. Therefore, primary research data were obtained from instructional plans, classroom observations and FGI records.

Classroom observations

Classroom observations allowed all participants to share ideas and stimulate dialogue and reflection on technological integration activities. In addition to benefitting the pre-service teachers, the teaching demonstration of each mentor also benefitted other mentors involved in the CPD program. A model of CPD for mutual learning was constructed, based on observations of and by mentor teachers and pre-service teachers. All teaching activities were video-recorded. Moreover, all participants other than the instructor were asked to either observe the lessons or view the video records.

Particular emphasis was placed on classroom observations, in which an observation sheet was created to describe teaching activities that use technology by using the basic constructs of the TPACK (including PCK, TCK, and TPK), based on the evaluative framework of Pamuk (2012). The observation sheet consists of four columns, i.e., TCK TPK, PCK and TPACK descriptions. All participants were instructed to fill in the sheet. For example, the TCK column in an observation sheet appeared in which the instructional material was designed as animated pictures.

Focus group interviews

According to Krueger (1988), focus groups evaluate programs and the generation of insights that can contribute to efforts to develop strategies. Accordingly, by using FGIs, this study evaluated changes in how participants integrate technology into their work and also formed related perspectives collectively. The interview questions were based on the analysis results of instructional plans, recorded videotapes, and records based on classroom observations. Although certain questions were based on an analysis of instructional plans, videotapes, and classroom observation sheets to facilitate data triangulation, most FGI topics involved technology integration because this study deals with collaborative professional development as reflected by changes between the initial and final perspectives regarding this topic, based on practical experiences. Sample questions are as follows.

- Please briefly state your (mentors and pre-service teachers) instructional plan and its proposed rationale. How are technology, teaching strategies, and content knowledge related to each other.
- Did you encounter challenges in designing and conducting activities that integrated technology? If so, how did you overcome such challenges?
- Did you discuss technology integration activities with other participants, including the participant with whom you were paired? If so, what did you discuss?
- How did interacting with other participants help you to reflect on your perspectives on technology integration, if at all?

Data analysis

Data obtained from instructional plans, video-recorded classroom observations and FGIs were analyzed using a constant-comparison analysis method. To obtain confirmability, three trained researchers independently reviewed the instructional plans of each participant, developed analytic headings and codes based on TPACK constructs, and compared video-recorded classroom observations with the analytical results of instructional plans. The researchers subsequently formed a consensus on the technology integration performance of instructors. When a pattern of relative clues emerged, the researchers again reviewed the analytical results to confirm any trends among the various data. The extent to which instructors adopt TPACK provides a contrast with classroom observation records in terms of indicating similarities and differences between the opinions of instructors regarding technological integration activities. These differences may reflect the ability of instructors or creative technology integration-related ideas. Comparing instructional plans and video-recorded classroom observations revealed specific events of interest,
subsequently revealing patterns in the data on technology integration activities, which were attributed to collective professional development. Above results were confirmed through FGIs. The above analysis was conducted during each of the four CPD cycles.

To ensure achieve credible qualitative research, the performance of each participant, in relation to their instructional plan, actual teaching, observation sheet and communication in FGIs, was triangulated across the four cycles to clarify the changes in technology integration based on TPACK. Positive changes reflect professional development in technology integration.

Results

Above results demonstrate the effectiveness of three aspects of the CPD program in technology integration. These aspects involved the change in the teaching activities of mentor teachers, from presenting a TCK to building various TCK bases, the change in the focus of the pre-service teachers from TK to TCK and PCK, as well as the fact that the CPD program improved the professional development of pre-service teachers more than that of mentor teachers in TPACK utilization.

Changes in the ways in which mentors integrate technology in their teaching practices

When designing instructional plans, three mentor teachers initially used PowerPoint presentations to introduce instructional materials. During the first FGI, two mentor teachers considered the effectiveness of presenting abstract concepts using PowerPoint slides because the combination of pictures and words can support comprehension. Although mentor teachers recognized that technology-based teaching improves learning comprehension and facilitates teacher-student interaction during teaching, they did not integrate technology integration into every lesson. A mentor expressed the following during the first FGI:

Students must occasionally perform a science experiment... I do not think technology-based teaching is appropriate in all science lessons, unless abstract concepts are taught. I am accustomed to presenting pictures with descriptive text. (Mentor A, 1st FGI)

As for the use of TPACK constructs, the mentor teachers presented TCK concepts during lectures. At the beginning of this CPD program, mentor teachers mastered the subject contents and used simple technological skills to lecture using a digital presentation. The digital material attracts student interests, based on an observation sheet provided by a pre-service teacher.

Based on observation of the teaching activities of pre-service teachers, the mentor teachers admitted that the pre-service teachers had superior technology skills. For example, the pre-service teachers made animated pictures, and were more proficient at performing Internet searches. A mentor teacher thus asked her pre-service teacher for assistance during the second and third cycles.

After class, I asked her if she could send me the animated pictures. Maybe I should make similar materials with her guidance. I’ll use animated pictures to teach students in the future. (Mentor A, 2nd FGI)

Besides presenting additional materials obtained online, mentor teachers slightly adjusted their presentations, such as by using animated pictures in PowerPoint slides; whereas, previously, they had only used static pictures. Analysis of observations of student responses to teacher questions revealed that teachers enhance student motivation and comprehension of course contents. Two mentor teachers stated the following:

I raised some questions on slides to encourage the students to think independently. My questions and answers were immediately displayed as animation effects, explaining why no time was wasted in writing on a blackboard. Students felt excited when they answered correctly. (Mentor B, 3rd FGI)
My technological skills are improving. After viewing the slides made by the pre-service teachers, I realized that PowerPoint has many functions that I had never used previously. I have now learned creative ways to display contents and will try to use them in the future. (Mentor A, 4th FGI)

Mentor teachers recognized that the pre-service teachers outperformed them in terms of technological skills, subsequently motivating them to actively learn technological skills from those pre-service teachers, as reflected in the following comment:

I want to use software other than PowerPoint, such as Cyberlink PowerDirector, and Photoshop. I can use software to edit or produce videos for students. (Mentor C, 4th FGI)

Once having improved their technological skills by collaborating with and learning from pre-service teachers, the mentors began to add animated pictures to their PowerPoint slides in order to improve the concentration and motivation of their students. Mentors increasingly adopt various technologies to assist students in comprehending abstract concepts and increasing their learning motivation. Therefore, mentor teachers changed the way in which they went from using a single TCK to utilizing various TCK approaches in teaching activities.

Development of the ability of pre-service teachers to integrate technology

Initially, the teaching activities of the pre-service teachers using technology corresponded to the models of technological integration used by their mentors, which involved using digital presentations to support their lectures, as reflected in the following comment:

Using slides reminds me to slow down while teaching. Additionally, the animated pictures ensured that my lecture went more smoothly. (Intern C, 1st FGI)

However, the pre-service teachers admitted that they taught more poorly than their mentors when they used technology because they concentrated too much on how to use it. Following observation and reflection, they began to recognize how to improve their teaching skills. An observation sheet from a mentor teacher noted how the pre-service teacher with whom they were paired presented digital material together with their lecture, yet spent too much time standing in front of their computer, and did not interact much with their students. The pre-service teacher also reflected upon their teaching experience in the second FGI:

Comparing my teaching with that of my mentor made me realize that I simply stood by the podium when using slides. I failed to interact with the students, explaining the ineffectiveness of my approach. (Intern B, 2nd FGI)

I observed that my mentor did not use the computer continuously throughout the class and also that his teaching was more flexible than before. Using slides to introduce mathematical steps is occasionally difficult. I saw that my mentor used the blackboard whenever deemed necessary, such as when questioning students. I plan to make this approach a model for my own teaching method. (Intern C, 2nd FGI)

The instructional plans of pre-service teachers in the third cycle included more technological material, such as films. However, as they had previously wanted, the pre-service teachers began to perform multiple teaching activities, and also became concerned with technologies such as animated pictures, which were most appropriate for specific subject content.

Biology as a subject includes varied content knowledge. I cannot adopt the same method to teach all chapters of the textbook. My mentor thus recommended that I combine questions with animated pictures to more closely interact with the students. (Intern C, 3rd FGI)

Mentor C identified changes in the instructional strategies of pre-service teachers using technology that was related to TPK when using the observation sheets. Mentor A also confirmed changes in the instructional strategies of pre-service teachers when discussing instructional plans with two pre-service teachers. Throughout the CPD program, the
pre-service teachers considered student characteristics when designing activities that involved technology, as expressed in the following comments:

*During course design, I ignored the prior knowledge of students and did not consider how appropriate they were for my course design. I must therefore be more aware of the ability of my students (Intern C, 3rd FGI)*

*A teaching activity that uses only one teaching method is incomplete. Teaching methods should therefore be revised to match student learning styles. (Intern A, 4th FGI)*

During the initial stage of the CPD program, pre-service teachers relied on the technology integration model of their mentors. Pre-service teachers focused on TK applications, and their lectures were based on slide presentations. Classroom observations and FGI discussions facilitated reflection by the pre-service teachers on their use of technology. After the third FGI, pre-service teachers began to combine subject contents with animated pictures and considered multiple teaching methods that suited the learning abilities of their students.

**Benefits of the CPD program for mentors and pre-service teachers**

During the CPD program, including the after class period, collaboration among mentor teachers was insufficient to develop TPACK, although they observed each other’s teaching and recorded several descriptions of teaching activities in the observation sheets. Their interactions were also limited to technological skills rather than subject contents or instructional strategy. One mentor teacher made the following comment.

*We exchanged digital multimedia by e-mail…. With our different office locations and busy schedules, we rarely discussed among each other. (Mentor A, 3rd FGI)*

The frequency of mentor-mentor interactions remained unchanged during the CPD program, while that of mentor-intern interactions increased. Before the program, mentor teachers assisted pre-service teachers in their classroom performance. In response to problems, pre-service teachers could approach any mentor teacher for assistance. The program increased the frequency of mentor-intern interactions, and especially increased the frequency of interactions concerning technology integration. Notably, besides enabling regular discussions on subject content, interactions between mentors and interns helped both parties to develop new interests. The pre-service teachers focused on developing pedagogical knowledge, while the mentors focused on using various technological functions to present knowledge, as reflected by the following comments:

*I would often wonder about the effectiveness of my instructional strategies, explaining why I often called on my mentor. My mentor also suggested how I could lecture more effectively. (Intern B, 2nd FGI)*

*Given my hectic schedule, I have no spare time to design digital materials using PowerPoint software. My pre-service teacher offered assistance after we discussed the subject contents. Afterward, we often discussed how to present effective digital materials in order to improve learning. (Mentor C, 2nd FGI)*

Pre-service teachers recognized that cooperation encouraged and strengthened peer-mentorship, as reflected by the following comments:

*Despite my fear of teaching poorly, I seldom sought my mentor for assistance. Pre-service teacher C always encouraged me and helped whenever I ran into difficulties in the classroom, especially in relation to software utilization. I am happy to learn from other pre-service teachers. (Intern A, 4th FGI)*

*Before teaching, I asked pre-service teacher C about lesson design. We discussed the program requirements and shared ideas. I found this discussion helpful in integrating technology in my classroom. (Intern B, 2nd FGI)*

Analytical results indicate that both mentor teachers and pre-service teachers developed professionally. However, mentors collaborated only on exchanging digital material. Notably, pre-service teachers were critical to the CPD program. The pre-service teachers, especially Intern C, were superior to their mentors in technological skills. Like
video clips, these technological skills not only motivated mentor teachers to apply technology in teaching, but also enabled them to help their mentors and other pre-service teachers design technological applications for teaching. Additionally, the pre-service teachers actively sought further instruction from mentor teachers and seized additional opportunities than previously to discuss technology-integrating activities with other pre-service teachers. By increasing learning opportunities and peer-mentoring opportunities, pre-service teachers recognized their own professional development in integrating technology.

In summary, pre-service teachers benefit from the CPD program more than the mentors. In particular, the pre-service teachers benefited from various opportunities to learn technology integration-related applications, while the mentors lacked sufficient peer collaboration in technology integration during the CPD program.

Discussion

This study examines the collaborative professional development of a group of mentors and pre-service teachers in technology integration. During the CPD program, the findings regarding poor collaborations of mentor teachers in technology integration differ from those of Tondeur et al. (2013). According to their results, collaborating with colleagues improves technology integration ability. The poor collaboration is owing to that the mentor teachers had limited motivation to work with other mentor teachers, as mentioned elsewhere (Pawan & Ortlof, 2011; San Martin-Rodríguez, et al., 2005), thereby limiting the exchange of ideas regarding technology application. In Taiwan, in-service teachers are accustomed to working alone, thereby limiting their willingness to discuss teaching practices with each other. Technological skills are also insufficient for mentors to integrate technology into their classroom instruction. In this study, the technological skills of pre-service teachers can fulfill the requirements of mentor teachers. This study identifies the benefits of collaboration between mentor teachers and pre-service teachers (Chen, 2012; Goodnough, et al., 2009). When supported by pre-service teachers in using technology and designing digital materials, mentor teachers can easily combine various technological applications and subject contents. Since pre-service teachers support the use of technology in ways related to subject contents, the technology integration activities of mentor teachers tend to construct the TCK base. Therefore, during this study, mentor teachers went from using a single TCK to constructing various TCK bases.

Our results further indicate that the technology-based teaching activities of pre-service teachers shifted the focus from TK to TCK and TPK, despite the difficulty of identifying overall TPACK development. Similar to a study of Koh and Divaharan (2011), the pre-service teachers focused on TK at the start of the CPD program. The program extended the opportunities enjoyed by pre-service teachers to observe the teaching activities of mentors and other pre-service teachers. Those activities were discussed as well. In particular, pre-service teachers became mentors for each other, which is consistent with previous findings (Le Cornu, 2005; Sundli, 2007), and acted as critical friends and providers of feedback. These pre-service teachers easily form collaborative relationships and obtain more support and feedback from their peers than from their mentors. Guidance from mentors with respect to pedagogy knowledge and interactions with skilled peers enabled pre-service teachers to constantly practice the application of TCK and TPK, subsequently developing professionally in technology integration concepts related to TPACK.

Pre-service teachers undoubtedly strive to obtain practical experiences and adopt various instructional strategies to prepare for their future careers. Highly motivated, the pre-service teachers in the CPD program actively engaged with all mentor teachers and other pre-service teachers at school and shifted the focus of technology integration activities from TK to TCK and TPK. The mentor teachers examined in this study were motivated only to apply technology. With the assistance of pre-service teachers, experienced mentors with sufficient PCK can develop technology integration ideas faster than ever. Although mentor teachers did not implement sufficient technology integration activities, pre-service teachers learned about the methods of technology integration through discussion with their mentors. Since action is superior to knowledge, the CPD program enhanced the professional development of pre-service teachers in relation to technology integration more than it did for the mentor teachers.

Based on the TPACK constructs, technology integration activities involve more pedagogical activities than simply lecturing. Insufficient in-service training limits the effectiveness of teachers in integrating technology in their classroom. As a school-based teacher education program, the CPD program benefits both mentor teachers and pre-service teachers by extending collaborative opportunities, thus enabling all teachers to acquire knowledge skills from each other. Observing the instruction and discussions of pre-service teachers motivated mentors to integrate
technology in their teaching practices. The above perspective regarding the extent to which the active learning performance of pre-service teachers affects mentors appears to facilitate the professional development of in-service teachers in relation to technology integration.

Although previous investigations have demonstrated the effectiveness of collaboration between a pair of pre-service and mentor teacher (e.g., van Velzen et al., 2012), this study uniquely identifies that pre-service teachers gain more from the CPD program than mentor teachers do. Conversely, the collaborations among mentors were poor, slightly limiting applications to a school-based mentoring team.

In this study, several mentor-intern relationships provide numerous opportunities for mutual learning by mentors and interns. Pre-service teachers were also involved in peer mentoring and were willing to provide mutual support. The CPD program is based on collaboration between mentor teachers and pre-service teachers in a school. Undoubtedly, the program can be applied for traditional school-based teacher education, which provides few opportunities for meaningful collaboration with other educators. When the CPD program is implemented in a school-based teacher education situation, pre-service teachers skilled in using computers could facilitate the implementation of the mentors’ ideas on TPACK application and further obtain the guidance from mentors for integrating technology in the classroom. Both mentors and pre-service teachers develop professionally during the collaboration in technology integration.

Conclusions and future research

This study demonstrates that both pre-service teachers and mentor teachers developed professionally in relation to technology integration through the program implemented here. In particular, the mentor teachers changed their teaching activities when they had the support of pre-service teachers: from presenting knowledge of technological contents to constructing various TPK approaches. The pre-service teachers shifted to using TCK and PCK, having originally focused on TK applications and the application of slide presentations to teaching. Moreover, the mentors developed technology integration ideas with the assistance of their pre-service teachers faster than previously. Afterwards, the pre-service teachers learned more about the methods of TPACK applications through discussion with all mentors as well as interactions with skilled peers in a school-based mentoring team.

Notably, the CPD program benefited pre-service teachers more than mentor teachers since the former sought more opportunities for the practical application of TPACK than the mentor teachers did, who only exchanged technological applications.

This study developed a strategy in which pre-service teachers seemed to be an initiator for CPD in a mentoring team. We recommend that future research efforts identify the feasibility of the strategic perspective in school-based teacher education. Future research should also investigate the factors resulting in poor collaborations of the mentors in a school-based mentoring team.

References


Personal Learning Network Clusters: A Comparison Between Mathematics and Computer Science Students

Ansie Harding1 and Johann Engelbrecht2*

1Department of Mathematics and Applied Mathematics, University of Pretoria, South Africa // 2Department of Science, Mathematics and Technology Education, University of Pretoria, South Africa // ansie.harding@up.ac.za // johann.engelbrecht@up.ac.za

*Corresponding author

(Submitted March 11, 2014; Revised October 30, 2014; Accepted January 26, 2015)

ABSTRACT

Personal learning environments (PLEs) and personal learning networks (PLNs) are well-known concepts. A personal learning network cluster is a small group of people who regularly interact academically and whose PLNs have a non-empty intersection that includes all the other members. At university level PLN clusters form spontaneously among students and are known to last over an extended period of time. Little is known regarding the workings of these PLN clusters of students. The claim is that these PLN clusters are at the heart of student learning and are aligned with the current trend of a knowledge-pull community of learning. In this paper we investigate the activities and characteristics of PLN clusters in two different fields of study at a South African university, namely mathematics and computer science. We discuss the benefits that these clusters offer, investigate the mashup of activities and tools and we contrast experiences in the two fields of study. It is the commonalities rather than differences that are striking between the two groups of students. Although computer science students lean more towards digital communication, both groups impress with the pride they take in their PLN clusters and are vocal in describing the benefits that these clusters offer.

Keywords

Personal learning network clusters, Learning management systems, Learning communities, Collaborative learning

Background

This study sets out to investigate the phenomenon of spontaneous formation of small clusters of students, mostly at the onset of their study years, interacting academically as well as on a personal level over an extended period of time. We refer to these groups as personal learning network (PLN) clusters. The setting is in one of the principal universities in South Africa where we investigate and compare two groups of students, mathematics and computer science students, and report on their experiences within their PLN clusters.

Roger and Johnson (1988) surmise that how the students perceive and interact with one another is a neglected aspect of instruction. Although much training time is devoted to helping teachers arrange appropriate interactions between students and materials, and some time is spent on how teachers should interact with students, little or no time is spent on how students should interact with one another. This premise is particularly true for students in PLN clusters and it is in this sparseness of knowledge on how students interact amongst themselves that the main contribution of this paper lies.

We colour the background setting of this paper through a discussion of collaborative learning, and the various faces of technology enhanced learning, and how collaborative and technology-enhanced learning are combined within PLN clusters.

Collaborative learning

A model for collaborative learning that has typically been practised is of assigning small groups of students within a larger group with the purpose of discussion or for completing an assignment. The responsibility of dealing with the smooth running of structured collaborative learning mainly rests on the shoulders of the instructor who is often faced with issues such as fair selection of groups, how to design activities for collaborative learning, how to deal with conflicts within groups and how to fairly evaluate individual work. In a large course, formal group work becomes
increasingly hard to administer. The group itself may experience internal difficulties such as starting tasks late, conflict between team members due to personality clashes, and tasks not being completed on time.

**Technology enhanced learning**

Technology Enhanced Learning (TEL) is at the order of the day, all learning that involves technology is technology enhanced and collaborative learning dovetails with TEL. TEL has many faces, some of which we proceed to discuss.

**Learning management systems (LMS)**

LMSs have been adopted by universities all over world at a remarkable rate (Coates, James, & Baldwin, 2005). A case can certainly be made for enhanced learning resulting from utilising the features of an LMS appropriately.

Graf and Kinshuk (2007) analysed students’ performance and behaviour in a course in which they provided adaptivity based on students’ learning styles in the LMS and found that with this add-on the LMS proved effective. Vovides, Sanchez-Alonso, Mitropoulou, and Nickmans (2007) claim that the powerful features of LMS systems are often underutilized and plead for developing LMS systems that offer new opportunities to personalize instructional support in the form of learning objects with metadata, including features and functionalities such as intelligent learning mentors or tutors.

Inept use of LMSs has led to criticism. Siemens (2004) states:

> “Learning Management Systems are often viewed as being the starting point (or critical component) of any e-learning or blended learning program. This perspective is valid from a management and control standpoint, but antithetical to the way in which most people learn today” (p. 1).

Posting static content on an LMS is cautioned. Chatti et al. (2010a) bluntly state that the LMS-centric model of learning has failed to achieve performance improvement and innovation. In most cases an initially paper-based learning resource is simply converted into digital form and a classroom training event is transformed into an online course. The pattern is of modularisation of courses and isolation into discrete units. LMS driven models follow a one-size-fits-all approach and suffer from an inability to satisfy the heterogeneous needs of many learners.

Coates et al. (2005) state:

> “It is important that steps are taken to identify how online LMS can be used to augment and complement rather than substitute for an institution’s core teaching objectives” (p. 33).

Garcia-Peñalvo, Conde, Alier, & Casany (2011) argue that with LMSs users have reached a plateau of productivity and stability and that this stability and maturity of the LMS may become a resistance factor working against the introduction of innovations. Within the PLE and PLN paradigms, new transformative tools, services and ways of learning are already in use, which cannot be ignored.

The above reasons stated for the “failure” of LMS’s in the learning environment may be somewhat misplaced. LMSs can be used productively in a technology-enhanced learning environment if used in a pedagogically sound manner. LMSs have features that enable collaborative learning, such as discussion forums, blogs etc. to be used for TEL and within PLN clusters.

**Using technology in collaboration models**

**Network learning**

All learning is situated in practice and all practice is essentially social in nature (Swan & Shea, 2005) and therefore commonly resides in networks. The notion of network learning is supported even stronger by a more social and
connected Web that has arisen over the past years. Web users absorb knowledge very quickly across different channels and they prefer random on-demand access to all kinds of knowledge disseminated over the internet. They are in close contact with their friends using networks to share and create new knowledge. This new Web generation is often termed Web 2.0. (Chatti et al., 2010a).

The concept of an open network learning environment was discussed by Tu, Sujo-Montes, Yen, Chan, & Blocher (2012).

“Open network learning environments are digital environments that empower learners to participate in creative endeavours, conduct social networking, organize/reorganize social contents, and manage social acts by connecting people, resources, and tools by integrating Web 2.0 tools to design environments that are totally transparent, or open to public view; the same architecture can be used to design the degree of openness users feel is necessary to the situation” (p. 14).

They report on an online course designed to empower learners to construct their own personal learning environments within open network learning environments. They conclude that effective instructions should prepare “online” learners to become “network” or “open network” learners. We subsequently link our study to this notion.

Personal Learning Environments (PLE) and Personal Learning Networks (PLN)

Personalized learning is linked to the idea that learners do not have control over what is taught but do have control over what is learnt (Tobin, 2000). Learners want to control what they learn and for this purpose create a personal learning environment (PLE). Attwell (2007) defines a PLE as a collection of all the tools we use in everyday life for learning. PLEs enable a learner-controlled integration of different learning tools and services into a personalized space. The space takes a learner-centric approach and is characterized by the use of a set of services and tools that belong to and are controlled by the learners. Rather than integrating different services into a centralized system, the idea is to provide a learner with a plethora of different services and hand over control to the learner to select, use, and mashup the services the way he/she deems fit. A PLE driven approach does not only provide personal spaces, which belong to and are controlled by the user, but also requires a social context by offering means to connect with other personal spaces for effective knowledge sharing and collaborative creation (Chatti et al., 2010a).

Ivanova (2009) describes a blended learning course during which a learning network is formed, following the modified Rogers’ (2003) model for competence development lifecycle in a learning network. She found the main challenges in forming this learning network are to (1) provide sustainable value to students, and (2) stimulate students to contribute their knowledge, insights and experiences on a continuous basis. She found that social networks contribute to the processes by which learners meet and share their competencies. Building a PLE is found to be a core for PLN deployment.

Wilson, Liber, Johnson, Beauvoir, Sharples and Milligan (2007) argue that a PLE should focus on coordinating connections between the user and a wide range of services. It should recognize the need to integrate experiences in a range of environments, including education, work, and leisure activity.

“Unlike the VLE [virtual learning environment], the PLE is concerned with sharing resources, not protecting them, and emphasizes the use of creative commons licenses, enabling editing, modification, and republishing of resources. Rather than pre-packaged learning objects, the resources collected and accessed using the PLE are more typically weblog postings, reviews, comments, and other communication artefacts” (p 32).

Conde, García, Alier and Casany (2011) advise that the problems experienced with LMSs can be solved with the integration of LMSs into PLEs, thus integrating informal learning tools and contexts and formal environments. Similarly, in order to deal with how PLEs interact with institutional LMSs, Derek Morisson in Scater (2011) ventures the notion of a PLE docking into the “mother ship” (an LMS) every so often, to bring or get content.

Related to the concept of a PLE, and sometimes indistinguishable from, is the idea of a personal learning network (PLN). Whereas PLEs are the tools, artefacts, processes and physical connections that allow learners to control and
manage their learning, PLNs extend this framework to include the human connections that are mediated through PLEs (Couros, 2010). While there is a growing field of research and thinking behind the concept of a PLE, the academic research on personal learning networks (PLNs) is much more anecdotal. Couros (2010) defines a PLN as the sum of all social capital and connections that result in the development of a personal learning environment. PLNs offer an environment where people and tools and communities and resources interact in a very loose kind of way (Wilson, 2008). The longevity of PLNs is one of its most important features. Time-based, course-centric communities die, probably because these communities are based around courses and not around communal learning. PLNs cultivate sustained, long-term learning.

A PLN consists of the people involved with the use of tools only implied - an informal network of the people a learner interacts with and derives knowledge from in a PLE. In a PLN a person makes a connection with another person with the specific intent that some learning will occur because of that connection.

**Personal learning network clusters**

We define a *personal learning network cluster* as a small group of people who regularly interact academically and whose PLNs have a non-empty intersection that includes all the other members. Where a student’s PLN could consist of a number of other students not all doing the same subjects, a PLN cluster would be a subset of the student’s PLN, a smaller group of students that are together to address a particular subject or topic. In this paper we focus on PLN clusters that spontaneously form among students and we investigate how they function in two fields of study – mathematics and computer science. These clusters usually stay intact collaboratively over the duration of its members’ years of studying and often beyond. Students in a cluster use a number of tools to communicate and learn while using social media, mobile phone technology and learning management systems, among other platforms for learning purposes.

**Mashups**

The creation of rich learning mashups (normally web applications that integrate complementary elements from more than one source) currently associated with collaborative learning resulted from advances in digital media and data in the early 21st century. Wild, Kalz and Palmér (2010) describe mashups as “the ‘frankensteining’ of software artefacts and data” (p. 3). They describe the development of a technological framework enabling students to build their own personal learning environments by composing web-based tools into a single user experience, getting involved in collaborative activities, sharing their designs with peers and adapting their designs to reflect their experience of the learning process.

There is an opinion that technology enhanced learning (TEL) has not succeeded in revolutionizing education and the learning process. A speculative reason is that most current initiatives take a technology-push approach in which learning content is pushed to a predefined group of learners in a closed environment (Chatti, Agustiawan, Jarke, & Specht, 2010a). A fundamental shift toward a more open and learner-pull model for learning is needed - a shift toward a more personalized, social, open, dynamic, emergent and knowledge-pull model as opposed to the one-size fits all, centralized, static, top-down and knowledge-push models of traditional learning. Chatti et al. (2010b) suggest the 3P learning model as a new approach for addressing the growing complexity and constant change of knowledge required, tailored for the new generation. The 3P learning model encompasses three core elements: personalisation, participation and knowledge-pull. It is in these three core elements that PLN clusters are shown to shine.

**User requirements**

Students join PLN clusters for reasons that are most likely not explicitly formulated in their own minds. Such implied needs are addressed in successfully functioning PLNs and if not members would leave or become inactive. As these PLN clusters spontaneously form as a natural conglomeration of students through convenience of interaction on a social or academic level, the user requirements are not stipulated at the onset. The common goal of wanting to succeed at university is a draw card for like-minded students to join forces, as is the natural phenomenon
Hogg, Hohman, and Rivera (2008) offer three reasons why people join groups. The first of these reasons is that people have a need to belong, and that self-esteem acts as a meter of successful cluster belonging (sociometer model). The second reason is that it is comforting to share our world views with like-minded others because it provides us with a sense of meaningful existence (terror management theory). The third reason is that people have a basic need to reduce uncertainty about themselves and their place in the world, and that group identification can reduce such uncertainty (uncertainty-identity theory).

Benefits of study groups are informally documented via blogs (e.g., Desmond, 2014). Such benefits include counteracting procrastination because of regular meeting times, learning faster through group momentum, getting fresh perspectives on a topic, learning new study skills through observation, breaking the monotony of studying on your own, filling in gaps in your own learning through comparing notes etc., practicing for a career because of experiencing group dynamics.

This study was conducted on existing PLN clusters and we can therefore not determine what the user requirements of these students were for joining these clusters. We can however determine why students stay in these clusters and what they perceive as the positive features so as to post hoc address user requirements.

The study of PLN clusters aims to alert and enlighten educators to this valuable learning model in which both collaboration and TLE are essential features.

**Approach**

We follow a case study methodology in investigating the phenomenon of PLN clusters, focusing on a single university in South Africa. PLN clusters are studied from two perspectives – one from within the technology rich field of computer science and the other from within mathematics, a field that is less rich technology wise but requiring use of symbolic exposition. By investigating and comparing activities of PLN clusters in mathematics and computer science, commonalities are extracted and differences are reported on to gain a deeper understanding of this phenomenon. The sample space consists, on the one hand, of students enrolled for a second year mathematics course, either studying for an engineering degree or for a science degree, and on the other hand a group of second year computer science students of which a few specialize in multimedia. A simple, voluntary first questionnaire was issued to both groups to determine whether a student was part of a group of close friends who came together informally for study purposes. The term personal learning network was not used. Second year groups were chosen as first years were still finding their feet and the possibility of existence of stable study groups was minimal. Another reason for focusing on second year students is because engineering students do mathematics only up to second year level and we wanted to include them in order to address a cross section of students. In order to match the academic level, second year computer science students were included in the sample.

In total 111 mathematics responses were received of which 49 indicated that they were part of a PLN cluster. Of these 22 students indicated willingness to receive a second, follow-up questionnaire via email regarding the functioning of their PLN cluster. A total of 13 students responded to the second questionnaire.

A total of 54 responses were received from computer science students of which 39 indicated that they were part of PLN cluster, all willing to participate in a follow-up questionnaire. A total of 17 students responded to a follow-up email questionnaire.

These numbers should not be seen as a representative sample. The purpose of the study was to investigate the working of PLN clusters in order to gain insight into this phenomenon and to do a comparison between experiences of mathematics and computer science students and for this purpose the samples sufficed. It should be noted that the computer science student response percentage was considerably higher and could point to a difference in attitude, possibly being more prepared to share information.
In the second questionnaire students were asked about the size of their clusters, how long the cluster had been going, how they got together and if the cluster has changed since it started. They had to explain how the cluster functions, how and how often they communicated, what technology was used and what their discussions were about. Lastly they had to mention benefits of the collaboration, what they personally got from belonging to the cluster and whether they preferred a spontaneous cluster to a group formed by the teacher.

Results

Cluster structure

The general trend is for clusters to form at the onset of the first year of study, due to being friends earlier or due to meeting in class. Mathematics clusters seem to consist of 3 – 5 members in general while computer science clusters seem to be slightly bigger with groups of 6 – 8 members not uncommon. Both mathematics and computer science clusters have regular face-to-face meetings, even for a mathematics cluster of nine members. The fact that computer science students have the same schedule and therefore have coinciding spare time makes face-to-face meetings easier to arrange. An interesting single phenomenon recorded is a computer science cluster of 30 members, a cluster that expanded from an initial 15 members to double its original size. This is a WhatsApp group with limited face to face interaction, sharing notes and previous exam papers through a Google drive folder. An advantage of such a large cluster is that when a question is posed someone is bound to have the answer. This cluster is by far the largest and probably more aligned with social media trends but seems to function well in terms of a cluster. Clusters are dynamic and undergo change but not dramatically so. Certain members participate less often from time to time due to work constraints; a member may leave the cluster because he finds that he studies better on his own; here and there a member is added along the way, or a member may quit studying and be replaced by another. The stability of the clusters represented by the responses was noticeable.

Understanding of what a PLN is

The two different group responses to what their understanding of a personal learning network is, is characteristic of the study fields. Mathematics students admit in general to not being familiar with the term personal learning network. Only one student could describe it as a group of people who know each other on a personal level and work together to optimize their learning environment. The majority of computer science students has an understanding of the term and have come across it in their environment. The “people” aspect is mentioned repeatedly.

A bunch of people helping each other to learn and understand the concepts.
It is your group of people that helps you to broaden your knowledge base and better your understanding of certain things.
It is an informal group of people working towards the same educational goal, helping each other.
I would assume it means a network of resources that is not provided by the institution, in our case the university.

Communication and activities

Both face to face and digital communication appear to be integral to all cluster activities. Mobile phone communication, especially via a WhatsApp group seems to be most common, with frequency varying from once a week to multiple times per day until “late hours.” Facebook is also commonly used, as well as Skype and Google Talk. E-mail communication is mentioned as an option, more frequently used by mathematics students due to the “difficulties of using symbols on a phone.” Google docs and Google Drive (or the equivalent Dropbox) are used among computer science students in particular for sharing documents. Face to face sessions are either scheduled for a fixed time(s) per week or on a needs basis, often triggered by a due assignment. Meetings are set up typically via WhatsApp. Clusters are also known to meet informally on campus during free times, often on a daily basis, in the library on campus, at each other’s homes or even in the cafeteria. Among the mathematics cohort face to face
communication seems to be the primary activity with technology often used only for organisational purposes whereas among the more technologically oriented computer science group face to face and digital communication are equally well used, both for learning purposes.

All activities are self-generated and indicative of a knowledge-pull approach. Cluster activities for mathematics students centre on problem solving, mainly preparing for practical sessions or tests and exams. An important component of the technology side of the communication channel is that clusters have access to a helpline at all hours of the day. One student describes it as “using collective knowledge to achieve our goals”. A healthy approach to mastering the work is illustrated by

Each person tries work on their own, if they don’t succeed any of the other people in the group are contacted to help the individual.

The practice of collecting exam-type questions for when the cluster gets together and try to solve those problems seems to be common for mathematics students but is not mentioned by computer science students. Computer science students mainly do assignments and projects in their clusters. An interesting approach is that when any member of a group is struggling particularly with certain topics an emergency call is sent and the group meets on campus on request. If no one can solve the problem one member is delegated to contact the lecturer for help. YouTube videos as a learning resource features strongly and PLN clusters often share links to useful videos on a particularly difficult topic. This way of dealing with problems can account for the phenomenon that has been discussed locally among lecturers (informally) of experiencing fewer office visits, especially in mathematics. It is surmised that the umbilical cord of support between lecturer and students is weakening, a premise worth investigating.

Mathematics students experience the added problem of dealing with the symbolic representation of mathematics. Students commonly photograph their notes and problems and communicate these, especially when they need to exchange a significant amount of mathematics. Posing a problem on WhatsApp could include text such as “x to the power of 2,” “x^2” or “sin theta” while creativity comes into play with personalized names for symbols such as “pig’s tail” and “devil’s fork” for sigma and psi. Some mobile phones have a “fancy characters smiley pack” that comes in useful. The odd cluster finds mathematics communication via mobile phones, Facebook and WhatsApp just too cumbersome and use technology only for organisational purposes. Phoning about mathematics problems seems to be something of the past. Computer science students do not experience this type of obstacle.

**Role of an LMS**

Mathematics students make use of an LMS running on the Blackboard platform, while computer science students make use of a locally developed learning management system. It was clear that an LMS as a tool for learning is not rated highly, especially among mathematics students. Mathematics students exclusively use the LMS for obtaining information, memos, lecture notes and marks.

I try to avoid it as much as possible, the amount of worthwhile notes and discussions posted is disappointing. It makes more sense to personally discuss anything work related if at all possible.

Such negativity could also be a reflection on teachers who do not utilize the learning possibilities of the LMS optimally.

Computer science students are more prone to make use of the discussion tool of the LMS (eight of 17 respondents). They feel that computer science courses are their most difficult and they interact with questions posted on the discussion board, if not by posting then by observing.

It becomes very useful when you have a question to ask and its possibly been answered already on the discussion board.
Benefits

Both groups of students are generous in listing benefits of PLN clusters. Enhanced social skills, better grasping of concepts, motivation, support and understanding are among the frequently mentioned benefits by both mathematics and computer science students.

Motivation

The motivation provided through the cluster interaction is repeatedly mentioned as a benefit of PLN clusters on a personal level as well as knowing there are other people with similar concerns, “knowing that I am not alone in my struggle.” In a cluster you will find

- A person motivating you to complete assignments
- People that know how it feels to study
- More minds to solve the problem

Social support and people skills

One of the main benefits of belonging to a PLN cluster, mentioned by both groups, is getting support from caring friends and most of all a sense of belonging. Working in a cluster is perceived to strengthen friendship ties; it provides social enrichment and a comforting feeling that one has a reliable and trustworthy group of people around you, getting help and giving help. It offers a secure environment where you share what you know and refine your knowledge.

Both groups mention people skills as a gain from PLN interaction. Learning to be patient, to work together with other people, to explain to others, are some of these benefits as well as coming to know that “it is far more pleasant to study with fellow students than to study alone.” It also enhances discipline, as members cannot afford to be distracted.

- I grasp aspects of math faster and it has improved my marks as well as my people skills.
- Working together enhances social skills

Communication opportunities

Another benefit, mentioned by a mathematics student, is being able to speak your mind without the fear of sounding like a fool. This thread also appears among computer science responses. One student feels that his understanding improves because they collectively argue about what a question might mean whilst another says that “bad ideas get weeded out quickly” with no hard feelings experienced. The benefit that clusters offer in terms of immediate feedback on attempts and issues is mentioned, especially by computer science students. They equate the group’s shared experience to better quality of work.

Mathematics students claim that interaction improves your communication skills and helps you acquire problem solving skills. Most clusters require of members to prepare before the meeting and it therefore forces them to keep up to date with the work. In addition, meetings offer more contact time with the subject.

- I am usually the one that explains to the others but I find that this consolidates my own knowledge.

Exposure to different perspectives

Exposure to multiple points of view and different perspectives are mentioned by many respondents as a gain from the cluster interaction. Someone else’s perspective could be the key to finding a solution to a problem. Computer
Science students mention particular benefits regarding programming assignments obtained within a pleasant environment:

It helps a lot to hear other peoples’ opinions and see different perspectives. IT has a lot of different ways to do the same thing, thus it is very beneficial to come into contact with different way. IT aims to program in the most efficient way and other people in the group might have a more efficient way than me.

You see the problem you’re facing through a different perspective and light heartedness when difficult questions arise (this is very important as it gives you the will to continue).

Mathematics students mention the benefits of exposure to different approaches to problem solving. One member may follow a more algebraic approach whereas another may follow a more visual approach and they can learn from each other. The following quotes illustrate the practical benefits:

Sharing proofs is quite useful; sometimes they require an approach that just doesn’t occur to you, but that might to someone else.
I would have had to work much harder in order to pass if I did not work with them.
We learn from how other people think.

Other practical benefits

More practical benefits include access to class notes, study notes, information on dates, explanations from group members and sharing of resources (“even if it is just a useful YouTube video”). Respondents also mention gains such as tips on topics dealt with in practical classes from members who may have attended the practical class earlier in the week as well as a way to catch up on notes, hand-outs and announcements that you may have missed. Tangible benefits are mentioned in terms of improved marks and using meetings as an evaluation of how well you understand the work.

Conclusions

This study set out to investigate the phenomenon of spontaneous PLN clusters among two groups of students, mathematics and computer science students, with the latter being more technology inclined. The overwhelming impression is of how important these groups are in the academic lives of these students. They speak with the pride of the interaction; they take ownership of their clusters and are vocal in describing the benefits that these clusters offer. PLN clusters are truly at the heart of student learning, they are little researched gems in the landscape of learning, and the importance of these should not be underestimated. It is a component of academic life that teaching staff have no input into and yet the functioning thereof realizes the ideals of independent learning within a collaborative, technology enhanced environment, while and stimulating discussion on the work, aspects of learning that lecturers have strived to instill for many years.

Commonalities and differences between the two groups were noted. Computer science students have richer mashups of activities, leaning more towards digital communication as could be expected. Yet the importance of face-to-face communication among these students came as a surprise. There is evidence of face-to-face meetings not just happening by chance but being arranged for a definite purpose. Mathematics clusters are more prone to face-to-face communication. Yet it is the similarities rather than the differences between the two groups that are striking. Both sets of clusters have the academic interest of the members at heart and cluster activities centre around this aspect.

There is no doubt that the main aim of the PLN clusters is to advance learning within a supported environment with the social aspect playing an integral but secondary role. Yet there is strong emphasis on the personal gains resulting from belonging to a PLN cluster. Post hoc user requirements can be formulated from the gains expressed by students in PLN clusters.

• Sense of belonging. PLN clusters offer peace of mind on a personal level because of belonging to a circle of caring people who are willing to assist when required, also resulting in motivation, support and understanding.
• **Freedom of expression.** It is a protective environment amongst peers where a member can express himself freely and conflicting opinions are dealt with in a mature manner without fear.

• **Exposure to different perspectives.** When mastering a subject different perspectives deepen understanding, an opportunity available in PLN clusters.

• **People skills.** Functioning successfully in a PLN cluster cultivates skills that will be of value both in a work and personal environment.

• **Communication opportunities.** Constant communication is possible and it is noticeable that the almost permanently available helpline is one of the most valued features. It is an effective TEL environment.

PLN clusters are exemplary in answering in the learning requirement needs of the 21st century. It complies to Web 2.0 concepts and encompasses the three core elements of personalisation, participation and knowledge-pull of the 3P model (Chatti et al., 2010b). Learning within these clusters support the notion that “...the most resilient, adaptive, and effective learning involves individual interest as well as social support to overcome adversity and provide recognition” (Ito et al., 2013, p. 4). Collaborative learning has certainly existed for a prolonged period but has gained impetus in the form of PLN clusters with current technology advances. The role of the lecturer with respect to PLN clusters is undefined and his/her support function seems to be diminishing, an aspect that should be viewed positively.

The study has not set out to determine how prevalent PLN clusters are, proposed as a future project, as is expanding the study to other universities. There certainly would be individuals who do prefer to work totally independently and there could be clusters that disintegrate for whatever reasons. Investigation of these phenomena is also left for future studies. The main finding of the paper is that PLN clusters have value for student learning, in both the two fields investigated. The activities within the PLN clusters differ somewhat between the two groups but the structure and way of functioning are essentially the same.

The formation of PLN clusters is evidence of spontaneous formation of knowledge-pull communities, separated from the formal academic programme. Life-long learning is at the order of the day and the longevity of the PLN clusters is testimony to crossing the boundaries of courses and semesters.

The tension between individual independence and collective cooperation has been debated in literature (Wagner, 1995). Paulsen (2003) claims that online education can foster both freedom for the individual and group cooperation and he uses the term collective freedom. The fact that participation in PLN is voluntary, it is initiated by the students themselves and not a forced activity may contribute to a measure of collective freedom for students.

The increasingly large groups that universities are dealing with make structured collaborative learning practices difficult to conduct. Yet the virtues of collaborative learning and the skills that are fostered by it are exactly those observed in the PLN clusters. PLN clusters are part of a larger phenomenon, defined as the “open movement” of which Couros (2010) says:

> “The open movement is an informal, worldwide phenomenon characterized by the tendency of individuals and groups to work, collaborate and publish in ways that favour accessibility, sharing, transparency and interoperability” (p. 110).

Our study has shown that the PLN clusters observed also comply with all six elements that Kaufman, Sutow and Dunn (1997) identified for successful collaborative learning namely positive interdependence, social skills, face-to-face verbal interaction, individual accountability, group processing, and appropriate grouping. There is also evidence of the presence of the four key factors that define a sense of community as stated by McMillan and Chavis (1986) namely membership, influence, fulfillment of individuals’ needs and shared events and emotional connections.

The study provides a case study example of personal learning environments of which Attwell (2007) says:

> “The space takes a learner-centric approach and is characterized by the use of a set of services and tools that belong to and are controlled by the learners. Rather than integrating different services into a centralized system, the idea is to provide a learner with a plethora of different services and hand over control to the learner to select, use, and mashup the services the way he deems fit.”
The importance of this study lies in being exposed to the voice of students that function within personal learning network clusters and taking cognisance of the inner workings of these clusters. The benefits of belonging to a PLN cluster are notable and should be advocated to students at the onset of their academic careers.

References


Couros, A. (2010). Developing personal networks for open and social learning, Emerging Technologies in Distance Education, 6, 109-128.


Make E-Learning Effortless! Impact of a Redesigned User Interface on Usability through the Application of an Affordance Design Approach

**Hyungjoo Park and Hae-Deok Song**
Department of Education, Chung-Ang University, Heukseok-ro 84, Dongjak-gu, Seoul, Korea // yungjoo@live.cau.ac.kr // hsong@cau.ac.kr

*Corresponding author

(Submitted April 7, 2014; Revised January 6, 2015; Accepted February 23, 2015)

**ABSTRACT**
Given that a user interface interacts with users, a critical factor to be considered in improving the usability of an e-learning user interface is user-friendliness. Affordances enable users to more easily approach and engage in learning tasks because they strengthen positive, activating emotions. However, most studies on affordances limit themselves to an examination of the affordance attributes of e-learning tools rather than determining how to increase such attributes. A design approach is needed to improve affordances for e-learning user interfaces. Using Maier and Fadel’s Affordance-Based Design methodology as a framework, the researchers in this study identified affordance factors, suggested affordance design strategies for the user interface, and redesigned an affordable user interface prototype. The identified affordance factors and strategies were reviewed and validated in Delphi meetings whose members were teachers, e-learning specialists, and educational researchers. The effects of the redesigned user interface on usability were evaluated by fifth-grade participating in the experimental study. The results show that affordances led users to experience positive emotions, and as a result, use the interface effectively, efficiently, and satisfactorily. Implications were discussed for designing strategies to enhance the affordances of the user interfaces of e-learning and other learning technology tools.

**Keywords**
E-learning content, User interface, Usability, Affordance-Based Design methodology

**Introduction**
Studies on affordances in the field of e-learning have mainly examined the affordance characteristics of certain e-learning technologies. These studies looked at affordance attributes such as accessibility, entertainment exchange, information repositories in asynchronous video conferencing tools (Krauskopf, Zahn, & Hesse, 2012), multimedia access and collection, communication, and representation of thought and knowledge in handheld devices (Song, 2011), enhanced spatial knowledge representation for experiential learning, improved contextualization of learning and richer effective collaboration learning for 3-D virtual environments (Dalgarno & Lee 2010), hints and tips for solving problems related to blogs (Robertson, 2011), and educational affordances of ubiquitous learning environments (Fu, Chu, & Kang, 2013; Tan, Lin, Chu, & Liu, 2013). However, very few studies on the development of affordable e-learning technologies have been conducted. One of the reasons for this is a lack of design guidelines for improving the affordances of e-learning user interfaces.

Many user interfaces have been developed based on structured and iterative methodologies. These methodologies consist of a series of processes involving analysis, user task analysis, heuristic evaluation, formative evaluation, and summative comparative evaluation (Shneiderman, Plaisant, Cohen, & Jacobs, 2009). However, this functional design approach does not pay attention to affordances, properties of an object that induce specific behaviors on the part of users, and is therefore limited in that intended affordances are not systematically considered in the design process. Given that affordances are complex to identify and difficult to induce through the design process, an alternative design method is needed for improving affordance values. To address this issue, Maier and Fadel (2009) suggest Affordance-Based Design (ABD), which helps designers to simultaneously analyze users’ needs, consider the affordance requirements of learning tasks, and identify the affordances of artifact components such as technological components simultaneously during the design process. This method has been actively used in reengineering architecture to improve affordance values. However, few studies have identified and improved affordances in the e-learning field.

Studies on cognitive information processing theories indicate that our cognitive systems have limited capacity and thus, learners must select those that best match their goals (Clark & Mayer, 2011). Employing an Affordance-Based
Design method is expected to help designers to select a minimalist approach that avoids overloading working memory by identifying and systematically increasing affordances that are pertinent to a specific technology for targeted learning tasks. Thus, the purpose of this study was to make e-learning user interface affordable with affordance support. To achieve this purpose, design strategies were suggested by looking at the improvement in the usability of an existing e-learning user interface from an affordance perspective. The effects on pedagogical usability were then examined.

Theoretical framework

Usability of the e-learning content user interface and affordances

A user interface affects the efficient use of e-learning content because it functions as an information channel that mediates the relationships between users and artifacts. “Usability” is the degree of ease with which the system can be used and with which it promotes learning and, thus, is important in user-centered interface design (Nielsen, 1993). Given that a user interface interacts with users through their direct and cognitive perceptions (McGrenere & Ho, 2000), a critical factor to be considered for improving the usability of an e-learning content user interface is to make it more “affordable”; that is, to enhance the affordances that it provides. When users perceive the user interface as more affordable, they are more likely to actively use it to achieve learning goals. Learning can be a real emotional challenge. In fact, many aspiring online learners fail to progress beyond their first attempt to use unfamiliar software. Thus, there is a need to enhance positive emotion during e-learning. Designing user-friendly interfaces can help learners better interact with e-learning content.

Gibson (1979) used the term “affordances” to describe the properties of possible behaviors by which an agent (e.g., human being, animal) interacts with its environment (e.g., the world, a society). The environment provides the agent or animals with directly perceivable information that is valuable and meaningful. These pieces of information can be referred to as “affordances.” Affordance characteristics include (a) complementarity, because an affordance expresses a relationship between the artifact and the user, (b) imperfection, because the affordances of the artifact can differ with respect to different users, (c) polarity, because affordances can either be positive or negative, (d) multiplicity, because an object possesses multiple affordances, and (e) quality, that the artifact should produce the desired results or behaviors (Maier, 2011).

Norman (1988) presented the concept of “perceived affordances,” further developing the idea proposed by Gibson (1979). According to Norman, affordances exist independently of a user’s perceptions, but behavior is brought about by the user’s perceptions. Thus, the existence of affordances in itself has no meaning; rather, affordances become meaningful when they are perceived. However, an agent’s perception is possible only when special properties, or affordances, are present in the environment (McGrenere & Ho, 2000).

Hartson (2003) further divided Gibson’s affordances into physical, cognitive, and sensory types. “Physical affordances,” as initially described by Gibson, help users to interact physically with an interface. “Cognitive affordances,” which are similar to Norman’s perceived affordances, help users to understand something. Finally, “sensory affordances” help users to perceive cognitive or physical affordances using their senses. Sensory affordances may exist independently or coexist with cognitive and physical affordances, helping to induce user behaviors through interaction (Hartson, 2003; Song & Park 2009). These categories help designers to capture important affordance factors in designing user interfaces.

Increasing the affordances of an e-learning user interface with Affordance-Based Design

Attempts to improve affordances were originally carried out in the field of ecology but are now being made in the areas of spatial, environmental, and information systems (Hartson, 2003; Maier, 2011; McGrenere & Ho, 2000; Norman, 1988). Despite the importance of improving affordances, very few design approaches have been suggested. One important challenge for designers is to consider different affordance values. To address this issue, Maier and Fadel (2009) proposed an Affordance-Based Design (ABD) methodology. ABD focuses on designer-user interaction, conceptualizing users as the agents of the behaviors associated with the artifacts being designed. In contrast,
functional design focuses on what a designer wants the artifact to do and thus is effective in designing artifacts that perform intended tasks (Maier, 2011). Because function refers to “the general input/output relationship of a system whose purpose is to perform a task” (Pahl & Beitz, 1996), it is limited in that users cannot be involved in the design process. In addition, it does not consider unexpected influences. Affordance-Based Design can provide designers with a theoretical foundation that looks at design components holistically by considering both function and user together and by taking into account not only desired affordances but also undesired affordances. How then, can we design or redesign an artifact with this design method?

Affordance-Based Design begins with motivation, such as a need for a novel idea or a redesign. Once such a motivation is seen to exist, the Affordance-Based Design process involves the following steps: (a) determining the affordances that the artifact should and should not have, (b) conceiving the artifacts’ overall architecture, (c) analyzing and redefining the affordances of the components, (d) selecting a preferred architecture that will enhance desired affordances and minimize undesired affordances, and (e) determining affordance structures and design affordance components (Maier, 2011). According to Maier, the first four steps are used to conceive of and select affordance artifacts in designing a new architecture. The last step, determining affordance structures and design affordances, should be considered in redesigning an existing architecture. An Affordance Structure Matrix (ASM) (Maier & Fadel, 2009) can be a useful tool in redesigning a user interface. It allows designers to identify and prioritize the important affordance factors that need to be considered for each object. Specific affordance design strategies can then be established based on this information. Maier and Fadel applied this method in reengineering architecture and suggested that the method is effective in improving affordance values. However, no attempt has been made to do so in redesigning e-learning user-interfaces. Thus, the following research questions, encompassing two distinct phases, were generated: How can an e-learning user interface be redesigned through the application of an Affordance-Based Design approach? What is the impact of the redesigned user interface on effectiveness, efficiency, and satisfaction?

**Phase 1: Redesigning user interface through an Affordance-Based Design**

**Context**

The e-learning content used in this study was from the EDUNET Cyber Home Learning System, Korea’s largest education portal, which supports the distribution and utilization of a diverse range of high-quality educational content. The topic selected was “What is the Internet and Information Society?” under the subject heading “Information Communication and Technology.” This topic was chosen from among the twenty topics listed under this subject heading because students’ interest in the topic is high and the corresponding e-learning user interface included a variety of operation methods, navigation buttons, messages, help guides, caption functions, screen layouts, and links to induce specific learning behaviors. The e-learning user interface was examined to see how affordable it was in aiding students to perform intended learning actions and how useful it was helping them to learn content.

**Research procedure**

The e-learning user interface was redesigned through an Affordance-Based Design approach using Maier and Fadel’s (2007) methodology as a framework. First, as shown in Table 1, affordance structures were identified. A literature review suggested important components of e-learning content and their associated affordance factors. The affordance factors then were determined through Delphi meetings. Eight experts (two professors in Instructional Technology, two professors in Information Technology, two e-learning design specialists, one e-learning content-development expert, and one elementary-school teacher) participated as Delphi experts. At the beginning of the Delphi meeting, an introduction explained the concepts of affordances and Affordance-Based Design to the experts. Once they understood these concepts, they participated in an expert heuristic evaluation to identify the e-learning content components and e-learning affordance factors and prioritize the identified affordance factors; a content validity index was used to prioritize the identified affordance components. Second, affordance design strategies were suggested. Finally, the design strategies were incorporated into the development of a new user interface prototype.
Table 1. Design steps, purpose, and numbers of participants in the study

<table>
<thead>
<tr>
<th>Design step</th>
<th>Purpose</th>
<th>Participants</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying affordance factors</td>
<td>Identify components for targeted e-learning user interface</td>
<td>Researchers</td>
<td>Literature review</td>
</tr>
<tr>
<td></td>
<td>Identify their associated affordance factors</td>
<td></td>
<td>ASM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Delphi</td>
</tr>
<tr>
<td>Identifying affordance design</td>
<td>Evaluate and prioritize affordance factors</td>
<td>Eight experts with</td>
<td>Delphi</td>
</tr>
<tr>
<td>strategies</td>
<td></td>
<td>specialties related to the</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>e-learning content</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>interface</td>
<td></td>
</tr>
<tr>
<td>Developing the user</td>
<td>Create a new user interface prototype</td>
<td>Researchers</td>
<td>Experimental</td>
</tr>
<tr>
<td>interface prototype</td>
<td></td>
<td></td>
<td>study</td>
</tr>
</tbody>
</table>

Findings

Affordance structures

The initial step in Affordance-Based Design, identifying affordance structures, includes two main activities: identifying the components of the targeted e-learning user interface and identifying their associated factors. Most important in this step is how clearly specific e-learning components can be associated with particular affordance factors so that the strengths and weaknesses of each component can be evaluated in terms of those factors. To help with this process, an ASM was used. Table 2 shows an example of the application of an ASM to e-learning content. First, the components of an artifact are placed in the upper portion of the ASM, where the components that are able to operate from the artifact must be separated and arranged. For this purpose, the components of the EDUNET Cyber Home Learning user interface were analyzed and sorted. The Delphi experts examined and isolated the important components of the e-learning user interface. The resulting components included overall mood, titles, lists, location of the learner, learning content, media feature controls, screen changes, captions, tips, behavior-inducing messages, and learning support tools.

Table 2. Affordance system matrix

<table>
<thead>
<tr>
<th>Affordance factors</th>
<th>E-learning Components</th>
<th>Overall mood</th>
<th>Title</th>
<th>Lists</th>
<th>Location of the learner</th>
<th>Learning content</th>
<th>Media feature control</th>
<th>Screen changes</th>
<th>Captions</th>
<th>Tips inducing messages</th>
<th>Behavior-Learning support tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accurate representation</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>Convenient operation</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>Effective layout</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>Prevention of errors</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>Categorized contents</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>Induction of interactions</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>Immediate reaction</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>Provision of tips</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>Appropriate visual effects</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>Appropriate auditory effects</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>Abnormal termination</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>Invalid connection</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>Lack of movement</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>Complex operation</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>Inability to perform operation</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>Unresponsiveness</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>
Next, the affordance factors are placed on the left side of the ASM as well. The affordance classification criteria suggested by suggests by Hartson (2003) are used in this study. Hartson argued that three types of affordance components exist: cognitive, physical, and sensory affordances. The Delphi experts generated both positive and negative specific affordance components in terms of these three types. Content validity indices for the relevance of these affordance components were calculated by three experts in the field of e-learning. The components with scores of over CVI .90 were selected as the final components. The cognitive type of affordance included five components: accurate representation, categorized content, provision of tips, inaccurate content, and confusion. The physical type of affordance included nine components: convenient operation, induction of interaction, immediate interaction, accurate representation, categorized content, provision of tips, inaccurate content, and confusion. The sensory type of affordance included three components: effective layout, appropriate visual effects, and appropriate auditory effects.

**Affordance design strategies**

To identify affordance design strategies, the affordance factors of the e-learning components first needed to be evaluated and prioritized. As shown in Table 2, the identified components were marked differently on the ASM according to whether their affordance values were positive or negative. The positive artifact-user affordance factors were marked with a “+”, and the negative artifact-user affordance factors were marked with a “−”. Next, priorities for affordances were determined based on the calculations for marked affordance factors. The sum and percentage of components marked with a “+” and the sum and percentage of those marked with a “−” were then calculated and noted. The number of “+” marks was added up and noted under “sum of + affordances,” and the percentage of affordances was then calculated. The “percentage of + affordances” and the “percentage of − affordances” were the percentage values that were obtained by dividing the “sum of + affordances” and the “sum of − affordances” by the “total +/- affordances.” The “percentage difference” is the difference generated by subtracting the “percentage of − affordances” from the “percentage of + affordances.” The affordance design factors had to be determined based on the percentage difference between positive and negative affordances; components with a small positive value or a negative value would need to be improved (Maier, 2011; Maier & Fadel, 2009). Once the affordance factors were evaluated, design strategies needed to be developed to improve these components with a negative value. As shown in Table 2, the e-learning content component with the lowest percentage difference was the “media features control,” “location of the learner,” “list,” “screen changes,” “learning support tools,” “learning contents,” and “title” came next. Affordance design strategies needed to be developed to reduce negative aspects of the interface and to make the use of the identified components more effective and efficient.

Affordance design strategies were suggested for the affordance factors with a percentage difference of 50% or less. A summary of the design strategies suggested to improve affordances for the identified e-learning components is provided in Table 3. For example, for the media features control, strategies for reducing negative-affordance factors such as unexpected termination had to be developed. For this reason, strategies such as “ensure that the learning window does not close unexpectedly,” “ensure error-free operation,” and “design buttons so that learners can easily perceive and operate them” were suggested to avoid negative affordances, whereas “ensure that operation methods are recognized as similar to those in the real world” and “ensure immediate processing after operation” were suggested as positive strategies.
<table>
<thead>
<tr>
<th>E-learning component (%)</th>
<th>Affordance factors</th>
<th>Direction of affordance</th>
<th>Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Media feature control (-11%)</td>
<td>Convenient operation</td>
<td>+</td>
<td>Ensure that operation methods are recognized as similar to those in the real world</td>
</tr>
<tr>
<td></td>
<td>Prevention of errors</td>
<td></td>
<td>Ensure immediate processing after operation</td>
</tr>
<tr>
<td></td>
<td>Immediate reaction</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Provision of tips</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unexpected termination</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Invalid connection</td>
<td></td>
<td>Ensure that the learning window does not close unexpectedly</td>
</tr>
<tr>
<td></td>
<td>Lack of movement</td>
<td></td>
<td>Ensure error-free operation</td>
</tr>
<tr>
<td></td>
<td>Complex operations</td>
<td></td>
<td>Design buttons so that learners can easily perceive and operate them</td>
</tr>
<tr>
<td></td>
<td>Inability to perform operations</td>
<td></td>
<td>Design buttons to operate as learners expect them to</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ensure valid connection to learning content when playing learning media</td>
</tr>
<tr>
<td>Overall mood (0%)</td>
<td>Accurate representation</td>
<td>+</td>
<td>Clearly present the overall layout of the e-learning content by area</td>
</tr>
<tr>
<td></td>
<td>Effective layout</td>
<td></td>
<td>Ensure that the overall layout is clear and simple</td>
</tr>
<tr>
<td></td>
<td>Categorized contents</td>
<td></td>
<td>Ensure visual comfort</td>
</tr>
<tr>
<td></td>
<td>Appropriate visual effects</td>
<td></td>
<td>Simplify components</td>
</tr>
<tr>
<td></td>
<td>Complex operations</td>
<td></td>
<td>Ensure that there are no operations errors</td>
</tr>
<tr>
<td></td>
<td>Inability to perform operations</td>
<td></td>
<td>Structure content for easier identification</td>
</tr>
<tr>
<td></td>
<td>Inaccurate content</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lists (11%)</td>
<td>Accurate representation</td>
<td>+</td>
<td>Present the contents by category and ensure that the layout is effective</td>
</tr>
<tr>
<td></td>
<td>Convenient operations</td>
<td></td>
<td>Represent the e-learning content consistently throughout and simplify the layout for easy comprehension</td>
</tr>
<tr>
<td></td>
<td>Effective layout</td>
<td></td>
<td>Ensure valid connection between links and content</td>
</tr>
<tr>
<td></td>
<td>Immediate reaction</td>
<td></td>
<td>Ensure that the transition is smooth and error-free when the learner moves from one learning progress stage to another</td>
</tr>
<tr>
<td></td>
<td>Appropriate visual effects</td>
<td></td>
<td>Accurately present information on the learner’s location</td>
</tr>
<tr>
<td></td>
<td>Invalid connection</td>
<td></td>
<td>Arrange lists for easy use</td>
</tr>
<tr>
<td></td>
<td>Lack of movement</td>
<td></td>
<td>Ensure that the lists are arranged by category for easy identification of the learning hierarchy</td>
</tr>
<tr>
<td></td>
<td>Unresponsiveness</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screen changes (11%)</td>
<td>Convenient operation</td>
<td>+</td>
<td>Ensure that operation methods are recognized as similar to those in the real world</td>
</tr>
<tr>
<td></td>
<td>Effective layout</td>
<td></td>
<td>Ensure immediate processing after operation</td>
</tr>
<tr>
<td></td>
<td>Prevention of errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Captions (14%)</td>
<td>Learning support tools (14%)</td>
<td>Learning content (33%)</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td></td>
</tr>
</tbody>
</table>
| - Invalid connection  
- Lack of movement  
- Complex operations  
- Inability to perform operations | - Accurate representation  
- Effective layout  
- Prevention of errors  
- Provision of tips  
- No movement  
- Inability to perform operation  
- Unresponsiveness | - Accurate representation  
- Categorized contents  
- Induction of interactions  
- Provision of tips  
- Appropriate visual effects  
- Appropriate auditory effects  
- Lack of movement  
- Unresponsiveness  
- Inaccurate content  
- Confusion |
| + Present messages that induce learner behaviors  
- Locate learning support tools at places where they can be easily seen by the learner  
- Provide detailed help guides that are easy to understand | + Present messages that induce learner behaviors  
- Locate learning support tools in places where they can be easily seen by the learner  
- Provide detailed help guides that are easy to understand | + Present accurate learning content without errors  
- Present appropriate amount of information per screen  
- Structure and present contents around core contents  
- Provide appropriate interactions when presenting learning content  
- Provide feedback after interaction  
- Ensure error-free operation when learners interact with content  
- Ensure that screens are designed with appropriate amounts of content  
- Provide correct learning content without errors |
Title (50%) • Accurate representation + • Use a title that is related to the learning content • Effective layout • Place items so that learners will easily notice them • Categorized contents • Ensure that titles include core contents • Inaccurate content - • Content is phrased accurately without errors • Make learning hierarchy easy to identify by title

User interface prototype

Based on the results of the suggested affordance design strategies, a prototype for a new e-learning content user interface was developed by programmers into an actual prototype platform to be used in the usability evaluation stage. New e-learning content was incorporated from the unit on information technology training in EDUNET’s Cyber Home Learning System. This new content was developed by employing Affordance-Based Design strategies. The main learning task was related to the safe use of the Internet.

The suggested affordance design strategies were included in the redesigned e-learning user interface to improve the identified affordances for the interface. Figure 1 shows sample screens from both the existing Cyber Home Learning content and the newly designed content. Many changes were incorporated into the new user interface. The structure of the menus was changed. The user interface of the existing EDUNET content has two main menus: (A) Index, and (B) Status Progress Bar. The Index provides a list of learning contents and can be assessed by clicking on the appropriate menu. The Status Progress Bar is located at the bottom of the screen and includes page-navigation, multimedia-play, and page-viewing options.

The new re-designed user interface has a different menu structure, with (A) Index, (B) Status Progress Bar, and (C) Learning Support Tools. All of the menus were located on the right side of the screen in order to clearly present the overall layout of the e-learning content by area and create an effective layout, one of the affordance factors of the Overall Mood e-learning component. Learning content in the new Index (A) is arranged by category for easy identification of the learning hierarchy to reduce confusion, one of the negative affordance factors of the “Lists” e-learning component. The structure of the new Status Progress Bar (B) was redesigned so that learners can easily perceive and operate it, reducing complex operations, one of the negative affordance factors that affect learners using the Media Features Control. Script view and audio control functions were also added to help learners process information efficiently. The Learner Support Tools (C) were newly added to ensure proper connections with the learning support tools and reduce the incidence of invalid connections, one of the negative affordance factors of the Learning Support Tools.

Old design: Existing user interface

New design: Redesigned user interface

Figure 1. Sample screen showing the existing and redesigned user interfaces
Phase 2: Impact of redesigned user interface on usability

Participants

Evaluation of the user interface prototype was carried out through an experimental study that allowed the researchers to determine which affordance features of the user interface were improved. For this purpose, the evaluation of the user interface prototype, 171 fifth-grade student participants were recruited from five intact classrooms of 34-36 students each at a metropolitan elementary school located in Seoul, Korea. The participants were 48 % female and 52 % male. Given that there were 23 items included in the usability test, the number of participants was more than five times the number of test items and thus should be sufficient for further statistical testing. The pre-test score results from the national achievement test showed no significant difference between the participating classes.

Procedure

The study was conducted in the students’ regular classroom during an Information Technology Training lesson. Two types of e-learning content (existing, new) were used to examine the effects of the proposed affordance design strategies. Students used the interface while studying e-learning content. The experiments were conducted over three class hours in an elementary-school computer room. After completion of the e-learning content, the learners in both groups were asked to participate in a web-based online survey that assessed usability through the perceived effectiveness and efficiency of the interface, and students’ satisfaction with it.

Assessment instrument

To assess usability in a web-based online survey, three factors were included in the test instrument: effectiveness, efficiency, and satisfaction. For the effectiveness and efficiency factors, related question items were modified from those used by Costabile, Marsico, Lanzilotti, Plantamura, and Roselli (2005) and by Silius and Tervakari (n.d.). For the satisfaction factor, relevant question items were modified from Nokelainen’s (2006) usability assessment questions. In an effort to enhance readability and consistent interpretation by the intended audience, the survey instrument was pilot-tested with a sample of seven children in the fourth grade. Some terminology that was perceived as difficult by these younger children was revised. The final survey included 23 items scored on a 5-point Likert-type scale: eleven items referred to effectiveness (e.g., “Important information is emphasized and easily seen”); “The contents are organized around core concepts”; “The learning content is clearly organized”; “Buttons with similar functions are grouped together”; “Tips are provided, and asking questions is allowed”). Eight questions referred to efficiency (e.g., “The tools that are necessary for learning are easy to find”; “Various methods of communication are provided”; “It is possible to immediately move to the desired learning content”; “There are no broken hyperlinks”). Four items referred to satisfaction (e.g., “The learning content was beneficial”; “I would like to use this type of e-learning again”; “I would like to introduce this type of e-learning to my friends”). This assessment instrument had a Cronbach’s α coefficient of .92.

Findings

Students in the affordance-based e-learning content group had a significantly higher usability score ($M = 4.18, SD = .45$) than students in the existing e-learning content group ($M = 3.63, SD = .77$); $t(110) = -5.42, p < .01$. Specifically, students in the affordance-based e-learning content group had a significantly higher effectiveness usability score ($M = 4.37, SD = .47$) than those in the existing e-learning content group ($M = 3.72, SD = .85$); $t(105) = -5.90, p < .01$, and those in the affordance-based e-learning content group had a significantly higher efficiency score ($M = 3.94, SD = .60$) than those in the existing e-learning content group ($M = 3.54, SD = .80$); $t(130) = -3.53, p < .01$. Finally, students in the affordance-based e-learning content group had a significantly higher satisfaction score ($M = 4.14, SD = .71$) than those in the existing e-learning content group ($M = 3.57, SD = .87$); $t(138) = -4.60, p < .01$. 

193
Conclusion

With the ubiquity of advanced technology in the early 21st century, most people may have no particular difficulty in navigating user interfaces. However, when learners perceive an interface as affordable, they are likely to operate the interface effortlessly and, therefore, to show desired learning behaviors. Thus, the purpose of this study was to improve redesign affordance value of an e-learning contents user interface and to examine the impact of redesigned user interface on usability.

Despite the importance of affordance design in user interfaces, most studies on affordances have only examined the affordance characteristics of e-learning tools (Dalgaro & Lee 2010; Krauskopf, Zahn, & Hesse, 2012; Robertson, 2011; Song, 2011). A design approach that will help to improve the affordance attributes of e-learning programs is also needed. Improvements will be possible when the properties that influence learner behavior are systematically considered from an affordance perspective. This study was conducted in two phases. In the redesign phase, by using Maier and Fadel’s (2009) Affordance-Based Design methodology as a framework, the researchers identified affordance factors and affordance strategies and developed a user interface prototype. In the evaluation phase, the effects of the redesigned user interface prototype on usability were tested. One of the significant contributions of this study was the application of a new design approach that increases affordance values in the field of e-learning. The systematic affordance design approach of this study can provide e-learning instructional designers with a guide to exploring affordances pertinent to the critical components of an e-learning user interface, improving the components of the e-learning user interface by reducing negative impacts on learning and enhancing positive impacts simultaneously. However, the methodology needs to be further validated in the future studies. Presenting new content through a redesigned user interface, as was done in this study, might in itself improve scores, regardless of improved affordance factors. Thus, validity of the methodology used here requires further consideration. In addition, users were not involved in redesign process itself and thus, students own' perspectives might not be fully considered.

The study findings show that the new e-learning interfaces earned significantly higher scores on all three components of usability (effectiveness, efficiency, and satisfaction). This can be explained by the mechanism of affect on cognitive learning. According to Pekrun, Goetz, Titz, and Perry (2002), there are a couple of pathways from a student’s affective conditions to his or her cognitive processing. When individual learners experience positive affect, they are more likely to recall information, process information efficiently, and be intrinsically motivated to carry out a particular task. In fact, students report that the redesigned user interface led them to have positive affect because they could easily locate important information and access tools. In particular, this empathic design could be useful for designing interfaces for children. The results of this study showed that elementary-school students using the new e-learning interface tended to easily grasp important information, clearly understand organized learning content, operate learning support tools without trouble, easily find necessary information, and immediately move to desired learning content. The results also show that these students perceived the use of the new, user-friendly e-learning content as more valuable.

The literature suggests that the knowledge and information processing capability of individuals increase with age from childhood to adolescence (Byrnes, 2001). For instance, children progress from being able to solve simple word problems in Grade 4 to being able to perform mathematical manipulation and reasoning by Grade 12 (Reese, Miller, Mazzeo, & Dossey, 1997). Given the limited capacity for abstract thinking of fifth-grade students like those who participated in this study, Affordance-Based Design may work as a critical vehicle for improving the usability of e-learning content. However, since usability is typically measured by time to task, number of errors, time to learn/relearn, and the subjective experience of users, the implications of this study are limited by its exclusive reliance on users’ subjective perceptions. A variety of usability measures should be employed to validate the findings of future studies.

A critical challenge in increasing affordance values is to identify the relevant affordance factors. The affordance factors of the e-learning user interface were identified at the beginning of the design processes in this study. Three types of affordance classification criteria were used, in accordance with Hartson’s (2003) affordance classification: cognitive, physical, and sensory. These are essential to creating e-learning content because the three types of affordance in this study are directly related to the e-learning process. However, given that more learning-related affordance factors such as educational affordances are being explored in current studies, future studies should look for more specific types of affordances to make e-learning content more effective (Kirshner, 2002). This attempt will
also provide e-learning instructional designers with valid criteria for creating effective learning-oriented e-learning interfaces.

The affordance strategies used to redesign the e-learning user interface in this study had a positive impact on usability, indicating that if such affordances are properly increased, e-learning programs content that encourage more appropriate learner behavior can be developed. The growing use of new learning tools, such as smart-phones, tablets, and e-books, in educational settings increases the need to apply affordance design strategies to various media to allow learners to respond actively to changes in their learning environment. Affordance-Based Design should result in more user-friendly interfaces for new learning tools.

References


MONTO: A Machine-Readable Ontology for Teaching Word Problems in Mathematics

Aparna Lalingkar*, Chandrashekar Ramanathan and Srinivasan Ramani
International Institute of Information Technology, Bangalore // aparna.l@iiitb.ac.in // rc@iiitb.ac.in // ramani.srini@gmail.com

*Corresponding author

(Submitted April 21, 2014; Revised November 26, 2014; Accepted February 23, 2015)

ABSTRACT
The Indian National Curriculum Framework has as one of its objectives the development of mathematical thinking and problem solving ability. However, recent studies conducted in Indian metros have expressed concern about students’ mathematics learning. Except in some private coaching academies, regular classroom teaching does not include problem solving in mathematics, but is limited to mere practice exercises and drills of known exercises. For describing mathematical thinking, Schoenfeld gave a framework containing four components: resources, heuristics, controls and beliefs. Beginning in childhood we develop an ontology for the ideas we learn, and this ontology evolves as we continue learning. Ontologies used for teaching need to incorporate elements of mathematical thinking popularized by problem solving experts. So teaching that makes use of such ontologies of problems, problem solving strategies, and tasks would be beneficial to students. In this paper we identify the gaps in the literature on teaching problem solving, and discuss how and why ontologies can be used for teaching problem solving in mathematics at the high school level. As a proof of concept, we describe the method by which an ontology named MONTO has been created for teaching problem solving in mathematics, and give examples of its use. We describe the MONTO ontology and compare it with some other teaching ontologies described in the literature. We developed and evaluated the MONTO ontology for Surface Area and Volume (3D Solids) problems taught as part of the national curriculum in India, and the results obtained were satisfactory: MONTO was found to be 94% robust against unseen problems in different curricula for the same domain.

Keywords
Teaching problem solving, Pedagogy ontology, Ontology-based student model, Application of ontology, Knowledge representation

Introduction
The Indian National Curriculum Framework (NCF) (NCF, 2005) and the National Council of Teachers of Mathematics (NCTM, 1980) both focus on development of problem solving in school mathematics. A majority of research work done on problem solving in mathematics has been conducted in the US, and yet, according to Trends in International Mathematics and Science Study (TIMSS -2007), East Asian students perform better than the US students (Gonzales et al., 2008). It is true that, over the past decade, a few Indian students have won gold and silver medals in International Mathematics Olympiads; however, studies of students’ mathematical learning conducted in Indian metros show that even students from top schools perform below the international average (EI Reports, 2009). The Program for International Student Assessment (PISA) (PISA, 2012) reported that the mathematical proficiency of students from Indian states such as Tamil Nadu and Himachal Pradesh was lower than the average of countries of the Organization for Economic Co-operation and Development (OECD: http://www.oecd.org/).

A problem is defined as “to search consciously for some action appropriate to attain a clearly conceived, but not immediately attainable aim” (Polya, 1981). For this study, the term mathematical problem refers to a problem that is solved by using mathematical models, formulas, mathematical logic, and rules. In the historical review of problem solving, there is a dichotomy between the terms problem solving and doing exercises. The term problem solving refers to the use of various heuristic strategies, pattern searching, and control functions for selecting the appropriate strategy, whereas doing exercises refers to the use of known procedures and methods (Schoenfeld, 1985). For the scope of this paper, we consider mathematical problems at K-12 level. Problem solving is introduced at school level when students learn word problems including a real world scenario. At that point students experience difficulties because they need to understand the real world scenario, connect it to the mathematical language and convert it into the mathematical model to solve. Hence, the paper focuses on teaching word problems. Despite more than seven decades of work in teaching problem solving (Polya, 1981; Polya, 1946; Schoenfeld, 1985; Silver, 1985; Marshall,
1995; Jonassen, 2011), classroom teaching of solving mathematical problems at school level has remained a great challenge.

The paper is organized as follows. Initially, the theory of mathematical thinking, knowledge and behavior is discussed, drawing from the literature of mathematics education, to give a theoretical framework for the argument. Secondly, the state of the art for teaching word problem solving is discussed, and the gaps are identified. The next section contains a discussion of the methodology for developing ontology and of the proposed ontology, considering one particular domain with an example. The next section compares the proposed ontology with existing teaching ontologies described in the literature. The final section includes an evaluation of MONTO, the conclusion, and possible future work.

State of the art

Mathematical thinking, knowledge, understanding

Once students are taught mathematical problem solving, they should be able to explore patterns and seek solutions to the problems and not just memorize procedures, and should be able to formulate conjectures and not merely do exercises. The implicit objectives of teaching mathematical problem solving at school level – such as development of mathematical thinking, logical thinking, and critical thinking – are expected to be achieved after some years. While teaching mathematical problem solving, Schoenfeld (1985) conducted a study of students and came up with a framework for the analysis of mathematical thinking and behavior that contains four components (Table 1).

<table>
<thead>
<tr>
<th>Resources: Basic concepts and mathematical knowledge</th>
<th>Cognitive aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control: Logic and efficiency in combining resources and heuristics for taking decisions involved in solving a problem</td>
<td>Meta-cognitive aspect</td>
</tr>
<tr>
<td>Belief Systems: One’s opinion and belief about mathematics and mathematical problem solving</td>
<td>Motivational aspect</td>
</tr>
</tbody>
</table>

Table 1. Mathematical thinking framework

<table>
<thead>
<tr>
<th>Resources: Basic concepts and mathematical knowledge</th>
<th>Cognitive aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control: Logic and efficiency in combining resources and heuristics for taking decisions involved in solving a problem</td>
<td>Meta-cognitive aspect</td>
</tr>
<tr>
<td>Belief Systems: One’s opinion and belief about mathematics and mathematical problem solving</td>
<td>Motivational aspect</td>
</tr>
</tbody>
</table>

Table 2. Types of knowledge (Thomas, 1988; Wong et al., 2007; Mayer, 1998)

<table>
<thead>
<tr>
<th>Conceptual Knowledge/ Semantic Knowledge: Knowledge of mathematical facts and relations among the facts</th>
<th>Cognitive aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextual Knowledge/ Linguistic Knowledge : Knowledge of how things work in real world scenarios; Knowledge of story situations</td>
<td>Cognitive aspect</td>
</tr>
<tr>
<td>Procedural Knowledge/ Syntactic Knowledge/ Schematic Knowledge: Knowledge of the strategies for solving the problem; Knowledge of the components of various correct procedures; Knowledge of the syntax or structure of the problem</td>
<td>Meta-cognitive aspect</td>
</tr>
</tbody>
</table>

A student who has been taught mathematical problem solving is strong in analyzing a large amount of quantitative data, uses mathematics in practical ways, and is analytical both in thinking on her own and in examining the arguments put forward by others (Schoenfeld, 1992). An expert problem solver will have very good knowledge of the conceptual connections within the subject and of the procedures for representation and understanding of the problem, and good decision making ability for using the appropriate sets of concepts and procedures for solving the problem. Mayer (1998) and Jonassen (2000) discussed the cognitive, meta-cognitive and motivational aspects of problem solving in general. Providing accurate and appropriate information is the key to building knowledge blocks. Understanding of a mathematical problem depends on several factors, such as conceptual knowledge, contextual knowledge, and procedural knowledge (described in detail in Table 2). Many researchers have studied the understanding and teaching of word problems. Some selected findings are listed in Table 3. The understanding of natural language used in story problems with respect to mathematical language is very important for the understanding of word problems.
Table 3. Findings from previous work

<table>
<thead>
<tr>
<th>Key points</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creation of internal representation in the form of prepositional structure or schema</td>
<td>Mayer (1982); Hinsley et al. (1977)</td>
</tr>
<tr>
<td>Application of rules of algebra and arithmetic to the internal representation to solve a problem</td>
<td></td>
</tr>
<tr>
<td>Difficult and ambiguous language of the word problem leads to confusion among students</td>
<td>Cummins et al. (1988)</td>
</tr>
<tr>
<td>Importance of semantic understanding by students for better problem solving performance</td>
<td>Riley et al. (1984); Carpenter &amp; Moser (1982)</td>
</tr>
<tr>
<td>Studied schema formation of word problems and proposed various schema</td>
<td>Kintsch &amp; Greeno (1985); Bakman (2007); Mukherjee et al. (2007)</td>
</tr>
<tr>
<td>Expert problem solvers perceive “deep knowledge structure” and novice problem solvers perceive “surface knowledge structure”</td>
<td>Schoenfeld &amp; Herrmann (1982)</td>
</tr>
</tbody>
</table>

**Deep knowledge structure:** mathematical principles needed for solving a problem

**Surface knowledge structure:** wording and phrases used in the problem

Training in problem solving changes novices’ perception of knowledge structure from surface to deep

**Cognitive Load:** Iterative information elements which a student needs to use in order to do complex cognitive tasks involved in solving a problem

Sweller (1988)

Multiple representations of graphical, textual, and other problems facilitate learning and reduce the cognitive load

Chandler & Sweller (1991)

Systems designed for tutoring of mathematics:

**ACTIVEMATH:** ontology-based interactive-exploratory learning environment, similar to a hyperbook

Melis et al. (2001);

**Cognitive Tutors:** based on ACT-R theory and use production rules

Anderson (1993);

**Khan Academy:** distributed learning environment and blended learning environment for teaching basic concepts of mathematics and science

Sahlman & Kind (2012)

**Designing problem solving learning environment:** defining problem schema, use of analogy, understanding causal relationships (i.e., semantic and schematic relationships), questioning to support problem solving, modeling of problems, use of meta-cognitive regulations for self-regulated learning, modeling of students

Jonassen (2011)

The understanding of mathematical terms is challenging for students irrespective of the medium or language of instruction. The importance of schema formation in the understanding and solving of word problems is highlighted in the state of the art. For student modeling, one needs to capture a student’s previous knowledge and monitor the ongoing performance for diagnosis of a problem and generation of constructive feedback. Based on the literature review, we have developed a framework given in Table 4 which includes a number of criteria based on knowledge representation and meta-cognitive parameters: controls, modeling of problem strategies, questioning to support problem solving, modeling of problems and modeling of students. Existing systems and schema-based simulations were evaluated against this framework and the gaps were found (please see Table 5). These gaps provided the rationale for further study and led to the research question: “How is ontology to be used for teaching the solving of mathematical word problems at the high school level?”

Table 4. List of parameters and detailed criteria of the framework (Lalingkar et al., 2014)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Representation</td>
<td>Linguistic Knowledge</td>
</tr>
<tr>
<td></td>
<td>Semantic Knowledge</td>
</tr>
<tr>
<td></td>
<td>Syntactic Knowledge</td>
</tr>
<tr>
<td></td>
<td>Multiple Knowledge Representation</td>
</tr>
<tr>
<td>Meta-cognitive Aspects</td>
<td>Controls/ Decision Making</td>
</tr>
<tr>
<td></td>
<td>Modeling of Problem Strategies</td>
</tr>
</tbody>
</table>
Use of ontology

Ontology is a specification of a conceptualization (Gruber, 1993). Conceptualization means structured interpretation of a part of the world in which people think and communicate in the world. Borst (1997) modified the definition to “ontology is a formal specification of a shared conceptualization.” Inclusion of the two words “shared” and “formal” in the modified definition emphasized, respectively, the agreed upon conceptualization of the ontology that can be specified by a group, and machine readability of the ontology. Ontology captures domain knowledge in a generic way and provides a common understanding agreed upon by a group (Gomez-Perez & Benjamin, 1999).

Ontology is usually organized in the form of a taxonomy which has classes, sub-classes, relations, functions, axioms and instances. The closed world and open world are concepts used in ontology. The closed world is the world about which we know everything; so whatever we want to know can be known from the information already available in the ontology. The open world is the world about which some things are unknown to us; still, we can predict or infer from the available information about this world from the ontology. By using the features of ontology (just mentioned) and first order predicate logic, we can make inferences, and we can handle both closed world and open world assumptions. Chandrasekaran and Josephson (1999) mentioned that ontology can represent beliefs, goals, hypotheses, and predictions about a domain, and describe plans and activities. Thus we can create ontology for modeling domain knowledge, problems, problem solving strategies and the tasks that students need to undertake while solving a particular problem and teaching problem solving.

From childhood on, we develop a kind of ontology for the things we learn and the relationships among them in our minds. This ontology evolves as we continue learning. This is the natural way of learning, and so it would be useful to teach by using a combination of various ontologies. Also, while presenting the state of the art of ontology for education, Dicheva et al. (2005) have shown ontology’s usefulness as a tool for pedagogy, knowledge representation and information retrieval, and also as a cognitive tool.

Methodology to create ontology

Noy and McGuinness (2001) have given the life cycle for ontology development (see Figure 1). After defining the problem and selection of the domain and scope for the ontology, we searched ontologies available for mathematics. We were able to find some parts of the ontology used in ActiveMath (Melis et al., 2001) and RamSys (Díez & Gil, 2008). However, the focus of those ontologies was not on teaching of word problems in the selected domain. Hence, we decided to develop MONTO, a machine-readable ontology for teaching word problems in mathematics. MONTO
is described and discussed in the next section. For creating the domain ontology, we first analyzed the domain from the NCERT textbook to enumerate important terms such as Shape, Measurement, Unit, One-Dimensional, Two-Dimensional, Three-Dimensional, etc. These important concepts provided the basis for creating a class-subclass hierarchical structure. Thus, the main classes for the domain, such as Shape, Measurement, Unit, and Formula, were defined. Also, the focus of the application of the ontology was “teaching of word problems,” so other classes that are used for pedagogy and the system were also defined. The most general classes were defined first, followed by the sub-classes of the main classes. This is called the top-down or specialization approach. Defining sub-classes first and then moving to the super-classes is called the bottom-up or generalization approach. Many times people use a combination of both the approaches. While thinking about defining classes and sub-classes, we followed a combination of both the approaches, but while developing the actual ontology in the editor we followed the top-down approach.

After defining the class-subclass hierarchical structure, the properties or slots by which those classes were connected were defined. For instance, the Shape and Measurement class was connected by three properties `hasOneDMeasurement (Shape, Measurement)`, `hasTwoDMeasurement (Shape, Measurement)` and `hasThreeDMeasurement (Shape, Measurement)`. Also, some slots such as data type properties were given restrictions on their values called as cardinality of the slot/property. For instance, data type property `isCompleted` or `isActive` can take only Boolean values, i.e., True or False. Some data type properties such as `difficulty (Problem, Value)` can take only three values: easy, medium and difficult. “Domain” and “range” for object properties and “domain” for data type properties was also defined while defining the slots. Once the core ontology was defined, then it was populated manually by the instances. For example:

`Shape --> ThreeDShape --> RightCircularCylinder:CylindricalLampShade`

Here Shape is a class with ThreeDShape as a sub-class. RightCircularCylinder is a sub-class of ThreeDShape and CylindricalLampShade is the instance of that sub-class. Please see Figure 2 for a small extract of the domain ontology. The above example demonstrates part of the knowledge structure of the domain knowledge, and provides some more understanding of it. Easy machine readability and the general nature of ontology allow it to be reused for various purposes. Crowley and Medvedeva (2003) mention as one of the advantages of using ontology that knowledge in the form of ontology can be easily modified without making any changes in the back end code.
MONTO: Ontology for teaching word problems in mathematics

The proposed MONTO ontology is divided into three main parts: system/pedagogy ontology, task ontology, and domain ontology. The following sub-sections describe these in detail and justify the process of creation of MONTO.

System/ Pedagogy ontology

Information processing and learning to think are the essence of problem solving (Joyce & Well, 2007). The Analysis-Synthesis method of teaching is used for teaching problem solving, geometry constructions, proofs and riders. In Analysis-Synthesis, by asking the questions during the analysis, the teacher tries to devise a plan to approach that problem. For synthesis, a part of the execution of the plan is taken from students, and based on this the teacher asks questions about the content. Greeno (1980) mentioned that for teaching problem solving, the first need is to teach how to analyze a problem and then to teach how to carry out the plan that is derived from the analysis of the problem, which is supportive to the Analysis-Synthesis method of teaching.

The System/Pedagogy ontology is an ontology which can be used for tutoring problem solving based on the principles described above. The general form of the system ontology (See Figure 3) includes Problem class (this can further be divided into sub-classes as per the type of problems), ProblemSchema class, Question class, Answer class, Bug Class, Feedback Class and User class. To facilitate analysis of a given problem, analytical questions can be asked to a student, and those questions are included as individuals of the Question class. The student’s response to each question may be one of a number of possible answers. Those answers are stored under the class Answer. The answers can be more than one, and depending on the answer the next question can be asked. In this way the student can be tutored from her own knowledge level. In order to give feedback based on the student response, each answer includes a feedback, which is stored as a part of Feedback class.

All these classes are connected by object properties such as hasQuestion(Problem, Question), has Option(Answer, Question), hasFeedback.(Answer, Feedback). Bug class has two sub-classes named MisConceptions and MissingConcepts. Missing concepts are concepts which the expert has but the student does not have, and misconceptions belong to the student and not the expert. For each wrong answer there are expected bugs (either a MisConception or a MissingConcept) that are related to the Answer class by a relation expectedBug(Answer, Bug). Each bug has remedial feedback. Each problem type has a schematic structure which is captured in the
ProblemSchema class, which also gives semantic knowledge of the problem. Thus we can show a student a Schematic, Semantic and Linguistic Knowledge representation of each individual problem, which can lead to a firmer understanding of the word problem. All parts are explained with an example. Each of these classes has a data type property named value, with “String” as a data type.

Figure 3. System/Pedagogy ontology

Task/Strategy ontology

Figure 4. Task/Strategy ontology
The term “strategy” means a way of solving a problem. Every problem can be solved by multiple strategies. For every strategy, students need to do multiple tasks. Task ontology (See Figure 4) contains the classes \textit{Strategy} and \textit{Task}.

The \textit{Strategy} class is connected with the \textit{Problem} class and the \textit{Task} class by the relations \textit{hasStrategy(Problem, Strategy)} and \textit{hasTask(Strategy, Task)} respectively. There can be sub-strategies and sub-tasks as well. Each problem has given items and to-find items that are collected under the \textit{GivenExpression} class and the \textit{ToFindExpression} class, which are connected separately to the objects in the domain ontology. Many times, Given and ToFind items are not provided explicitly in a problem, and by identifying them missing concepts or misconceptions of students can be captured. Also, at each task level one can identify students’ missing concepts or misconceptions. Hence all these classes are also connected to the Bug class.

\textbf{User model ontology/ student ontology}

All the actions of students are connected to the \textit{Bug} class, and each bug has a remedial feedback. So at any stage the student’s missing concepts and misconceptions are captured. Also, the information about problems the student has completed, strategies the student has applied, tasks the student has undertaken, questions the system has asked the student during analysis of the problem, and the answers the student has selected are captured in the student ontology (See Figure 5).

The \textit{User} class in the system ontology is connected with the \textit{Problem} class and \textit{Feedback} class. Through the feedback to student and the information stored in the \textit{User} class (i.e., the student’s learning profile), one can build a student model that can be used for further intelligent tutoring. The student learning profile comprises the strengths and weaknesses of a student in that particular domain. The \textit{Problem} class is the major class, which connects system ontology, task ontology, and domain ontology, and all these ontologies together can be used for building the student model for an adaptive intelligent tutoring of word problem solving in mathematics.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{User model ontology/ student ontology}
\end{figure}
MONTO for mensuration domain explained with an example

An analysis of a total of 151 problems of Mensuration in Class IX and X mathematics textbook of CBSE (Central Board of Secondary Education) published by NCERT India shows that these problems can be divided into several types. A total of six shapes and 16 types of combinations of shapes are covered. For each shape, the problems can be divided into two major types: surface area problems (total 68) and volume problems (total 83).

Mary wants to decorate her Christmas tree. She wants to place the tree on a wooden box covered with colored paper with picture of Santa Claus on it. She must know the exact quantity of paper to buy for this purpose. If the box has length, breadth and height as 80cm, 40cm, and 20 cm respectively, how many square sheets of paper of side 40 cm would she require?

Figure 6. Example of problem statement for covering problem

Surface area problems can be divided into further subtypes, such as problems with lateral surface area, total surface area and surface area with fewer faces. The real world scenarios for each of these problem types consist of problems such as: (a) covering problems, where, for example, the surface area of the shape is covered or painted or whitewashed with some material or some object; (b) building problems, where, for example, a cuboidal-shaped structure is built with cuboidal-shaped bricks; (c) painting problems; and so on. In such problems, we either need to find the surface area covered, or the number of objects required covering the surface, or the cost of covering/painting/whitewashing/building etc. Thus, we have 151 problems that can be divided into 31 types of problems. For the purpose of illustration, a covering problem involving a cuboidal-shape is discussed here to demonstrate the use of ontology in teaching how to solve it. The problem statement is given in Figure 6.

Q1) What do we need to find in the given problem?

A11: Quantity of paint required to paint the wooden box A12: Number of square sheets required to cover the wooden box A13: Quantity of white colour required to white wash the wooden box A14: Number of square tiles required to cover the wooden box

Q2) How do you find number of square sheets required to cover the wooden box?

A21: Area of the square sheet A22: Surface area of the wooden box A23: Volume of the wooden box A24: (Surface area of wooden box)/(Area of square sheet)

Q3) How do you find surface area of the wooden box?

A31: 2 x (Breadth x Height + Height x Length) A32: (Length x Breadth x Height) A33: 2 x (Length x Breadth + Breadth x Height + Height x Length) A34: (Length x Breadth) + 2 x (Breadth x Height + Height x Length)

Q4) How do you find area of square sheet?

A41: (Radius x Radius) A42: (Breadth x Height) A43: (Lateral Height x Height) A44: (Side x Side)

Q5) Which quantities are given in the problem to find surface area of wooden box?

A51: Lateral height of wooden box, Radius of wooden box, Height of the wooden box A52: Length, breadth and height of wooden box A53: Side of wooden box A54: Side and radius of wooden box

Q6) Which quantities are given in the problem to find area of square sheet?

A61: Length and Breadth of the square sheet A62: Radius of the square sheet A63: Side of the square sheet A64: Base and height of the square sheet

Figure 7. Example of set of questions to be asked to analyze the given problem
Usually when a problem is given, students are expected to analyze it first. In order to guide the student in the analysis, the system ontology will have the following set of questions and the choices that can be offered to students to answer those questions. The choices for the answer will have some distracters and the correct answer. There can also be feedback for each choice. The distracters may detect some misconception or mistake of the student.

In Figure 7 is shown a set of questions that can be asked to the student to teach how to solve the given problem. The distracters are close to the information in the problem, with either insufficient information to find the required quantity or with the wrong concept. The correct answers are A12, A24, A33, A41, A54, A63. Now, the advantage of this approach is that the student’s knowledge or learning profile (which includes the student’s knowledge of concepts or misconceptions or wrong answers) can be found and appropriate feedback can be provided to the student. Feedback to the system about the student’s knowledge profile can add information to the student model. From the domain perspective, we have ontological knowledge that WoodenBox is an instanceOf Cuboid class, which has Length, Breadth and Height, and SquareSheet is an instance of Square class that has side property. Every Cuboid has SA and every Square has an Area, which has a certain formula. All this domain knowledge is there in the domain ontology to manipulate the knowledge model. The ontology includes two strategies for solving the problem.

The first strategy is the standard way of solving this type of problem, which comprises tasks such as: (a) Finding area of wooden box by using formula; (b) Finding area of one square sheet; (c) Finding number of square sheets required to cover the wooden box. The second strategy involves using the following tasks: (a) Finding area of each face of the wooden box; (b) Finding area of one square sheet; (c) Finding number of square sheets required to cover each face of the wooden box; and (d) Finding total number of square sheets required to cover the wooden box. The task ontology captures these strategies. The student’s problem solving skill/strategy can be monitored and tracked by using this task ontology. The information gathered by tracking can be stored in the student model ontology, which can be used for further intelligent feedback and tutoring to the same student. Also, the CoveringProblem Schema is shown in Figure 8, where it is explained how schematic, semantic and linguistic knowledge gets captured (See Figure 9). For the above example of a covering problem, a part of the problem schema is used.

The instance of sub-class of the parts of the problem schema provides semantic knowledge and linguistic knowledge of the problem. For instance, there is a lot of extra information in the above problem which an average student may get distracted by while solving the problem. However, the linguistic, semantic and schematic knowledge representation provides help for extracting useful information from the real world scenario according to what is required for solving a particular type of problem. The major advantage of such a schema is that it helps students to understand semantic relations in the domain knowledge and converts linguistic knowledge into the mathematical knowledge required to solve a problem. In the above example, only one way of solving is shown, but in the ontology one can include a great number of ways (strategies) to solve a problem.
Comparison of MONTO with related work

Usually ontology is used for information processing and retrieval in AI, ML and Web applications. However, earlier work on the use of ontology for teaching in general is limited. Ontology is used for teaching in some systems, such as ActiveMath (Melis et al., 2001), a web-based tutoring system for high school and college level mathematics, and SlideTutor (Crowley & Medvedeva, 2003), an intelligent tutoring system for microscopic diagnosis of inflammatory diseases of the skin. ActiveMath uses ontology for knowledge representation and annotation of learning objects (LOs), which contributes to the re-usability and interoperability of the content. This domain ontology can provide knowledge representation in the proposed ontology for teaching problem solving. It does not include the use of LOs, but this can be incorporated according to implementation needs.

The major difference between MONTO and the ActiveMath ontology is that MONTO diagnoses the mistakes with a greater degree of granularity. SlideTutor uses various types of ontologies, such as domain ontology, domain task ontology, ontology of abstract problem solving methods, and pedagogical ontology. Errors and hints are modeled on the instructional module (i.e., on pedagogical ontology) to provide personalized training. Use of these ontologies offers flexibility in the instructional layer, scalability, reusability, and easy maintenance. Application of these ontologies is very similar to the proposed application of MONTO, the problem solving tutoring ontology.

The system ontology in MONTO is called pedagogic ontology. The difference is in the nature of the domain knowledge. The medical domain is qualitative in nature, whereas the mathematical domain is quantitative. So the granularity in which mistakes are diagnosed by MONTO is much higher as compared to qualitative medical domain used in SlideTutor. There also exists an ontology named OMNIBUS that has been used in SMARTIES, an innovative intelligent tutor authoring system discussed in (Mizoguchi et al., 2010). The OMNIBUS ontology is about target-world modeling, which contains an “is-a” hierarchy of not only the target concepts but also the states, actions, processes, events, and learning and instructional theories. The objective of the creation of the OMNIBUS ontology is much wider as compared to the problem solving tutoring ontology proposed in this paper. It is being used mainly for authoring of intelligent tutoring systems that need to integrate Learning theories and Instructional theories and not for any particular domain knowledge. However, MONTO can be used as a back-end for creating ITS for teaching word problems in mathematics. Table 6 shows the gaps filled by MONTO.
Table 6. MONTO filling the gaps in the evaluation rubric

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Linguistic KR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Semantic KR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Syntactic KR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Multiple KR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modeling of problem strategies</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questioning</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modeling of problems</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Modeling of students</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X (Missing and Misconceptions)</td>
</tr>
</tbody>
</table>

Evaluation of MONTO

For evaluation of any knowledge sharing technology or ontology, (Gomez-Perez, 2004) suggested three parts: (a) Verification of the ontology; (b) Validation of the ontology; and (c) Assessment of the completeness of the ontology.

Verification

Verification of ontology refers to technical activity that guarantees correctness of the created ontology. While creating the ontology, the consistency check has been done for the classes, sub-classes, and properties through which the connection took place. No inconsistency was found in the ontology. The ontology is created manually and the instances are populated manually. Only one particular domain was considered for ontology creation, so the scope was kept limited to that domain. In total there are 933 classes in the ontology, of which 890 classes are specific for the selected domain. In total there are 64 properties defined, out of which 46 are object properties and 18 are data type properties. Out of 46 object properties, 20 are inverse functional. A greater number of inverse functional properties would have improved the efficiency of the ontology.

Validation

Validation in software refers to the task of checking a conceptual solution by actually implementing and realizing the conceptual solution. The concept of the proposed MONTO ontology was implemented by using an editor named Protégé, developed by Stanford University group. Of all the available ontology editors, Protégé was found to be the most efficient and useful editor. Protégé is open source and offers extensibility with various plug-ins. It imports ontologies in forms such as: OWL, RDF, XML, (RDF, UML, XML) backend; and exports ontologies in forms such as: XML, RDF, OWL, SWRL-IQ (Alatrish, 2012). For knowledge representation it supports frames, first order logic, SWRL and Meta classes. It offers in-built inference engines named PAL and DIG, but can also use other external inference engines such as RACER, FACT, FACT++, F-Logic and Pellet. Protégé has all the usability tools, such as graphical taxonomy, graphical views, zooms, collaborative working, and ontology libraries. The implementation includes SWRL rules such as:
The above SWRL rule finds which problems a user has completed for the learning mode. At each stage of the ontology development, feedback regarding correctness of the terminology used is validated. The terminology used for creating ontology is as per the W3C standards (defined by the World Wide Web Consortium).

**Completeness**

Completeness of the ontology refers to how complete the ontology is with respect to the body of knowledge from the same domain with respect to some unseen use-cases. In this case we consider it as an evaluation of the robustness of the ontology. Robustness subsumes completeness. In addition to capturing domain coverage, robustness also reflects the ability to withstand minor changes (perturbations) in the form of changes. To evaluate robustness of MONTO developed for the domain surface area and volume of 3D shapes of the class IX and X syllabus of CBSE curriculum, all problems of the same domain from the ICSE curriculum (total 121) and Maharashtra State Board Curriculum (total 41) were analyzed. A cross-domain evaluation of ontology has been done by fitting these problems into the ontology developed for CBSE curriculum problems. In this case, the ontology developed by using CBSE problems is considered as a baseline. Analysis and fitting of problems from both the curricula have been done by two different experts in the field. The results are given in Table 7 and Table 8. The other classes have been created by the ontology editor, so they are not considered in the evaluation.

![Figure 10. Garden pot with PentagonBasedPrism](image)

We may observe that only one type of problem and five sub-classes for the Shape class have been added in the analysis of ICSE board problems. The addition of five shapes was due to inclusion of three 2D shapes named Semicircle, Sector and Trapezium and a 3D shape named TrapeziumBasedPrizm. The respective additions in the Measurement class are due to these shapes. Similarly, only one 2D shape, named sector, was added in the ontology for Maharashtra State Board problems. The advantage of such additions is that it adds the quality of usability of the ontology with relatively small efforts. For instance, we added only one type of problem and four shapes together with respective measurements in ontology, but we can create 31 types of problems with those four shapes. Also, we can create different types of problems which are not covered in the syllabus but can be fitted into the ontology. For example, we can create a problem that involves finding the cost of painting a garden pot which has a PentagonBasedPrism shape with Trapezium-shaped lateral faces (see Figure 10), where the dimensions of the Trapezium and the cost of painting would be given.

<table>
<thead>
<tr>
<th>Names of classes in which edits has taken place</th>
<th>CBSE</th>
<th>ICSE</th>
<th>Difference of Classes in CBSE and ICSE</th>
<th>SSC (Maharashtra)</th>
<th>Difference of Classes in CBSE and SSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other Classes</td>
<td>43</td>
<td>48</td>
<td>5</td>
<td>44</td>
<td>1</td>
</tr>
<tr>
<td>Measurement</td>
<td>124</td>
<td>135</td>
<td>11</td>
<td>128</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 7. Evaluation of ontology across the boards for specified domain
<table>
<thead>
<tr>
<th></th>
<th>Value Unit</th>
<th>Shape</th>
<th>Unit</th>
<th>Given-Expression</th>
<th>Problem</th>
<th>Problem Schema</th>
<th>Question</th>
<th>Answer</th>
<th>Strategy</th>
<th>Task</th>
<th>To Find Expression</th>
<th>Total without other classes</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>117</td>
<td>128</td>
<td>11</td>
<td>121</td>
<td>4</td>
<td>34</td>
<td>34</td>
<td>32</td>
<td>31</td>
<td>139</td>
<td>34</td>
<td>890</td>
<td>933</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34</td>
<td>35</td>
<td>4</td>
<td>35</td>
<td>151</td>
<td>35</td>
<td>944</td>
<td>992</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>35</td>
<td>9</td>
<td>32</td>
<td>151</td>
<td>1</td>
<td>54</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>15</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 8. Percentage of edits and robustness of baseline ontology

<table>
<thead>
<tr>
<th>Board</th>
<th>Total number of problems analyzed</th>
<th>% of robustness of baseline ontology</th>
<th>% of edits of baseline ontology</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICSE</td>
<td>121</td>
<td>94%</td>
<td>6%</td>
</tr>
<tr>
<td>SSC (Maharashtra State Board)</td>
<td>41</td>
<td>98%</td>
<td>2%</td>
</tr>
</tbody>
</table>

If an above-average student has learned all the types and shapes, then that student can apply that knowledge to solve this problem. This illustrates the use of ontology for the generation of new problems which are not given in the syllabus and would require more application level skill. As compared to 121 problems from the ICSE board, the baseline ontology developed by analyzing 151 CBSE problems was found to be 94% robust, with 6% edits. Similarly, as compared to 41 problems from Maharashtra State Board, the baseline ontology was found to be 98% robust with 2% edits. The results were quite satisfactory as the edits found were only in the domain ontology and no edits were found in the pedagogy ontology, task ontology and student model ontology.

**Research contributions**

**Implications to theory**

(i) The MONTO ontology fits very well into the framework given by Schoenfeld for development of mathematical thinking, i.e., resources, heuristics and control. (ii) The MONTO ontology can capture Semantic Knowledge and Schematic Knowledge of a word problem and show it to students for reinforcement. These knowledge extracts can be used for showing multiple knowledge representation to students, which may help them to reduce cognitive load. (iii) The MONTO ontology uses modeling of problems, modeling of problem strategies, asking analytical questions, capturing missing concepts and misconceptions and modeling of students, which is highlighted in the state of the art and has been used for getting student model information. (iv) The MONTO provide diagnosis of students’ mistakes and provide remedial feedback. (v) The MONTO is general in nature and any other mathematical domain can be fitted into it.

**Implications for teachers**

A smart learning environment (Lalingkar et al., 2014) is developed by using MONTO at the back-end. By using the information captured in the student model of MONTO, teachers can get learning analytics of students such as: the concepts learned by students, missing concepts and misconceptions of students, the problem types learned by students. Teachers can use these learning analytics to provide focused guidance to their students.
Limitations and future work

Manual creation of ontology is time-consuming and costly. However, semi-automatic creation of ontology by using natural language processing (NLP) techniques is possible. If domain experts are needed for creation of ontology, then the complex nature of ontology would pose a learning curve challenge for them. A possible future application of the MONTO ontology for teaching problem solving in mathematics is the following: Using NLP and machine learning (ML) techniques, similar problems can be generated; the system could be trained to capture the ways good students solve problems and the mistakes weak students make; the system could then be used for teaching problem solving to weak students. This may help to improve students’ beliefs about mathematics and problem solving in mathematics.

References


Constructing Proxy Variables to Measure Adult Learners’ Time Management Strategies in LMS

Il-Hyun Jo¹*, Dongho Kim² and Meehyun Yoon³

¹Department of Educational Technology, Ehwa Womans University, Seoul, Korea // ²Department of Career and Information Studies, University of Georgia, GA, USA // ³Department of Career and Information Studies, University of Georgia, GA, USA // ijo@ewha.ac.kr // dongho@uga.edu // meehyun@uga.edu

*Corresponding author

(Submitted May 5, 2014; Revised October 31, 2014; Accepted December 31, 2014)

ABSTRACT

This study describes the process of constructing proxy variables from recorded log data within a Learning Management System (LMS), which represents adult learners’ time management strategies in an online course. Based on previous research, three variables of total login time, login frequency, and regularity of login interval were selected as candidates from the data set, along with a guideline for manipulating the log data. According to the results of multiple regression analysis, which was conducted to determine whether the suggested variables actually predict learning performance, (ir)regularity of the login interval was correlative with and predictive of learning performance. As indicated in the previous research, the regularity of learning is a strong indicator for explaining learners’ consistent endeavors and awareness of learning. This study, which was primarily based on theoretical evidence, demonstrated the possibility of using learning analytics to address a learner’s specific competence in an online learning environment. Implications for the learning analytics field seeking a pedagogical theory-driven approach are discussed.

Keywords

Learning analytics, Educational data mining, Time management strategy, Adult education

Introduction

There are high demands within e-learning for adult learners. Over the past years, there has been an increase in online course enrollment among adult learners in order to obtain knowledge or develop professional skills (Cohen & Nycz, 2006; Hrastinski, 2008; Wang, Vogel, & Ran, 2011). In spite of a great deal of attention that has been paid to such online courses, learners have faced challenges when taking such online courses due to a lack of time management skills (Kwon, 2009). Time management strategies are increasingly required in the context of adult learning, because such learners are in the usual situation of being involved both in their study and their job at the same time; therefore, the successful completion of an online course depends on the efficient use of a given amount of time. Failure in online courses for adult learners has been reported to result from poor time management (Eastmond, 1998; Joo, Jang, & Lee, 2007).

To address these issues, utilizing log data of the learners can be considered. Log data reflect learners’ online behaviors, which are closely related to their learning style (Hwang & Wang, 2004; Sharma & Gupta, 2012). Instructors can detect the status of their learning processes at earlier stages (Hung & Zhang, 2008; Nanjiani, Kelly, & Kelly, 2004). If they can distinguish learning patterns in the early stage of an online course, it will be conducive to encouraging or guiding learners by providing them with appropriate instructional intervention (Brown, 2011).

Log data, which is saved as an unstructured data set, contains users’ log information within online systems. Users begin their studies when they log into a webpage, where they stay until they complete their learning. Because the log data detects every activity and click made by the users who are taking the courses, it can be used to represent how the learning processes occur on the web throughout the login duration. Furthermore, this information might be more genuine when compared to the data obtained from surveys, which rely highly on learners’ recall and subjective interpretations; thus, the possibility of distortion or low reliability need not be considered (Baker, 2010; Elias, 2011). However, log data alone cannot be transferred to meaningful learning processes without sophisticated interpretation of the theoretical aspects. The contribution of this study is to suggest an effective way to convert users’ log data into predictive indicators of learning performance based on a theoretical background.
The focuses of this study were two fold. First, “candidates for proxy variables” from the log data set that represent learners’ time management strategy were elicited, conceptual constructs that have long been considered to be a vital factor to high performance (Barling, Kelloway, & Cheung, 1996; Macan, Shahani, Dipboye, & Philips, 1990; Printrich, & DeGroot, 1990; Song, Singleton, Hill, & Koh, 2004; Zimmerman, 2002). Second, it was determined whether the elicited proxy variables predict learner performance in terms of verifying the empirical validity. The proxy variables identified can be used to detect the status of learners’ time management and to predict performance in other data sets from similar contexts.

Proxy variables to represent time management strategy

A proxy variable refers to an alternative that can be used when the actual variable is not measurable or not reliable (Jo & Kim, 2013; Wickens, 1972). Proxy variables have been widely used in the field of social science for the generation if prediction models (Durden & Ellis, 2003). Time management strategy was decided to be a constant habit of learners that is clearly manifested in their behaviors and accurately measured through analysis of the behavioral data, rather than use of self-reporting questionnaires that can lead to significant bias.

In order to establish reliable proxy variables, we attempted to build a solid theoretical foundation, which provides the rationale for manipulating the data set.

Since previous studies took different approaches to define time management strategy, candidates for proxy variables assumed to represent time management strategy were constructed. The following steps were taken to convert the unstructured data set into proxy variables (see Figure 1).

First, as a targeted conceptual construct, the concept of time management strategy was examined based on findings from previous studies. A body of literature on time management strategy of adult learners was examined, focusing on its definition, sub-factors, and impact.

The findings provided in previous research dictated the manipulation of data; further, necessary values, such as time spent on each activity and login points obtained from the data set, were extracted for further calculation. In essence, by considering the theoretical aspect, it was determined which things should be included from among the values in the data set.
A proxy variable, not identical to the targeted conceptual construct but optimally processed, can be applied to other data sets in order to determine whether learners’ behaviors within the systems represents the targeted conceptual construct. In this study, three variables were chosen on the basis of previous studies.

**Time management strategy**

A large body of research has argued that time management strategies are deeply associated with the ability to prioritize tasks (Jex & Elacqua, 1999; Kaufman-Scarborough & Lindquist, 1999; Blaxter & Tight, 1994). With higher levels of time management ability, learners are able to better organize the given tasks in the order of importance. Because all the participants had to receive a certain required score on the final exam to complete the course, it was determined that those with a higher level of time management strategy would put more effort into the course, which would result in more commitment to learning. This assumption is in accordance with previous studies, which have regarded time management as a technique for having sufficient time to accomplish the required tasks (Slaven & Totterdell, 1993; Woolfolk & Woolfolk, 1986). Many adult learners involved in both work and study succeed to achieve their academic goals when they spend extra time only for learning outside of work (Kember, 1995; Blaxter & Tight, 1994). From the perspective that regards it as a sub-element of self-regulated learning, many researchers have also emphasized the efficient use of time (Hofer, Yu, & Pintrich, 1998; Pintrich & DeGroot, 1990; Pintrich, Smith, Garcia, & McKeachie, 1993; Zimmerman, 2002). The concept incorporates a series of sub-factors, such as sufficient amount of time investment in tasks and active participation (Blaxter & Tight, 1994; Orpen, 1994; Woolfolk, & Woolfolk, 1986).

On the other hand, long-range planning has been considered to be a crucial factor for explaining time management strategy (Britton & Tesser, 1991; Eastmond, 1998). By sustaining consistent efforts based on a well-planned schedule, learners are able to maintain the expected conditions over the long-term, as well as achieve their expected learning outcome more effectively. It was assumed herein that learners who sustain consistent endeavors are more likely to study on a regular basis. That is, having constant awareness of their learning results in regular intervals among visits, because they want to keep themselves updated on new information. Regular study was considered to be the result of prioritization strategy, as it implies that the learners tried to be more engaged in their learning with consistent awareness.

![Figure 2. Selection of three variables](image-url)

**Proxy Variables (Data Process)**

- Total login time
  - Time spent on activities

**Login Frequency**

- The number of login points
  - Intervals between login points

**Time Management Strategy (Conceptual construct)**

- Prioritizing tasks
  - Sufficient amount of time investment

- Active participation
  - Persistency at tasks

- Well-planned time use

*Figure 2. Selection of three variables*
Based upon the theoretical evidence described above, potential proxy variables which represent time management strategy were chosen. They encompassed learners’ quantified endeavors, participation and consistency throughout the time of the study. The relationship between each potential proxy variable and sub-element of the time management strategy is presented in Figure 2, as a targeted conceptual construct.

Total login time

The degree to which learners invest their time has been recognized as a factor correlative with learning performance; moreover, a great deal of research has reported a strong relationship between total studying time and performance (Beaudoin, 2002; Carroll, 1963; Hwang & Wang, 2004; Thurmond, Wambach, & Connors, 2002).

Studying time for the online course was calculated based on “login time,” since the actual learning in an online course mostly occurs while being logged in. To support the significance of the variable, the conception of learning time as suggested by Cotton and Savard (1981) was adopted (see Figure 3). In the study, learning time was comprised of three parts: Allocated Time (AT), Time-on-Task (TOT) and Academic learning time (ALT). This conception is illustrated as follows.

![Figure 3. Three types of learning time](image)

In the original study, the authors argued that all three types of time are correlated with learning performance. This is in accordance with subsequent studies relating time spent on learning to learning achievement (Hwang & Wang, 2004). It can be assumed that the recorded login time can be considered as learning time, as it belongs to either the allocated time or time-on-task categories. In fact, the majority of previous studies used login duration as the measurement of learning time (e.g., Hung & Zhang, 2008; Hwang & Wang, 2004). However, determination of a more reliable learning time was attempted herein by combining time spent on each activity, such as watching lectures, reading materials, or posting questions. This is a more effective way than using total login duration, since the whole duration of the login might not be fully devoted to learning. For example, login time inevitably includes simple accesses which involve no learning activity. If each activity is combined, however, time spent on actual learning activities can be obtained exclusively.

![Figure 4. An example of calculating total login time](image)
Another data processing step was taken to deal with the fact that many learners logout simply by closing their browser without actually clicking a “sign out” button, which results the logout time not being recorded in the system. The time spent on the last activity before the next login was replaced with the mean of time spent on other activities within the same visit (see Figure 4). Although the exact amount of time spent on the last activity could not be measured, this provides a way of avoiding the inclusion of time passed without participation in learning activities.

Login frequency

How frequently learners participate in an online course has been considered to be an important factor predicting higher levels of performance. Piccoli, Ahmad, and Ives (2001) reported that learners’ login frequency into the LMS was highly correlated with course satisfaction in online learning. In a study measuring the correlation between frequency of participation and the grades achieved by college students, Davies and Graff (2005) argued that a significant association between frequency of participation in online activities and grades. Kang, Kim, and Park (2009) demonstrated that total login frequency in the LMS is directly connected with not only learning performance, but also attendance rate. Similarly, Hung and Zhang (2008) emphasized the significance of login frequency as a predictor of learner’s higher performance.

The instructor of the course analyzed in this study frequently uploaded learning material for learners to use in the class. It was therefore assumed that the more frequently learners logged into the LMS, the more newly updated and shared information they could obtain, which would lead to gaining a better understanding of the learning content as well as what needs to be prepared for the classes. The login frequency was calculated by adding up the number of individual student’s login time in the LMS.

The degree to which learners invest their time has been recognized as a factor correlative with learning performance; moreover, a great deal of research has reported a strong relationship between total studying time and performance (Beaudoin, 2002; Carroll, 1963; Hwang & Wang, 2004; Thurmond, Wambach, & Connors, 2002).

Studying time for the online course was calculated based on “login time,” since the actual learning in an online course mostly occurs while being logged in. To support the significance of the variable, the conception of learning time as suggested by Cotton and Savard (1981) was adopted (see Figure 3). In the study, learning time was comprised of three parts: Allocated Time (AT), Time-on-Task (TOT) and Academic learning time (ALT). This conception is illustrated as follows.

Regularity of login interval

Regularity of learning was presumed to represent learners’ consistent endeavors and awareness of their learning status throughout the course. Learners who want to achieve their learning goals while also engaging in outside work should exhibit perseverance with learning. It was assumed that learners with consistency in and awareness of their learning status study regularly. In fact, adult learners with multiple commitments to work, family, and study show the adoption of a regular routine to achieve their academic goals (Blaxter & Tight, 1994).

Regularity of login interval was calculated using standard deviation of the login interval. This basically indicates the “irregularity of the learning interval.” In the following figure (see Figure 5), the gap between A and B indicates the total course period, while \( \Delta t_{1,2} \) is the interval between the first and second login time, calculated by subtracting \( t_1 \) from \( t_2 \). In the same way, the nth interval can be obtained, presented as \( \Delta t_{n-1,n} \). Consequently, the mean of a learner’s login interval can be calculated, from which the variance is subsequently elicited, as indicated in Figure 6.

Figure 5. Concept of login interval
Mean of learning interval: 
\[ \overline{\Delta t_i} = \frac{\sum_{i=1}^{n-1} \Delta t_i}{n} \]

Standard deviation of learning interval: 
\[ S_i = \sqrt{\frac{\sum_{i=1}^{n-1} (t_i - \overline{\Delta t_i})^2}{n-1}} \]

*Figure 6. Calculation of mean and standard deviation of login interval*

Even when variation occurs in each learner’s login duration and login frequency, the regularity of the login interval can be the same (see Figure 7). However, considering that total login time and login frequency were included as independent variables in the predictive model, learning regularity can indicate learners’ efforts to keep themselves on right track throughout the course.

*Figure 7. Example of calculating login regularity*

**Research questions**

Whether or not e-learning is successful is primarily dependent upon how a learner who has autonomy in his/her own learning sustains consistent investment of time and effort. In particular, those adult learners who cannot fully engage only in their studies may become easily frustrated, as they suffer from perceived difficulties in managing time (Blaxter & Tight, 1994).

The purpose of this study was to propose proxy variables with regard to time management strategy by manipulating log data, as well as to examine whether the factors actually predict learning performance, as indicators of time management strategy. It is expected that such indicators will be conducive to providing precautionary feedbacks or
interventions within an early stage based on the learner’s status. In this study, the following research questions were addressed.

R1: How can candidates for proxy variables regarding adult learners’ time management strategy be elicited from a log data set?

R2: Do the suggested variables (total login time, login frequency and regularity of login interval) predict adult learners’ performance?

Methods

Participants and research context

The participants in this study were recruited from a commercial e-learning course entitled “Credit Derivative,” administered by a Korean e-learning company. All of the 200 participants (53 male, 147 female) work full-time in the financial business field. The course aims to provide learners with knowledge about derivative products, such as its definition, relevant laws, strategy for safe investment, and interpretation of index numbers. Learners who complete the course are expected to be prepared for client consultation. All the participants voluntarily participated in the course, even though they were financially supported by their employers to do so. This course was operated 100% online for one month, and was comprised of 12 modules that provide one hour video lectures covering part of the textbook.

Along with the asynchronous video lectures, bulletin boards and relevant resources were provided in the LMS for learners to download learning materials and post questions. Learners were supposed to study the content in advance of each lecture and take a test at the end of the course. They were allowed to replay all videos for review.

Measures and variables

Suggested independent variables

Log data were collected from the LMS by an automatic collection module embedded within the system. The Total login time, Login frequency and (ir)regularity of login interval were extracted as independent variables.

Dependent variable

Learning performance, the dependent variable in this study, was defined as the score obtained on the final test, which consisted of 20 multiple choice items, each equally worth 5 points (See Table 1). The scores from each question were collected and added together in order to obtain the total score. The total score was graded on a scale of one hundred points, and ranged from 40 to 100. A unit of measurement of time was the hour, and Total login time and (ir)regularity of learning interval was calculated accordingly.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total login time (hour)</td>
<td>38.18</td>
<td>14.46</td>
</tr>
<tr>
<td>Login frequency</td>
<td>46.31</td>
<td>12.56</td>
</tr>
<tr>
<td>(Ir)regularity of learning interval</td>
<td>2.92</td>
<td>2.05</td>
</tr>
<tr>
<td>Learning performance</td>
<td>77.92</td>
<td>15.09</td>
</tr>
</tbody>
</table>

Note. N = 200.
Results

Correlation analysis

Correlation analysis was conducted prior to multiple regression analysis to identify correlations among the independent and dependent variables (See Table 2). The results showed that (ir)regularity of the learning interval had a significant negative correlation with the final test score, defined as learning performance, and was also negatively correlated with Login frequency.

Table 2. Pearson’s correlation of suggested variables and learning performance

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total login time</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Login frequency</td>
<td>-.016</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Ir)regularity of login</td>
<td>-.064</td>
<td>-.181*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Learning performance</td>
<td>.035</td>
<td>.141*</td>
<td>-.597**</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. *p < .05. **p < .01, (2-tailed).

Multiple linear regression analysis

Multiple regression analysis can be used to determine the significance of independent variables (Pedhazur, 1997). Since the primary focus of this study was to find a proxy variable which represents learners’ time management strategy, multiple regression analysis was performed to determine which of the three candidates for proxy variables predict learning performance. The results are presented in Table 3.

The three suggested variables were found to account for 34.7% of the variance in learning performance ($F = 36.267, p < .01$). Of the three proxy variables, only (ir)regularity of learning interval was found to be a significant predictor of learning performance ($\beta = -.590, t = -10.115, p < .01$).

Table 3. Results of multiple linear regression analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized coefficients</th>
<th>Standardized coefficients</th>
<th>Collinearity statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
</tr>
<tr>
<td>(constant)</td>
<td>76.672</td>
<td>4.511</td>
<td></td>
</tr>
<tr>
<td>Total login time</td>
<td>-.002</td>
<td>.060</td>
<td>-.002</td>
</tr>
<tr>
<td>Total login frequency</td>
<td>.041</td>
<td>.070</td>
<td>.034</td>
</tr>
<tr>
<td>(Ir)regularity of login interval</td>
<td>-4.343</td>
<td>.429</td>
<td>-.590</td>
</tr>
</tbody>
</table>

$R^2$ (adj. $R^2$) = .357 (.347), $F = 36.267, p = .000$, Durbin-Watson = 1.692

Note. Dependent variable: Learning performance.

Discussion

This study primarily focused on providing a unique process to construct proxy variables based on previous studies and determining which candidate best predicts adult learners’ performance. The results of the correlation analysis revealed that the (ir)regularity of login interval factor had a significant correlation with and was a predictor of learning performance. This was also supported by the results of multiple regression analysis, which also indicated (ir)regularity of login interval as a predictor of learning performance.

Although learning time and learning frequency have both been emphasized in previous studies as predictive factors positively affecting learner performance, it is hard to assume that either login time or login frequency of adult learners are directly translated into their ability to sustain effort and time investment throughout a course. Some research reported a limited relation between studying time and learning performance. For example, Ha and colleagues (2006) insisted that total studying time is not related to learning performance, while Lee (2007) highlighted the fact that learning performance can be increased only when the learners fully concentrate on what they
do, regardless of the total amount of available time. Similarly, login frequency cannot fully explain whether the learners partook in meaningful learning. Although learners may access the LMS frequently, their actual studying time might be relatively short. It is also likely that they log into the LMS more intensively at a certain point rather than constantly participating in academic activities. Such tendencies have been reported to hinder well-planned learning, resulting from either procrastination or a lack of effective time use (Howell & Watson, 2007).

The regularity of the learning interval, unlike frequency, can provide critical evidence as to the fact that learners who log into the LMS more steadily from the beginning of a class to the end show better performance. This involves neither temporal access at a certain point nor merely one long time visit, but rather conscious learning with awareness over a relatively long term. Some studies have recognized regular participation as a vital key to success in learning (Britton & Tesser, 1991). Given that time management strategy has been considered to involve self-regulation, long-term planning and sustained efforts of the learners, the regularity of learning is expected to be a strong indicator of time management strategy.

**Conclusion**

This study was concerned with the fact that adult learners frequently fail in the completion of online courses due to poor time management strategy. Given the exponential growth of online learning in adult education, this issue should be given priority. Herein, a critical proxy variable that enables the prediction of at-risk adult learners was suggested, in terms of time management. From the perspective of adult learning, time management was investigated, as a key competency helping learners to achieve academic goals. The results indicated that the regularity of learning works as a powerful indicator of the learners’ time management strategy, as it represents consistent effort and awareness to sustain perseverance with the pursuit of learning in a manner that could not be fully explained by the other two variables examined. It was discovered that adult learners accomplish successful completion of an online course when they maintain consistent awareness throughout the course. We expect this finding will contribute to the area of online adult learning.

An innovative approach to constructing proxy variables was also provided herein. As opposed to previous studies which largely relied on an inductive approach to the analysis of data, we attempted to build a solid foundation that enables more reliable data processing. The data processing method employed in this research can be applied to similar contexts where learners mostly engage in lecture-based courses, such as in the Massive Open Online Course or Open Course Ware emerging as a new trend in higher and adult education (Holford, Jarvis, Milana, Waller, & Webb, 2014). The process of constructing proxy variables showed the possibility of application in further research to investigate more sophisticated proxy variables representing other conceptual factors.

Until now, a large amount of research in the field of learning analytics and educational data mining has been conducted in a data-driven way, and frequently, with scarce theoretical background (Ferguson, 2012). Recently, however, learning analytics research is placing increased emphasis on learning and teaching theories as well, in contrast to its strong roots as a data-driven approach (Zaïane, 2001). The social and pedagogical usage of learning analytics is currently being actively discussed (e.g., Baker & Corbett, 2014) as researchers search to define it as a separate field by which to improve learning opportunities away from the business area (Ferguson, 2012; Jo, Kim, & Yoon, 2014; Long & Siemens, 2011). We expect that the balanced consideration of theory and data will lead to advancement in data-oriented research.

This study had several limitations. First, the specific time use of the participants could not be tracked due to technical limitations. For example, it could not be determined whether the learners were playing or pausing the lectures. With the log data, which better mapped actual time use on more specific behaviours, a more accurate analysis to capture real studying time could be made possible. Despite the fact that the total login time was used as an indicator of extensive studying time, the learners may have been greatly distracted by irrelevant information, a factor which would result in lower learning performance compared to that indicated by the recorded time. If the learner’s specific time use in LMS can be tracked, thus allowing extraction of the actual studying time from the log data at every moment, it would be possible to investigate the relationship between studying time and learning performance more clearly.
Furthermore, this study relied entirely on empirical verification in order to elicit the proxy variable. Although empirical verification is meaningful in certain contexts, the problem of concurrent validity can be posed when applied to other contexts. For substantiating concurrent validity, the use of an existing questionnaire measuring time management strategies can be considered for comparison.

Our ultimate goal is to develop an integrated LMS that provides precautionary intervention based on the diagnosis of learners made from the use of proxy variables. Future research should address the design and development of such interventions. For example, an LMS equipped with a monitoring function which employs proxy variables can be developed so that learners can be provided with opportunities to monitor their status of learning with a formative assessment view, which would enable reflection on their level of accomplishment. The development of such a learning management interface can be extended not only to providing more detailed interventions, but also toward supporting the continuous self-development of learners.

Acknowledgements

This work was supported by a grant from the National Research Foundation of Korea, funded by the Korean Government (NRF-2013S1A5A2A03044140).

References


ABSTRACT

While investment in technology for use in the classroom is increasing, studies still do not reveal significant improvements in learning. Investigations have shown that aligning the design of learning experiences with students’ needs creates synergy in the teaching and learning classroom processes. This can be achieved through orchestration, the guided integration of conventional and digital resources. This study seeks to investigate how orchestration increases student learning while taking into consideration their different pace of learning and monitoring the standards that the curriculum requires of a teacher. This is done by adopting a model which incorporates both the requirements of a school system’s curriculum, as well as the students’ specific needs based on their effective learning. It also seeks to determine which elements enrich the learning experience within an orchestrated framework. This study experimentally shows a significant increase in student learning when orchestrations are made available to teachers, when teachers successfully integrate technology into the classroom and when systematic use of the available resources is made. Furthermore, this study demonstrates the importance of addressing weaknesses in each student’s base of knowledge by adapting the activities in the classroom to their individual pace of learning.

Keywords

Orchestration, Individual learning paces, Integration of digital and non-digital resources, Systematicity of resource use

Introduction

Around the world, significant investment in educational technology is being made (Kozma 2011; Kozma et al., 2014; Schulte, 2014). Independent research has concluded that providing each child with a computer fosters the development of digital skills (Beuermann, Cristia, Cruz-Aguayu, Cueto, & Malamud, 2013). However, there is still no evidence to suggest that having access to a computer improves learning in Math or Reading (Beuermann, Cristia, Cruz-Aguayu, Cueto, & Malamud, 2013).

Integrating technology into educational processes is not simply a case of providing access, but in fact a way of promoting improvements in student learning (Hernández-Ramos, 2005). This begs a rethinking of the supposed connection between the potential provided by these new tools and the existing needs of educational processes. This is worth considering from both the perspective of the teachers, teaching process, as well as of the students, learning process In this sense, there is a need to align the design of learning experiences, i.e., line up the coherence and consistency of the pedagogical activities, with the way in which teachers address student needs (Hermans et al., 2008; Beetham & Sharpe, 2013; Suárez et al., 2013).

Orchestration can provide such an alignment within the process of integrating technology into the classroom. The literature has described orchestration as a tool that provides: robust and innovative forms of teaching and learning, by taking classroom ecology into consideration; efficiency, by facilitating the use of familiar, tested resources; adoptability, as it helps the teacher organize the necessary resources; and adaptability, as teachers can rely on strategies for adapting the resources according to contingencies that may arise (Dillenbourg, Nussbaum, Dimitriadis, & Roschelle, 2013). In this regard, and within the framework of schools that have access to technology, this study sought to understand: Is there a significant increase in student learning when the teacher uses orchestration? In
relation to this, our second question aims to identify: *What elements are associated with a greater increase in learning within a context of orchestrated learning experiences?*

Kennedy (2005) points out that the taxonomy regarding the tasks that a teacher should carry out has been based more on idealized conceptualizations than on reality or the teachers’ needs. For example, the paces of a class i.e., the different students learning rhythms, or maintaining a favorable learning environment considering aspects as usage of time, space and classroom resources, are fundamental necessities not necessarily considered. However, good teaching is not just determined by school-level factors or a teacher’s knowledge, beliefs, and attitudes. It is also determined by considering the needs of the students, as well as other classroom- and student-level factors (OECD, 2009). It has been highlighted the importance of student and teacher needs and interests when faced with the challenge of creating favorable learning environments in the classroom (Slavin, 2006). Wu, Tu, Wu, Le and Reynolds (2012) discuss paying attention to diversity within the classroom, focusing their research mainly on the consideration of a student’s individual progress within the learning process, i.e., considering the different paces of learning within a classroom. The third question arises: *When systematically implementing strategies that consider student learning pace, what type of impact does this have?*

By answering these three questions, this study looks to bring new information to the existing literature on orchestration. It also looks to provide the tools to empower teachers when integrating technology into the curriculum. One of the main differences in the design of the orchestration used in this study is that it takes into account each student’s individual pace of learning. Furthermore, it also monitors the standards that the curriculum requires of a teacher. Several different kinds of orchestration have been suggested in the past (Jurow & Creighton, 2005; Fischer & Dillenbourg, 2006; Dillenbourg et al., 2009; Prieto et al., 2011; Sharples, 2013). However, none of these take into consideration the curriculum standards that teachers must work towards, nor the students’ specific needs based on their effective learning.

This study begins by detailing the intervention model. Secondly, it explains the experimental design and the subsequent results of the analysis. Finally, in light of these results, the research questions are answered and the findings are discussed based on the current literature.

**Intervention model**

To answer the research questions, an experimental design was developed based on the implementation of a pedagogical strategy to integrate digital and non-digital resources. This strategy featured two components:

**Orchestration**

Orchestration is aimed at integrating the teaching processes with the available technology in schools. The used orchestration in this study has been based on a definition that has been proposed by several authors. The work by Fischer and Dillenbourg (2006) refers to orchestration as the coordination of different activities with the use of different resources and in different contexts. To complement this, Hämäläinen and Oksanen (2012) refer to orchestration as a way of supporting the teacher by providing them with a set of guidelines. Nussbaum, Dillenbourg, Dimitriaidis and Roschelle (2013), suggest four advantages to using orchestration: (1) it allows the teaching-learning process to be viewed as a whole, taking into account specific contexts; (2) it improves efficiency by allowing teachers to work with tried and trusted resources, (3) it improves adoptability, allowing the adoption of new resources to be presented to the teacher in a well-organized, coherent and attractive fashion; (4) it improves adaptability, acknowledging the fact that the teaching process is often dependent on what happens in real time in the classroom. In this sense, orchestration provides the teachers with explicit strategies for how and when to be flexible with the use of a new resource. This component is directly related to the first two research questions: “*Is there a significant increase in student learning when the teacher uses orchestration?*” and “*What elements are associated with a greater increase in learning within a context of orchestrated learning experiences?*”

For this study, the concept of orchestration refers to a set of logistical and pedagogical guidelines which support the teacher’s role in the teaching process (Drijvers, Doorman, Boon, Reed, & Gravemeijer, 2010). These guidelines result in a step-by-step guide for using the available resources (both digital and non-digital). This is done by
associating the potential of each resource with different learning objectives and stages of the curriculum (Nussbaum & Diaz, 2013). Although orchestration is based on classroom logistics, the guidelines also focus on how to implement the teaching activity, stimulating the learning experience and promoting interaction between the participants (Nussbaum et al. 2013). In this sense, orchestration is not proposed as being a technical manual or instructional guide. Instead, it looks to provide the teacher with scaffolding for dealing with the flexibility required by their learning environment. The systematic use by teachers of this kind of orchestration will provide their students with an articulated learning experience that aims to meet a specific learning objective (Díaz, Nussbaum & Varela, 2014). This articulation will have a positive impact on the learning outcome as it addresses both the curriculum standards as well as the different paces of learning that are present in the classroom. In this sense, the learning outcome is understood as the progress made by students not just in terms of the curriculum, but also how they reinforce their existing base in order to soundly acquire new knowledge.

Software for curricular work and software for each student’s individual needs

The objective of the former is to complement the learning experiences and relates specifically to curricular material, while for the latter it is related to addressing a student’s individual weaknesses in mathematics by respecting each student’s individual learning pace. This component will allow us to answer the third research question: “When systematically implementing strategies that consider student learning pace, what type of impact does this have?”

The study was carried out over twelve weeks and covered Fractional Numbers for the 5th grade mathematics syllabus. The curricular material was broken down into the following six units and organized progressively: (1) Reading and writing fractions in daily life; (2) Types of fractions and ordering fractions; (3) Multiplication and division; (4) Addition and subtraction of fractions with a common denominator; (5) Least common multiple; (6) Addition and subtraction of fractions without a common denominator.

Orchestration model

The pedagogical strategy consisted of the orchestration of digital and non-digital resources. The purpose of this was to frame the use of technology within the curriculum, focusing the teacher’s job on the pedagogical aspects of a class and leaving the logistical aspects to prior organization, as suggested by Mercer and Littleton (2007). As mentioned above, these specifications were arranged as a micro-sequencing of specific actions for the teacher and it was left to them to decide whether to strictly adhere to the sequencing or simply use it as a model (Figure 1). In this manner, the strategy was shaped by guidelines and resources.

Guidelines

**Orchestration of the use of resources:** an orchestration book with logistical and pedagogical suggestions for each resource included in the strategy (indicated below) for each unit. The orchestrations were based on the model described by Nussbaum & Diaz (2013).

**Opportunity for group explanations:** group activities with the teacher and the whole class aimed at analyzing examples relating to the content before continuing with the individual work.
Resources

**Presentation:** digital resources which aimed to allow the students to discover and learn more about the topic by presenting it in an everyday narrative to which they could relate. This allowed the students to become significantly more familiar with the new concepts.

**Software to practice fractions digitally:** software based on practicing fractions in their various representations: graphic, symbolic, number line, etc. The contents of this software included the six units that were covered during the study period. The variety of exercises in this software was based on each student’s performance and therefore respected their different learning paces (Alcohoado et al., 2012).

**Fractions workbook:** based on worksheets for developing process and reasoning skills. To complement the digital practice, the worksheets focused on problem solving exercises which framed the concepts within a specific context. This turned out to be essential for recognizing how and when to apply the mathematical processes that had been covered during the classes.

**Mathematical challenge:** based on worksheets which included problems that were more difficult than those contained in the fractions workbook. All of these problems incorporated the use of algebra and their objective was to develop application, analysis and evaluation skills.

**Game-based activity:** a tool which allowed students to establish connections between game-based experiences and the experiences encountered in the presentation, software, and worksheets. These activities were carried out in pairs or groups of three and were included so as to reinforce the students’ ability to apply the material to new contexts.

**Software to practice operations:** students practiced at their own pace the different arithmetic operations working individually weaknesses that were detected by this tool (Alcoholado et al., 2011). In this way, the strategy not only focused on the specific curricular content, but also on the need to build a solid base of mathematical knowledge for each child.

This set of resources was implemented over a two-week period, which was how long it took to cover each unit.

**Training and coaching**

The experimental work included a transfer process designed to ensure that the pedagogical strategy was understood and adopted by both the teacher and the institution. Without this level of adoption, it would have been difficult to implement the work systematically. These four components were:

**Training:** designed based on the features of the pedagogical strategy, following the hypotheses of previous studies. For example, Lawless & Pellegrino (2007) state that teacher training is essential for effective access to updates for resources and strategies that integrate digital elements in order to reinforce the teaching process. This process included two group meetings with the teachers from the treatment group, each of them three hours long.

**Coaching in the classroom:** the transfer was extended to each teacher’s real work environment by carrying out three coaching sessions in the place where the technology would be integrated. This coaching was included because of studies such as the one done by Slavin (2006). Slavin proposes the need to consider what happens inside the classroom, as well as the social association between the actors. In particular, he highlights the importance of teacher and student needs and interests when faced with the challenge of creating favorable learning environments. The coaching was mainly based on the logistical administration of the classroom, making sure to keep the teacher focused on pedagogical and instructional issues and not on the handling of the technology itself. At the end of each of these sessions, the teachers were given feedback on their work, reinforcing the adoption of the strategy by analyzing which elements had already been incorporated and which were yet to be incorporated.

**Coaching for planning:** the orchestration suggests how to manage time, space and the use of resources, as well as when to make key interventions while the students work. In addition to this, each teacher was also visited twice in order to help them revise the orchestration based on opportunities and barriers that were specific to their context.
During these visits, elements and situations related to the curriculum and the different learning paces observed by teachers were reviewed.

**Coaching for institutional planning:** carried out three times during the project (in week 1, week 4, and week 8) and solely with the participation of representatives from the school’s administration, i.e., without the presence of teachers. The literature (Leithwood, Seashore, Anderson, & Wahistrom, 2004) reveals the importance of involving the administration in teaching and learning processes. Thus, in parallel to the work with the teachers, work was also done with a member of the administrative team from each school on scheduling the use of devices, both for the participating teacher as well as the rest of the teachers at the institution. Furthermore, some support tasks for teachers were reviewed, with the recommendation of assigning responsibilities for technical support to someone within the establishment.

**Methods**

The study was based on an experimental design with a randomly selected sample that was stratified according to environmental characteristics. These characteristics are explained in the section under Sample. This design was adopted so as to analyze the effect of using orchestration on the student learning process. This effect is understood as the increase in learning between a pre- and post-test, detailed in the section under Instruments.

**Sample**

The study was carried out with 5th grade students from a group of twenty urban public schools in Barranquilla, Colombia, which had the necessary technological infrastructure to integrate 1:1 computer-based experiences into student work. Of the twenty participating schools, ten were randomly assigned to the control group and ten to the treatment group. The random selection was stratified, meaning that the total sample of schools was fragmented into relatively homogenous strata. This was done by taking into account two characteristics that were observed during initial contact and that could have affected the results of the intervention: the gender composition of the school and school organization. In terms of the first characteristic, some of the schools were mixed (with girls and boys) \( n = 15 \) and others were all-girls \( n = 5 \). With regards to the second characteristic, the schools were classified into two groups according to how organized they were for the start of classes—if they had official registers, programmed schedules, and assigned sections \( n = 13 \), or not \( n = 7 \), when the initial assessment was conducted. These criteria provided an institutional management parameter before the intervention, something which has been discussed in several studies regarding the impact that this has on student learning (Leithwood, Seashore, Anderson, & Wahistrom, 2004). Once random selection was performed, it was verified that there were no statistically significant differences between the treatment and control groups in any of the variables given by the Secretariat of Education of Barranquilla in its database, such as: school size, ratio of teachers/students per class, socioeconomic status of the students, in addition to the two aforementioned characteristics (Table 1).

Despite using this technique, the treatment group ended up being composed of schools with a lower socioeconomic status than the control group. Although a general socioeconomic status is given for each school, it does not necessarily reflect the socioeconomic status of each specific student’s family. Furthermore, there was a difference in favor of the control group regarding curricular management and student learning, as reported by the 2012 SABER Test scores (a national test given by the Ministry of Education of Colombia) for 5th grade Math students.

The sample of participating students was composed of 702 children, the majority of which came from the southern city of Barranquilla. These students were enrolled in 5th grade at the start of the school year and ranged from 8 to 14 years of age (average age = 10.27), 60% of the total sample were girls. Given that 25% of the schools included in the sample were all-girl schools, 46.47% of the female students were concentrated in these schools. The final sample for analysis was composed of 531 students, due to the failure of some students to complete both the pre- and post-test. From the 171 children not considered in the sample, 73 were from the control group and 98 from the treatment group. The analyses were carried out using a final sample of the control group composed of 263 students (113 boys and 150 girls) and the treatment group of 268 students (95 boys and 173 girls).
Table 1. Sample

<table>
<thead>
<tr>
<th>School</th>
<th>Gender composition</th>
<th>School organization</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mixed</td>
<td>Unorganized</td>
<td>Control</td>
</tr>
<tr>
<td>2</td>
<td>Mixed</td>
<td>Unorganized</td>
<td>Control</td>
</tr>
<tr>
<td>3</td>
<td>Mixed</td>
<td>Organized</td>
<td>Control</td>
</tr>
<tr>
<td>4</td>
<td>Mixed</td>
<td>Unorganized</td>
<td>Control</td>
</tr>
<tr>
<td>5</td>
<td>Mixed</td>
<td>Organized</td>
<td>Control</td>
</tr>
<tr>
<td>6</td>
<td>All-girls</td>
<td>Organized</td>
<td>Control</td>
</tr>
<tr>
<td>7</td>
<td>All-girls</td>
<td>Organized</td>
<td>Control</td>
</tr>
<tr>
<td>8</td>
<td>Mixed</td>
<td>Unorganized</td>
<td>Control</td>
</tr>
<tr>
<td>9</td>
<td>Mixed</td>
<td>Organized</td>
<td>Control</td>
</tr>
<tr>
<td>10</td>
<td>Mixed</td>
<td>Organized</td>
<td>Control</td>
</tr>
<tr>
<td>11</td>
<td>Mixed</td>
<td>Organized</td>
<td>Treatment</td>
</tr>
<tr>
<td>12</td>
<td>Mixed</td>
<td>Organized</td>
<td>Treatment</td>
</tr>
<tr>
<td>13</td>
<td>All-girls</td>
<td>Organized</td>
<td>Treatment</td>
</tr>
<tr>
<td>14</td>
<td>Mixed</td>
<td>Unorganized</td>
<td>Treatment</td>
</tr>
<tr>
<td>15</td>
<td>Mixed</td>
<td>Organized</td>
<td>Treatment</td>
</tr>
<tr>
<td>16</td>
<td>All-girls</td>
<td>Organized</td>
<td>Treatment</td>
</tr>
<tr>
<td>17</td>
<td>All-girls</td>
<td>Organized</td>
<td>Treatment</td>
</tr>
<tr>
<td>18</td>
<td>Mixed</td>
<td>Unorganized</td>
<td>Treatment</td>
</tr>
<tr>
<td>19</td>
<td>Mixed</td>
<td>Unorganized</td>
<td>Treatment</td>
</tr>
<tr>
<td>20</td>
<td>Mixed</td>
<td>Organized</td>
<td>Treatment</td>
</tr>
</tbody>
</table>

There was no sampling of teachers as the analysis of this study was based on the students’ progress, i.e., increase in learning, controlling for the schools’ environmental characteristics as highlighted previously. In this sense, the teacher variable was considered as an independent variable for all of the analyses.

Instruments

Learning Progress Test: In order to establish the impact of the pedagogical strategy on student learning, both the control group and the treatment group were evaluated before and after the intervention (pre- and post-test). This test was developed by a team of external evaluators who had knowledge of the Colombian curriculum and local terminology. The test was designed based on the standards defined in the Colombian national curriculum for assessing knowledge of fractions. It was also divided into six categories, with each category featuring six questions. Each category was also directly related to one of the six units that comprised the curricular contents covered in this study. The final test therefore contained 36 questions. The reliability analysis showed a Cronbach’s Alpha (Cronbach, 1951) of .80, indicating high levels of reliability for the student sample of this study. Both the pre- and post-test were administered by an external evaluator in each school. The tests were held in the students’ own classroom, though their regular teacher was not present. Each test lasted for 60 minutes.

Logs: In order to register how each teacher adapted the orchestration to their specific needs, a log was used to record the use (or lack of use) of the different components for each session. This record allowed information to be gathered to support the coaching for planning, as well as to create an index for the systematicity of resource use for each unit. The purpose of this was so that it could be used as an independent variable with which to analyze increases in student learning.

Observation Guidelines: during the coaching sessions in the classroom, the presence or absence of different elements of the pedagogical strategy was recorded so that feedback could be given to the teachers with regards to what had been observed, as well as to generate an index for teaching practices related to the pedagogical strategy. This was done so that it could be used as an independent variable with which to analyze increases in student learning.
Statistical techniques

In order to answer the first research question: “Is there a significant increase in student learning when the teacher uses orchestration?” analysis of variance and analysis of covariance was used to compare the increase in learning between the control group and the treatment group.

Analysis of variance (ANOVA): As explained in the description of the experimental design, the sample was randomly selected. However, two variables were identified following the sample selection that could have affected the results of the intervention in the treatment and control schools. These variables were the school’s socioeconomic status and their learning and curricular management index, taken from their scores on the 2012 SABER Test in 5th grade Math (Colombian SABER Test for 5th grade Math students during 2012 [ICFES http://www.icfes.gov.co/]). In order to determine whether these differences were statistically significant, a comparison was made using an analysis of variance (ANOVA). This analysis was performed for each of the dependent variables, i.e., the school’s socioeconomic status and their learning and curricular management index.

Analysis of covariance (ANCOVA): having established heterogeneity between the control group and treatment group using the variables that were identified once the study had begun (the school’s socioeconomic status and their learning and curricular management index), the decision was made to compare the two groups using an analysis of covariance (ANCOVA). This analysis removes the heterogeneity between groups stemming from characteristics which are not controlled by the methodological design. In this case, the aim was to remove the heterogeneity of the 2012 SABER Test 5th grade Math results and the schools' socioeconomic status, thus removing the bias of these variables when comparing the final increase in learning. A second advantage of having used this technique is that it improved the strength of the study as a consequence of controlling for the aforementioned covariates (Morrison, 1990).

Pearson’s correlation: In order to answer the second and third research questions: “What elements are associated with a greater increase in learning within a context of orchestrated learning experiences?” and “When systematically implementing strategies that consider student learning pace, what type of impact does this have?” a series of analyses were performed using Pearson’s correlation. This technique is often used in studies of human behavior, where the magnitude of the correlation is of primordial interest (Bonett & Wright, 2000). The increase in learning for each sub-group of students within the treatment group was always used as a variable in these correlations. This was in addition to a range of other variables that are described below.

Construction of variables

The following sections describe how each variable was constructed and then later used to analyze increases in learning reached by students at the end of the study. These variables were focused on four dimensions: Student, Systematicity of Resource Use, Teachers and Schools. Data is shown in Tables 2, 3 and 4.

For the Students, final attendance records were collected from each student who participated in the study. Based on this record, the “Attendance” variable was defined by calculating the percentage of attendance over the total number of days for the twelve-week study. The second variable for students, “Initial test score,” was defined by considering the student’s previous knowledge of fractions as shown in the pre-test.

For the Systematicity of Resource Use, information was gathered using a record log filled out by each teacher during the twelve-week study. Through this log, teachers stated whether they used the resources associated with each unit completely, partially, or not at all, assigning 1 point, 0.5 points, or 0 points, respectively. Thus, the percentage of use for each resource was calculated for each of the six units. The resources that were subsequently used as variables are: Presentation, Group Explanations, Workbook, Fractions Software, Mathematical Challenge, Operations Software, and Game-based Activity. Data is shown in Table 2.
shown in Table 3.

For the Teachers, the field researcher used a check-list during the visits to record the presence, value 1, or absence, value 0, of indicators related to the work expected from teachers when implementing the strategy’s resources. These indicators were presented during the training period with teachers and also used as feedback indicators at the end of each visit. In this way, an average performance index was calculated for each teacher based on the three recorded visits. The criteria that were subsequently used as variables are: Subject Knowledge (knowledge of the mathematics covered), Technological Proficiency (user’s level of knowledge of computers), Class Structure (presence of activities relating to the introduction, practice and review sections of a class), Use of Orchestration (following the orchestration suggestions in the classroom), Explanation of Material (giving further explanations of concepts and procedures when required by the student learning pace), Feedback for Students (feedback given by the teacher in response to students’ answers), Group Management (classroom management, fostering a positive environment for learning), and Management of Digital Resources in the Classroom (implementing strategies for handing out digital devices in the classroom and collecting them back in, organizing spaces that are suitable for their use, etc.). Data is shown in Table 3.

<table>
<thead>
<tr>
<th>School</th>
<th>Presentation Score (x)</th>
<th>Group explanations Score (x)</th>
<th>Workbook Score (x)</th>
<th>Fractions software Score (x)</th>
<th>Mathematical challenge Score (x)</th>
<th>Operations software Score (x)</th>
<th>Game-based activity Score (x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>1.00</td>
<td>0.83</td>
<td>0.75</td>
<td>0.75</td>
<td>0.83</td>
<td>0.33</td>
<td>0.60</td>
</tr>
<tr>
<td>12</td>
<td>0.92</td>
<td>1.00</td>
<td>0.83</td>
<td>0.67</td>
<td>0.83</td>
<td>0.75</td>
<td>0.70</td>
</tr>
<tr>
<td>13</td>
<td>0.75</td>
<td>0.83</td>
<td>1.00</td>
<td>0.92</td>
<td>0.75</td>
<td>0.83</td>
<td>1.00</td>
</tr>
<tr>
<td>14</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.50</td>
</tr>
<tr>
<td>15</td>
<td>0.67</td>
<td>0.50</td>
<td>0.67</td>
<td>0.58</td>
<td>0.42</td>
<td>0.50</td>
<td>0.20</td>
</tr>
<tr>
<td>16</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.50</td>
</tr>
<tr>
<td>17</td>
<td>0.67</td>
<td>0.75</td>
<td>1.00</td>
<td>0.67</td>
<td>0.83</td>
<td>0.58</td>
<td>1.00</td>
</tr>
<tr>
<td>18</td>
<td>0.83</td>
<td>0.83</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.80</td>
</tr>
<tr>
<td>19</td>
<td>0.25</td>
<td>1.00</td>
<td>0.92</td>
<td>0.92</td>
<td>0.67</td>
<td>0.75</td>
<td>0.30</td>
</tr>
<tr>
<td>20</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.92</td>
<td>0.92</td>
<td>1.00</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Note. Max = 1; Min = 0.

The fourth dimension, School variables, includes contextual elements taken from the school management indicators. The first of these indexes, Leadership and management, was calculated considering three roles (1 if the role was present; 0 if absent): management (observing institutional backing of the time given to the participating teacher in order to prepare technology-based lessons), coordination (observing scheduling and availability of the technological equipment in the classes when the participating teacher needed them), and technical support (observing the operational maintenance of the technological equipment used during the students’ learning experiences). The score for this variable ranged from 0 to 3 points, considering the addition of the values obtained for each role. Data is shown in Table 4.

Table 3. Teachers

<table>
<thead>
<tr>
<th>School</th>
<th>Subject knowledge Score (x)</th>
<th>Technological proficiency Score (x)</th>
<th>Class structure Score (x)</th>
<th>Use of orchestration Score (x)</th>
<th>Explanation of material Score (x)</th>
<th>Feedback for students Score (x)</th>
<th>Group management Score (x)</th>
<th>Management of digital resources in the classroom Score (x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>0.00</td>
<td>0.00</td>
<td>0.63</td>
<td>0.67</td>
<td>1.00</td>
<td>1.00</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>12</td>
<td>0.33</td>
<td>0.33</td>
<td>0.56</td>
<td>0.89</td>
<td>1.00</td>
<td>0.89</td>
<td>0.78</td>
<td>0.67</td>
</tr>
<tr>
<td>13</td>
<td>1.00</td>
<td>0.66</td>
<td>0.52</td>
<td>0.78</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>14</td>
<td>1.00</td>
<td>1.00</td>
<td>0.63</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.89</td>
<td>0.89</td>
</tr>
<tr>
<td>15</td>
<td>0.33</td>
<td>0.33</td>
<td>0.59</td>
<td>0.78</td>
<td>0.89</td>
<td>1.00</td>
<td>0.89</td>
<td>0.83</td>
</tr>
<tr>
<td>16</td>
<td>1.00</td>
<td>1.00</td>
<td>0.74</td>
<td>0.89</td>
<td>0.89</td>
<td>1.00</td>
<td>0.89</td>
<td>1.00</td>
</tr>
<tr>
<td>17</td>
<td>0.66</td>
<td>0.66</td>
<td>0.67</td>
<td>1.00</td>
<td>0.89</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>18</td>
<td>0.66</td>
<td>1.00</td>
<td>0.56</td>
<td>1.00</td>
<td>0.89</td>
<td>1.00</td>
<td>0.89</td>
<td>0.89</td>
</tr>
<tr>
<td>19</td>
<td>0.66</td>
<td>0.66</td>
<td>0.67</td>
<td>0.67</td>
<td>1.00</td>
<td>0.89</td>
<td>1.00</td>
<td>0.67</td>
</tr>
<tr>
<td>20</td>
<td>0.00</td>
<td>0.00</td>
<td>0.19</td>
<td>0.67</td>
<td>0.33</td>
<td>0.56</td>
<td>0.33</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Note. Max = 1; Min = 0.
The second variable defined the Type of Computer present in the institution: 1 for netbooks; 2 for laptops; 3 for PCs in a computer lab. Data is shown in Table 4.

As a proxy for the level of learning achieved by the students from each school in the topics covered by this study, the third variable was the National test outcome in Math, obtained through the school’s average scores on the nationwide Colombian SABER Test for 5th grade Math students during 2012 (ICFES http://www.icfes.gov.co/). The SABER test is administered in 3rd, 5th and 9th grade, for both reading comprehension and mathematics. The 5th grade Math test was selected as it was conducted with the students of the previous generation of the school at 5th grade, being the best possible approximation for the level of learning for children of that age in each school. Data is shown in Table 4.

The final variable in this dimension is the socioeconomic status of the schools. Five social economics status are defined being 5 the highest; the schools present in the sample presents the lower middle class. This data was obtained from the Colombian Institute for the Evaluation of Education (ICFES http://www.icfes.gov.co/). Data is shown in Table 4.

<table>
<thead>
<tr>
<th>School</th>
<th>Leadership and management</th>
<th>Type of computer</th>
<th>National test outcome in math</th>
<th>Socioeconomic status of the schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Score (X)</td>
<td>Score (X)</td>
<td>Score (X)</td>
<td>Score (X)</td>
</tr>
<tr>
<td>11</td>
<td>2.00</td>
<td>2.00</td>
<td>290</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>3.00</td>
<td>2.00</td>
<td>322</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>3.00</td>
<td>3.00</td>
<td>311</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>2.00</td>
<td>2.00</td>
<td>299</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>2.00</td>
<td>2.00</td>
<td>200</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>0.00</td>
<td>1.00</td>
<td>223</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>3.00</td>
<td>1.00</td>
<td>315</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>0.00</td>
<td>1.00</td>
<td>301</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>1.00</td>
<td>2.00</td>
<td>285</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>2.00</td>
<td>1.00</td>
<td>229</td>
<td>1</td>
</tr>
</tbody>
</table>

Results

Control and treatment group analysis

In order to verify the homogeneity between the treatment and control groups, an analysis of variance (ANOVA) was carried out. The results of this showed significant differences between schools in terms of their latest test scores on the SABER Test in 5th grade Math (p < .001) and the socioeconomic status of each school (p < 0.001). By doing so, a lack of homogeneity between the groups was established. It was therefore decided that an analysis of covariance (ANCOVA) would be carried out in order to control for these variables and make substantial comparisons.

When using the two indicated variables as control variables, belonging to the treatment group turns out to favor increases in student learning (p < 0.00). Table 5 shows the SABER Test scores and socioeconomic status of each educational institution.

<table>
<thead>
<tr>
<th>Type</th>
<th>School</th>
<th>Socioeconomic status</th>
<th>Total for the group</th>
<th>SABER score</th>
<th>Total for the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>307</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>269</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>293</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>355</td>
<td>3297</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>259</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>347</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>360</td>
<td></td>
</tr>
</tbody>
</table>
Among the schools in the treatment group, differences were observed in terms of learning increases. A subsample was created to explore contextual variables associated with a greater increase in learning and based on a correlational analysis regarding three variables that showed statistical correlation. These were: Attendance (corr. = .20, \( p < .001 \)), Use of Orchestration (corr. = .19, \( p < .001 \)), and the Highest mean increase in learning between pre- and post-tests (corr. = .26, \( p < .001 \)) (variable taken from the difference in score between the two tests). The five schools that showed a correlation in these three criteria were called the treatment group subsample. For this subsample, the associations between the mean increase in learning at each school and the variables listed in the section under Statistical techniques – Construction of Variables were explored. The results of the correlational analysis are shown in Table 6.

### Table 6. Exploring associations with an increase in learning

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Variable</th>
<th>Correlation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>Attendance</td>
<td>.11</td>
<td>.30</td>
</tr>
<tr>
<td></td>
<td>Initial test score</td>
<td>-.35</td>
<td>.00 ***</td>
</tr>
<tr>
<td>Systematicity of resource use</td>
<td>Presentation</td>
<td>.08</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td>Group explanations</td>
<td>.29</td>
<td>.00 **</td>
</tr>
<tr>
<td></td>
<td>Workbook</td>
<td>.38</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>Fraction software</td>
<td>.53</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>Operations software</td>
<td>.55</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>Mathematical challenge</td>
<td>.25</td>
<td>.00 **</td>
</tr>
<tr>
<td></td>
<td>Game-based activity</td>
<td>.35</td>
<td>.00 ***</td>
</tr>
<tr>
<td>Teacher</td>
<td>Subject knowledge</td>
<td>.59</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>Technological proficiency</td>
<td>.50</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>Class structure</td>
<td>.33</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>Use of orchestration</td>
<td>.37</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>Explanation of material</td>
<td>.43</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>Feedback for students</td>
<td>.08</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>Group management</td>
<td>.12</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td>Management of digital resources</td>
<td>-.03</td>
<td>.73</td>
</tr>
<tr>
<td>School</td>
<td>Institutional management</td>
<td>-.05</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td>Type of computer</td>
<td>-.46</td>
<td>.00 ***</td>
</tr>
<tr>
<td></td>
<td>SABER 5th grade math score</td>
<td>.27</td>
<td>.00 **</td>
</tr>
<tr>
<td></td>
<td>Socioeconomic status of institution</td>
<td>-.03</td>
<td>.74</td>
</tr>
</tbody>
</table>

Note. ***\( p < .000 \), **\( p < .001 \), *\( p < .05 \).

For the first dimension, an association between a greater increase in learning and a lower score on the initial test (corr. = -.35, \( p < .001 \)) can be identified. This indicates that students who scored lower on the pre-test showed a greater increase in learning in the post-test. However, no association was observed with the class attendance variable.
For the second dimension (Systematicity of Resource Use), a positive association can be shown between a greater mean increase in learning and more frequent use of the workbook (corr. = .38, p < .001), fractions software (corr. = .53, p < .001), operations software (corr. = .55, p < .001), mathematical challenge (corr. = .25, p < .001), game-based activity (corr. = .35, p < .001) and group explanations (corr. = .29, p < .001). The only resource that did not show any association was presentation.

For the Teacher dimension, there is a positive association between teachers who were observed with higher levels of subject knowledge (corr. = .59, p < .001) and those with higher levels of technological proficiency (corr. = .50, p < .001). Similarly, there is a positive association between classes that follow a structure (introduction, practice and review) (corr. = .33, p < .001), and teachers who follow the orchestration suggestions (corr. = .37, p < .001) and focus on explaining the material instead of just covering it superficially (corr. = .43, p < .001). No association was observed between the feedback given by teachers to students, their group management skills, nor their management of digital resources in the classroom.

For the School dimension, the evidence shows that schools that use netbooks had a greater average increase in learning (corr. = -.46, p < .001). As indicated in the Methods section, the three types of computers present in schools were categorized on a scale ranging from 1 to 3, where 1 was assigned for netbooks and 2 for laptops (the subsample of schools did not contain cases with PCs in laboratory, i.e., category number 3). Thus, in this case, the negative sign reflects the inverse association of the netbook in category 1, indicating an association between the schools that used this type of device and a greater increase in student learning, as compared to schools that used notebooks. Additionally, the score achieved by schools on the 2012 SABER Test for 5th grade Math is also associated with better performance on the progress test used in this study (corr. = .27, p < .001). This means there is a positive association between schools that had a higher score on the national test and schools that saw a greater increase in student learning.

**Conclusion and discussion**

From the results of this study, it is possible to answer the three research questions that were posed. With regards to the first research question, “Is there a significant increase in student learning when the teacher uses orchestration?” it can be observed, within the context of this particular study, that there is indeed an increase. Students whose teachers had orchestrations at their disposal for integrating technology into the classroom showed a greater increase in learning between pre- and post-tests versus students whose teachers did not. In this sense, these findings enhance the appraisal of orchestration previously made by other researchers as being a key component of technology provision programs in educational settings. Orchestration has been established as the concrete link between technology and pedagogy, and its potential for achieving learning objectives has been revealed—in this particular case, in settings of lower socioeconomic status and with lower indices of curricular management (SABER Test). This disadvantage was established in the first analysis indicated in the Results section, by comparing the schools’ socioeconomic status in the control group and treatment group, and finding that the treatment schools had a significantly lower status (p < .001) than the control schools. Also, a comparison was made between groups regarding the curricular management index (SABER Test), again establishing a statistical difference (p < .001) between groups, showing that the treatment group comprised schools with a lower index than the control group schools.

With regards to the second question, “What elements are associated with a greater increase in learning within a context of orchestrated learning experiences?” there are a number of components that are directly associated with a greater increase in learning. To analyze these components, we will refer to the dimensions defined in the Methods section. First, regarding student-level variables and in line with the response to the first research question, the orchestrated pedagogical strategy described in this study is associated with a greater increase in learning for students who score lower on the initial test (pre-test). This shows a positive impact where students showed less previous knowledge of fractions. In this sense, the orchestrated strategy helped increase learning in students who were less prepared to learn more complex elements, i.e., students who were at a disadvantage when compared to their peers.

With regards to the systematicity of use of the resources included in the orchestrated strategy, it is worth noting that frequent use of most of the resources was associated with a greater increase in learning. Those resources included: opportunities for group explanations of the content being covered, workbooks based on worksheets, fraction software
that included the curricular flow associated with the six units, mathematical challenges demanding higher cognitive skills, game-based activities using cutouts and concrete experiences with the content, and the operations software which included the flow of mathematical base knowledge that followed each child’s work at their own individual pace. The only resource not associated with a greater increase in learning was presentation, which based the introduction of new material on everyday narratives.

In addition to this, a greater increase in student learning in the treatment group was also demonstrated at teacher level. Both knowledge of mathematics and technological proficiency were indicators associated with a greater increase in learning. Teachers who structured their classes with introduction, practice and review activities; followed the orchestration suggestions; and explained the material well instead of just covering the material quickly, were also associated with an increase in student learning. No positive or negative associations were observed between teacher feedback for students, group management, and management of digital resources in the classroom.

With regards to the school environment, only two variables showed a correlation: the type of computer present in the school and the average score on the 2012 SABER Test for 5th grade Math. The use of netbooks was directly and positively associated with a greater increase in learning. This may be due to the mobility of the technology in the classroom, the students’ natural learning environment, without necessarily having to move somewhere else to work on computers. Finally, the significant correlation with the SABER Test scores may be due to better institutional management of the curriculum for student learning (http://www.icfes.gov.co/).

Finally, to respond to the last research question, “When systematically implementing strategies that consider student learning pace, what type of impact does this have?” we analyze issues related to the systematicity of resource use, paying close attention to the correlation reported by the use of the two software included in the strategy. Both software consistently reported a strong correlation between systematic use and an increase in student learning. This shows a positive association between the integration of curricular material and the material reviewed to address students’ individual weaknesses, taking into consideration the different learning paces present in the classroom.

When analyzing each software separately, the correlation of greater systematicity in the use of the fractions software, which covered curricular material, is 53%. This is an important figure when analyzing each strategy component separately, as while the remaining resources also has a positive association, it is to a lesser extent. This association could also be explained by the fact that the fractions software was designed based on the six units covered during the study’s twelve-week period. In addition to this, the pre- and post-tests measured knowledge of those six units. The operations software respected each student’s learning pace, helping them to consolidate algorithms that are a basic element of mathematical thinking and key to acquiring new and more abstract, complex concepts. As a result, there was a 55% correlation between this software and the students that demonstrated a greater increase in learning. This demonstrates the importance when covering curriculum contents of not just working on standards defined by ministerial programs, but also addressing the mathematical deficiencies that students may present and, in doing so, better consolidating their new learning.

It is worth highlighting the importance of digital content in regards to this last question, continuing with NESTA’s line of thought when they make a direct allusion to the use of technology rather than the technology itself (Luckin, Bligh, Manches, Ainsworth, Crook, & Noss 2012). The content of the hardware intended for student learning is also important and must be orchestrated or coordinated according to other learning experiences and ideally contain a component of work that is individually paced. By considering achievements and progress in relation to the individual student instead of set learning standards, this type of individual work can have an impact not just on learning but also on motivation.

The results from this study are similar to the results of a study carried out by BECTA (Somekh et al., 2011). The main difference between the two lies in how the digital component in both studies addressed the students’ and teachers’ needs. In the BECTA study, the digital work only took into account the students’ needs by allowing them to progress at their own rate. In our study, however, we also took into consideration the teachers’ needs to advance through the curriculum with the students. Orchestration, which was not present in the BECTA study, was of fundamental importance to us as it took into consideration the duality of student vs. teacher needs in the classroom. This in turn allowed us to incorporate digital and conventional resources more coherently. Considering that our study compared groups that worked with and without orchestration, while the BECTA study looked at incorporating a digital component without orchestration and integrated support for the school community, which we did not, future
studies should analyze the value added by orchestration in contexts such as the one in which BECTA carried out their study.

The findings discussed in this study suggest further investigation of orchestration and integration of curricular contents, as well as individual learning paces in the classroom. This should be done within two new frameworks. The first of these would be to conduct the study over a longer period of time, i.e., at least a full school year, which could cover learning associated with curricular contents for the whole grade level. In addition, the analyses could be fully reviewed with the systematicity of resource use maintained over time. Second, it would be interesting to analyze this experience in other areas of learning, expanding the scope to include other areas of the curriculum.

Acknowledgements

The research was funded by the Inter-American Development Bank (IADB) and supported by the Center for Research on Educational Policy and Practice, Grant CIE01-CONICYT, from Chile.

References


Towards a Social Networks Model for Online Learning & Performance

Kon Shing Kenneth Chung* and Walter Christian Paredes
Complex Systems Research Group, Project Management Program, The University of Sydney, Australia // ken.chung@sydney.edu.au // walter.paredes.e@gmail.com

*Corresponding author

(Submitted May 28, 2014; Revised December 16, 2014; Accepted December 22, 2014)

ABSTRACT
In this study, we develop a theoretical model to investigate the association between social network properties, “content richness” (CR) in academic learning discourse, and performance. CR is the extent to which one contributes content that is meaningful, insightful and constructive to aid learning and by social network properties we refer to its structural, position and relationship attributes. Analysis of data collected from an e-learning environment shows that rather than performance, social learning correlates with properties of social networks: (i) structure (density, inter-group and intra-network communication) and (ii) position (efficiency), and (iii) relationship (tie strength). In particular, individuals who communicate with internal group members rather than external members express higher tendencies of “content richness” in social learning. The contribution of this study is three-fold: (i) a theoretical development of a social network based model for understanding learning and performance, which addresses the lack of empirical validation of current models in social learning; (ii) the construction of a novel metric called “content richness” as a surrogate indicator of social learning; and (iii) demonstration of how the use of social network analysis and computational text-mining approaches can be used to operationalize the model for studying learning and performance. In conclusion, a useful implication of the study is that the model fosters understanding social factors that influence learning and performance in project management. The study concludes that associations between social network properties and the extent to which interactions are “content-rich” in e-Learning domains cannot be discounted in the learning process and must therefore be accounted for in the organizational learning design.

Keywords
Social networks, Structural holes, Strength of weak ties, Individual learning, Group learning, Social learning, Performance, Situated learning, Connectivism, e-Learning, Learning analytics

Introduction
New ways of social interaction, learning and performance (in the form of learning outcomes) have been heavily influenced by the appearance of different Internet technologies and Massive Open Online Courses (MOOCs) such as Coursera and Khan Academy. These platforms provide stimulating and interactive channels of communication that foster the creation and exchange of user-generated content for learning. Learning is a social process of progressive knowledge acquisition that is shaped by individuals and their interaction with others who can contribute new ideas, opinions and experiences (Rosen, 2010). Thus, social networks within which people interact play an important role in the learning process by expanding the possibilities of learners to reach new sources of information, and by providing (existing and latent) channels for open collaboration among individuals (Haythornthwaite, 2002b; Greenhow, 2011). Despite the attractive advantages presented by scholars about the impact of technology in the learning process, there is still a lack of understanding of the dynamics of social interaction within learning communities. Therefore, the motivating questions that inspire this study are (i) Is there an interplay between social networks, learning and performance? (ii) If so, what is the role of social learning in the inherent relationship between properties of social networks and performance? (iii) How does one quantify and measure learning within a social context? (iv) How does one account for social network properties of structure, relations and position in modelling learning for the purpose of learning analytics?

In this exploratory study, we develop a theoretical model based on social learning and social network theories to understand how knowledge professionals engage in learning and performance, both as individuals and as groups. The study also focuses on how an individual’s levels of participation and depth of engagement in the learning process are impacted by social interactions. The following section conducts a review of relevant learning theories and social network theories to arrive at a social networks model for understanding learning and performance. The model is then tested within an e-learning domain in a Group of Eight (Go8) university in Australia. The Go8 (group of Eight) comprises eight leading universities in Australia (www.go8.edu.au). The section on Results discusses the findings of
the study, followed by a discussion of these results in light of theory. Implications of the study, limitations and avenue for future research are provided in the conclusion.

**Conceptual foundations**

There are many different theories describing the general aspects and perspectives of human learning. According to Lave and Wenger (1991), *Situated Learning Theory* (SLT) supports the notion that learning takes place in social situations where individuals develop skills by interacting with others who can provide them with insights about existing knowledge and previous personal experiences within a “community of practice.” They argue that knowledge is acquired in a context that normally involves the practical use of that knowledge, in what we commonly know as “learn by doing.” Thus, social interactions play a fundamental role in order to involve learners in a community that has a determined behavior and beliefs. In enabling situated learning, social technologies in particular have great potential and educational value due to their inherent capacity to increase learners’ motivation and engagement through participation and knowledge creation (Greenhow, 2011). Due to its notion of community, SLT presents an interesting perspective to analyze learning and performance in that it is closely aligned with the social networks perspective where a social phenomenon such as learning can be understood in terms of the interactions amongst the learners and their mentors.

**Towards a model for learning and performance through social networks**

Social network analysis (SNA) enables the study of social systems from a structural perspective through the identification of behavioral patterns based on node and tie attributes (Freeman, 2006). One of the fundamental principles motivating studies regarding the relationship among learning, performance and social networks is that an individual’s social structure can influence an individual’s access to valuable resources (Leavitt, 1951; Brass, 1984; Coleman, 1988; Ibarra, 1993; Borgatti, 2005). Those resources that are rich in new information and knowledge, and depending on the level of engagement of the individual, can be conducive and translated into considerable improvements for performance and learning (Chung et al., 2009, 2010). Borgatti and Cross (2003) defined a formal model to explain social behavior during the learning process and the importance of relations to facilitate the search of information. In distributed settings, the density of interactions within social networks of learners facilitate rich information exchange and allows for developing a sense of belonging of the learners in a distributed setting (Haythornthwaite, 2002a). In the case of Yoon (2011), social network graphs (or sociograms) of student interactions were used in learning settings for students to make informed decisions about their choice of interactions with others, which in turn assisted in their understanding of complex socioscientific issues. Similarly, Tomsic and Suthers (2006) investigated the social network structure of Hawaiian police officers and how the introduction of an online discussion tool affected their learning their booking operation and increased collaboration across districts. They concluded that formation of new collaborative ties is more significant for learning rather than raw frequency of interaction. In a more recent study, Hommes et al. (2012) studied the influence of social networks, motivation, social integration and prior performance on the learning of 301 medical students. While they claimed to be the first to use social networks to study its effect on student learning, a number of key authors (Haythornthwaite, 2002a; Siemens et al., 2011; Paredes et al., 2012) in the domain of learning analytics researching in this area has been omitted. Furthermore, Hommes et al. (2012) utilised degree centrality only in their study to represent the social network construct and found it to be the key predictor for student learning. Thus, it is interesting to examine theories of network properties that explain how information is disseminated through ties within networks, and how network structures are conducive to the learning process and performance.

*Strength of weak ties theory*

One of the most seminal works related to information diffusion and learning is documented by Granovetter (1973). He reasoned that as individuals connect with closely-knit groups bound together by strong ties, information diffuses rapidly and becomes redundant. So although dense networks indicate high degree of communication, new
information come from weak ties, which act as bridges and connects one to new groups of people. In the context of learning, these weak ties are instrumental for new learning.

Krackhardt and Stern (1988) developed a measure to capture the extent to which ties are represented internally or externally within the groups. This is also referred to as the E-I index, which ranges from -1, meaning that ties are completely internal or intra-unit based, to 1, meaning that ties are completely external or inter-unit based. For instance, groups could be based on certain individual-based attributes such as group affiliation (e.g., office site location, or coursework or research student group, etc.) or roles (e.g., marketing, development, etc.). In organizational settings, groups that have ties that are externally based (i.e., high E-I indices) are more likely to implement large-scale organizational change and collaborate better (Nelson, 1989; McGrath et al., 2003), whereas groups with ties that are mostly internally based (i.e., low E-I indices) demonstrate higher resistance to external pressure (McGrath et al., 2003). It can thus be argued that when it comes to learning and performance, individuals who have E-I indices are those who demonstrate higher or richer levels of information exchange; such individuals will thus be those who are most likely to perform better.

**Structural holes theory**

Digressing from ties and groups, there is also a great deal of literature documenting the positive association between network position and performance. Burt postulates that one can optimize one’s network by efficiently maintaining one’s ties in non-redundant contacts such that one is effectively connected to a diverse variety of groups of closely connected contacts (e.g., a clique), where the groups themselves are not connected. The bridging of such holes within the network structure is termed structural holes, which yields information and control benefits. Thus, two network design principles in play here are efficiency and effectiveness. Efficiency entails maximizing the number of non-redundant contacts to maximize the gains through structural holes. Effectiveness means the preservation of primary contacts in the network considering also the contact’s diversity principle. As documented in the section on Measures below, the measure of efficiency incorporates effectiveness as well.

**Research model and hypotheses**

While most social network studies in the context of learning have generally been associated with individual and group performance or medium of communication, very few have incorporated learning as an important construct in
their model (Haythornthwaite, 2001; Yang et al., 2003; Chiocchio, 2007). Furthermore, given the unprecedented advancement and adoption of social technologies, this study provides much needed evidence in the learning and knowledge analytics domain to help understand how networks interact with technology to foster learning and performance in an information technology-dominant era (Bennett et al., 2010). It is anticipated that the theoretical model depicted below based on network and social learning theories would address these research gaps. Therefore, based on the discussion above, the following hypotheses are proposed, as depicted diagrammatically in Figure 1:

**H1:** Network structure: Density of an individual’s network is negatively associated with learning

**H2:** Network position: Efficiency of an individual’s network is positively associated with learning

**H3:** Network relation:

**H3a:** The extent to which an individual engages in communication within the network (intra-network communication) is positively associated with learning

**H3b:** The extent to which an individual contributes internally and externally to his group (inter-group communication) is positively associated with learning

**H3c:** Weak ties within an individual’s network is positively associated with learning

**H4:** Learning is positively associated with performance

### Context and methodology

#### The e-learning environment

The domain for this study is an e-learning environment where an online project management course was delivered for the Project Management Graduate Program in a leading “Group of Eight (Go8)” university in Australia. Being a postgraduate program that is similar to an MBA, the course was undertaken by 36 full-time working industry professionals ranging from healthcare to banking, information technology and engineering. Students were based locally, nationally and globally. The online mode of study required a high level of engagement with the course activities and peers through the university’s e-Learning website. In addition, the e-learning platform provides a channel for synchronous communication via chat, and asynchronous communication via group discussion boards. The discussion board is further classified into varying forums including “Public” and “Group” forums (which are private to the group). Within each forum, students can post and reply to messages. Assessments included an individual assignment (15%), an online quiz (10%), a group assignment (25%), and a final exam (50%). For the purpose of this study, group assignment marks were left out of the analysis as the primary focus is on individual learning and performance. Group interactions, however, were considered given that group communications added further insight into an individual’s level of engagement. Groups ranged from two to four members, with 12 groups in total. As the entire learning took place online, analyzing communication structures of such ‘virtual’ groups is another boon given that coordinating and collaborating with team members located in different time zones is one of the challenges of virtual collaboration and one of the capabilities expected to be developed by individuals as part of learning.

#### Data collection, storage and extraction

The main collaborative tool provided for group coordination and collaboration was the two discussion forums. The public forum focused upon solving general questions about the course, assignments and so on. The private forums, or group forums were created exclusively for internal group coordination and collaboration. Every student had the chance to post messages to the lecturer, other students using the public forum or directly to a group member using the private forums. In total there were 845 messages in the public forum, and a total of 722 across the private forum. Given that the purpose of this study is to model learning and performance of students only, interactions to and from the lecturer were discarded.

The data collection, storage and extraction comprised of the following steps and can be depicted in Figure 2 below:

1. Database development
2. Development of a Java program for network data extraction and measures calculation
3. Visualization of sociograms using social network analysis tools
4. Message content classification
5. Collation of network and attribute data for statistical analysis
Figure 2. Processes involved in data collection, storage and extraction

**Database development**

The information of the messages was extracted directly from the online discussion forums. Due to the unstructured nature of the message logs, a preliminary process of data preparation was executed. As a first step, a database schema (Figure 3) was designed so that the data could be loaded into a MySQL database. This provided us with the required flexibility to extract a variety of relational and attribute data about the levels of interaction.

Figure 3. Database schema for storage of social network (relational) data

**Development of a Java program for network data extraction and measures calculation**

Once the database was loaded, a small Java program was written to extract the information required to generate the node data and tie data required. This application provides some important functionality for this study that allows us to:

- Generate the message frequency matrix
- Calculate the contribution index (Gloor et al., 2003)
- Calculate the External-Internal Index (Krackhardt et al., 1988)
- Calculate the Content Richness Score
- Calculate Average Tie Strength
- Generate Node and Tie data
Visualization of sociograms using social network analysis tools

In addition, the Java program (Figure 4) facilitates the data extraction for the visualization of communication networks at various points in time (total of 6 periods) as shown in Figure 6.

![Java program for extracting relational data and network-level measures](image)

Figure 4. Java program for extracting relational data and network-level measures

After extraction, visualization analysis was done using UCINet and Netdraw to generate the network sociograms and statistics (Borgatti et al., 2002). The sociograms below show progressive visualizations of individuals and groups interactions and its evolution over time (classified into 6 periods) shown in Figure 5 & 6.

### Table 1. Number of new and total messages over 6 cumulative periods

<table>
<thead>
<tr>
<th>Period</th>
<th>From</th>
<th>To</th>
<th>Cumulative no. of new messages</th>
<th>Cumulative total no. of messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27/07/2009</td>
<td>16/08/2009</td>
<td>97</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>27/07/2009</td>
<td>06/09/2009</td>
<td>88</td>
<td>185</td>
</tr>
<tr>
<td>4</td>
<td>27/07/2009</td>
<td>18/10/2009</td>
<td>206</td>
<td>546</td>
</tr>
<tr>
<td>5</td>
<td>27/07/2009</td>
<td>08/11/2009</td>
<td>288</td>
<td>834</td>
</tr>
</tbody>
</table>

![Legend for groups (colour of node) and degree (shape of nodes)](image)

Figure 5. Legend for groups (colour of node) and degree (shape of nodes)
The UCINet program also calculated network-level measures such as density and structural holes (efficiency). Computation of other network-level measures (discussed in the following section) such as density, average tie strength, contribution index, external-internal index and content richness index were provided by the customised Java program.

Message content classification

As discussed before, given the lack of evidence of measures of social learning using a structural perspective, it is arguable that the main source of information about social learning in virtual teams resides in the richness of the dialogues between team members. A meaningful exchange of dialogue among team members and classmates is instrumental to enhancing their learning process. Thus, to develop a surrogate measure of social learning, for each message in the forums, we analyzed its content in order to identify behavioral patterns and sought to classify such patterns. The classification was based on past research that defined several methods to categorize different types of communication based on specific message features such as length (Licoppe et al., 2005), channel of dissemination (Peters, 1999; Licoppe et al., 2005), content (Pérez-Alcázar et al., 2003), and meaning (Gilbert et al., 2005), among others.
In this study, we have defined a classification method based on message content and meaning in order to categorize each message sent in the public and private forums. We term such categories “content richness.” The five categories of content richness, with non-exhaustive descriptions, as listed in Table 2, are:

- **Empty Message**: Inexistent content, file exchange without dialogue, greeting messages.
- **Team Building Message**: Team member personal introductions and very basic coordination. Final group closing activities, congratulations for group achievements and recognition for mutual cooperation. Team building messages support the creation of a sense of common goal and a shared set of beliefs and values.
- **Dissemination Message**: Information about group submissions and notifications about new document versions. Information dissemination is fundamental to keep all team members on track and aware of the status of project activities.
- **Coordination Message**: Team meeting coordination, a very important factor considering the time zone difference issue that some groups have to face.
- **Collaboration Message**: Messages that add value to the group work in terms of knowledge creation. Problem-solving dialogues, different insights about group work issues and project activities.

Thus, in terms of content richness, we consider the last category the most important; however, in terms of social learning all the other categories are also important as indicators of evidence of social learning. Every category was assigned a different weight that represents the content richness of the messages classified in each category (see Table 1). This parameter is considered later to calculate the content richness scores for an individual as explained further below.

<table>
<thead>
<tr>
<th>Weight</th>
<th>Content category</th>
<th>Message example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Empty</td>
<td>“Alright, see you later!” “Bye,” “Thanks.”</td>
</tr>
<tr>
<td>1</td>
<td>Team Building</td>
<td>“Excellent work, team,” “The last task has really got me enjoying this group work.”</td>
</tr>
<tr>
<td>2</td>
<td>Dissemination</td>
<td>“I submitted the last version of our report,” “The deadline has been extended.”</td>
</tr>
<tr>
<td>3</td>
<td>Coordination</td>
<td>“Let’s meet tomorrow at 7pm,” “I can write this section of the report. John, can you do the other part and Emily integrate it all?”</td>
</tr>
<tr>
<td>4</td>
<td>Collaboration</td>
<td>“Dear Peter, I think your answer to the question is correct. However, I found this article in which the authors analyse the issues from the different perspective. Please consider also…”</td>
</tr>
</tbody>
</table>

Collation of network and attribute data for statistical analysis

Finally, after all network measures were computed for each individual, the measures were stored as an attribute data in a spreadsheet format for further statistical analysis in SPSS. This master spreadsheet contained all network and attribute-level data for each individual.

Measures

We adopted an egocentric approach to collect data from the e-learning course social network. *Ego networks* are formed by a focal node called “ego,” the nodes to which that ego node is directly connected (alters), and the connections or *ties* among alters (Scott, 2000).

**Density**: Density $D$ is described as a measure of network cohesiveness and is defined as the relation of the existing number of ties to the maximum number of ties possible in a directed graph of $n$ nodes:

$$D = \frac{\sum_{i,j=1}^{n} x_{ij}}{n(n-1)}$$

where $x_{ij}$ is the value of the connection from $i$ to $j$. 

247
Efficiency: Effective size is a measure of the number of alters minus the average degree of the alters within the ego network (Burt, 1992). The effective size of an ego’s network has been defined as:

\[ 1 - \sum_{q \neq i, j} p_{iq} m_{jq} \]

where \( i \) is the ego, \( j \) is a primary contact, and \( q \) is also a primary contact who has strong ties with the ego \( i \) (represented by \( p_{iq} \)) and \( j \) (represented by \( m_{jq} \)). Efficiency is measured then by dividing the effective size by the number of alters in the ego’s network.

Contribution Index: Gloor et al. (2003) introduces the notion of Contribution Index (CI) to measure the level of participation in social learning settings. The contribution index is defined as:

\[ CI_i = \frac{\sum s_i - \sum r_i}{\sum s_i + \sum r_i} \]

where \( i \) is the contributor (ego), \( \sum s_i \) is the total of messages sent by \( i \) and \( \sum r_i \) is the total of messages received by \( i \). The contribution index value can be in a range from -1 to +1. If the learner mostly received messages, then his or her contribution index will be close to -1. On the other hand, if the learner mostly sent messages the contribution index will be close to +1. In terms of social learning we are looking for highly interactive dialogues. A contribution index near to 0 is indicative of a balanced dialogue of the learner with his or her team colleagues.

External-Internal Index: Krackhardt and Stern (1988) described a measure of interaction for both intra and inter work teams. External-Internal (E-I) index is a measure that compares the number and average strength of external ties to internal ties within different sub-groups in a network. The E-I index has been defined as follows:

\[ EI_q = \frac{\sum e_q - \sum i_q}{\sum e_q + \sum i_q} \]

where \( q \) is the ego, \( \sum e_q \) represents the total number of external messages sent by \( q \), and \( \sum i_q \) represents the total number of messages sent within the team. Similarly to the contribution index presented before, the E-I index ranges from -1 (only internal group communication) to +1 (indicating interaction only with external contacts rather than with members of the same team).

Content Richness Score: Each message of the dataset was classified according to the level of richness of its content. Each class was assigned with a weight that reflects the meaningfulness of the message in the dialogue. This is a novel contribution of the study - the Content Richness score. Content Richness (CR) is a measure of learning engagement in a dialogic context where the meaningful information exchange among team members drives the individual and group learning process and is thus evidence of learning. The CR score is defined as:

\[ CR_q = \frac{\sum_{i=1}^{n} mc_i}{n \times \max(mc)} \]

where \( q \) represents the ego, \( mc_i \) represents the message content value of the message \( i \), \( n \) is the total number of messages sent by \( q \), and \( \max(mc) \) is the maximum possible value of message content quality. The CR score range from 0 to 1. If a learner’s CR score is 0 that means that his or her level of contribution richness to the discussions was non-existent. On the other hand, a CR score of 1 means a highly meaningful participation and engagement in course and group discussions. The importance of the CR score for this study is that we hypothesized that the
learner’s performance is directly related to his or her level of engagement in a social learning environment, and this metric can help us to determine if there is such relation between these two indicators.

**Average Tie Strength:** In keeping with social network literature (Marsden et al., 1984), we used frequency of contact to calculate average tie strength. Therefore, for the purpose of this study, a learner’s tie strength has been measured as the average of all his or her tie strength (frequency of contact) to all other actors in the network.

**Performance Measures:** As stated previously, during the semester the students were assessed on three individual components - individual assignment (15%), online quiz (10%), and final exam (50%).

**Results**

As per Table 3, the mean of the CI is well balanced (CI = .015), and tends towards zero, which indicates an equally distributed rate of messages sent and received. In terms of E-I index, the mean value (E-I = -.872) indicates that there was more internal communication rather than dialogues external to the groups, which was expected given the large number of messages of internal group communication in the dataset. The mean CR score is quite high (CR = .667) which is indicative of meaningful content exchange within the groups.

<table>
<thead>
<tr>
<th>Table 3. Descriptive statistics of measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Engagement</td>
</tr>
<tr>
<td>Contribution Index (CI)</td>
</tr>
<tr>
<td>External-Internal Index (E-I)</td>
</tr>
<tr>
<td>Content Richness Score (CR)</td>
</tr>
<tr>
<td>Network Position</td>
</tr>
<tr>
<td>Efficiency</td>
</tr>
<tr>
<td>Network Structure</td>
</tr>
<tr>
<td>Density</td>
</tr>
<tr>
<td>Relations (Ties)</td>
</tr>
<tr>
<td>Average Tie Strength</td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>-.100</td>
</tr>
<tr>
<td>-.100</td>
</tr>
<tr>
<td>.000</td>
</tr>
<tr>
<td>.333</td>
</tr>
<tr>
<td>.000</td>
</tr>
<tr>
<td>.000</td>
</tr>
</tbody>
</table>

Pearson’s correlation analysis was used to test the hypotheses (see Table 4 for a detailed description of the correlations as at end of period 6 (i.e., overall semester)).

<table>
<thead>
<tr>
<th>Table 4. Correlation matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual assignment</td>
</tr>
<tr>
<td>Quiz</td>
</tr>
<tr>
<td>Exam</td>
</tr>
<tr>
<td>CI</td>
</tr>
<tr>
<td>E-I</td>
</tr>
<tr>
<td>CR</td>
</tr>
<tr>
<td>Efficiency</td>
</tr>
<tr>
<td>Tie strength</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual assignment</th>
<th>Quiz</th>
<th>Exam</th>
<th>CI</th>
<th>E-I</th>
<th>CR</th>
<th>Efficiency</th>
<th>Tie strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiz</td>
<td>.281*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exam</td>
<td>.379*</td>
<td>.885**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td>-.550**</td>
<td>.045</td>
<td>-.109</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-I</td>
<td>.035</td>
<td>.083</td>
<td>.171</td>
<td>-.464**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR</td>
<td>.311*</td>
<td>.341*</td>
<td>.253</td>
<td>.344*</td>
<td>-.354*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency</td>
<td>-.297*</td>
<td>.049</td>
<td>-.095</td>
<td>.304*</td>
<td>-.138</td>
<td>.394**</td>
<td></td>
</tr>
<tr>
<td>Density</td>
<td>.227</td>
<td>-.101</td>
<td>.069</td>
<td>-.318*</td>
<td>.173</td>
<td>-.406**</td>
<td>-.90**</td>
</tr>
<tr>
<td>Tie strength</td>
<td>.116</td>
<td>.44</td>
<td>.131</td>
<td>.72</td>
<td>-.271</td>
<td>.422**</td>
<td>.066</td>
</tr>
</tbody>
</table>

**Note:** *Correlation is significant at the 0.05 level (1-tailed). **Correlation is significant at the 0.01 level (1-tailed).

In summary, the significant results are summarized below:

- **Network Structure:** There is a significant negative relationship between network density and CR score, $r = -.406$, $p$ (one-tailed) $< 0.01$. In terms of social learning, this result makes sense because it is better to have a few meaningful dialogues, rather than many meaningless conversations. Given this result, we find support for $H1$. 

249
• **Network Position:** Network efficiency is significantly correlated with CR score, \( r = .394, p \) (one-tailed) < 0.01. Highly efficient learners are connected with contacts that are expected to provide good quality content (high CR), so they can fulfill the informational needs of the learner without having to look for other sources (Burt, 1992). This result provides us with evidence to support \( H2 \). Thus, when a learner has many sources of information, the cost of maintaining those connections increases as well as the level of information redundancy in the network. Consequently, his or her level of network efficiency decreases. Even though theoretically, a decrement in efficiency should affect learner’s performance, the results obtained do not allow us to conclude such phenomena (i.e., no statistically significant relationship between any of the performance variables and efficiency).

• **Engagement:** A significant positive correlation exists between CR and CI \( (r = .344, p < 0.05 \) (one-tailed)), thus lending support to \( H3a \). According to the Gloor’s definition of CI (Gloor et al., 2003), an optimal contributor would present a balanced rate of messages sent compared to the number of messages received. Therefore, the CI value should tend to be zero for optimal communication. The conclusion that we can make from this result is that learners with higher CR score send more messages than they receive, and as a consequence, there is no reciprocity in terms of meaningful content exchange for social learning.

• In addition, the CR score is significantly negatively correlated to E-I index, \( r = -.354, p < 0.05 \) (one-tailed), thus allowing us to reject \( H3b \). In the case of E-I index we are also looking for a balanced rate of internal and external communication (Krackhardt et al., 1988). A high E-I index indicates a relatively higher communication by an individual to those outside his group relative to those internal to his group. This is beneficial for avoiding “group think.” Accordingly, those who communicate more frequently internally within groups relative to externally outside groups are also engaged in higher or richer levels of communication. This can be attributed to the fact that the large number of internal group messages in comparison to the external ones influences the E-I index. In fact, there were more interactions between group members, rather than among external contacts, which indicates that learning as evidenced by content richness took place within groups rather than outside of groups. This result is very likely due to the large number of meaningful internal dialogues.

• **Network Ties:** There is a significant positive relationship between the average strength of ties and CR score, \( r = .422, p < 0.01 \) (one-tailed). Therefore, there is sufficient evidence to reject \( H3c \). The stronger the tie the more frequently the contacts occur. This implication means that contacts with high level of interaction tend to mutually exchange valuable information. These dialogues are rich in content and provide more in-depth insights about the topics of learning. Although Granovetter’s theory may not hold true in this circumstance, other researchers have claimed that strong ties are symbol of closeness and trust, which are two determinant components for social learning (Lave et al., 1991; Krackhardt, 1992).

• **Performance:** There is a significant positive relationship between CR score and the individual assignment marks, \( r = .311 \), as well as between CR score and the quiz mark, \( r = .341 \), both \( p < 0.05 \) (one-tailed). However, the exam result, which has the highest assessment weight in terms of learning outcome, does not seem to be significantly associated with CR score and for none of the engagement measures proposed in this study. Taken altogether, we consider that these results are somewhat indicative enough of how meaningful dialogic exchange among contacts can enhance learners’ performance. Therefore, we find partial support for \( H4 \).

The following table (Table 5) summarizes the results of the hypothesis testing:

<table>
<thead>
<tr>
<th>Hypothesis number</th>
<th>Brief description</th>
<th>Support/No support</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Network structure &amp; Learning</td>
<td>Support</td>
</tr>
<tr>
<td>H2</td>
<td>Network position &amp; Learning</td>
<td>Support</td>
</tr>
<tr>
<td>H3a</td>
<td>CI index &amp; Learning</td>
<td>Support</td>
</tr>
<tr>
<td>H3b</td>
<td>E-I Index &amp; Learning</td>
<td>No support</td>
</tr>
<tr>
<td>H3c</td>
<td>Network relations &amp; Learning</td>
<td>No support</td>
</tr>
<tr>
<td>H4</td>
<td>Performance and Learning</td>
<td>Partial support</td>
</tr>
</tbody>
</table>
Discussion

According to the results obtained after the correlation analysis we can argue that rather than performance, social learning is influenced by social network properties such as structure, relations and network position. In summary,

- denser networks are not conducive to learning;
- those highly efficient or who hold key brokerage positions are able to learn better than those in less efficient network positions.
- strong ties matter more so than weak ties as far as learning is concerned. This can be attributed to the complexity associated with learning and such findings are consistent with literature (Krackhardt, 1992).
- In terms of engagement, those with higher contribution indices demonstrate higher levels of learning as measured by content richness. Furthermore, those who communicate more frequently within their groups than to those outside the group demonstrated higher levels of learning.

The findings of this study provide further evidence of the importance of social networks in affecting learning. The results are aligned with the arguments presented by Lave and Wenger (Lave et al., 1991) in their situated learning theory, and with the principles that provide the basis of Siemens’ theory on connectivism (Siemens, 2004).

Content Richness (CR) was shown to be a good predictor of social learning due to the interesting findings that connect the measure with most social network characteristics modelled here. It makes sense that an individual with a less dense but more efficient network shows higher levels of engagement when interacting with his or her contacts (Burt, 1992). The exchange of meaningful information involves trust and benevolence amongst actors in the network and this is directly related to the strength of ties.

In terms of limitations, first of all, the size of the dataset raises a concern about the level of representation and generalization of the results. However, given the central limit theorem’s rule of thumb of a minimum of 30 samples and the fact that this study is an exploratory one, the results are indicative of the power of social networks in influencing learning and indirectly, performance, although not generalizable to the entire population. Secondly, the classification method for classifying content richness is in preliminary stages of maturity and questions may arise in terms of reliability and robustness. However, the development of a taxonomy or vocabulary in studies of linguistics and semantic data mining could allow the automatic identification of significant keywords in a message, which could be automatically classified in a predetermined set of categories. Content mining and information retrieval techniques can provide important benefits in this area of analysis.

Another limitation of this study is that most of the dialogues analyzed took place within groups, and as a consequence there is no clear evidence of interaction beyond groups, including those with the teacher (note that the purpose of this study was to model social learning amongst students only). It would be interesting to study the relation between the meanings of internal and external dialogues. However, when group formation is compulsory, according to our results we can state that the level of quality of the dialogues is significant. One important factor that should not be discarded is that group discussion boards are just one channel through which learners interacted during the semester. The messages in the discussion forum may have been followed up through use of other types of communication channels such as social media, video conferences, emails, face-to-face meetings and phone calls. These interactions, although nice to have for the study, have not been accounted for. In future, a holistic approach that accounts for such interactions, including interactions with teachers and even faculties, can be considered for further research.

Conclusion

Although many theories advocate the importance of social interaction in influencing learning and performance, empirical evidence is relatively few. In this study, we presented the development of a theoretical model for understanding the impact of social networks in learning and performance. The novelty of this research is driven by the construction of content-based measure called “content richness” which provides a new approach for measuring the level of engagement of learners in a social learning environment by exchanging meaningful information. We analyzed the communicational patterns of students and groups located in different cities, countries, and time zones. They were part of an online course that relied heavily on using the discussion forums as an important communication
tool. The results show that rather than performance, social learning is highly influenced by the learners’ network of contacts.

The implications of this study can be understood from a theoretical, methodological and practical standpoint. Although exploratory, this study challenges the theoretical notion of how social networks directly impact performance and questions the existence of antecedents and other confounding variables in the networks-performance equation. Methodologically, as described above, this study offers a novel approach to quantify social learning. Practically, the results are significant for educational and curriculum designers to facilitate the development of informal networks for improved learning. For instance, the number of opportunities to collaborate on a frequent basis over a semester may be considered as opposed to changing groups for varying assessments. For future research, it would be useful to run regression models with a larger dataset to test whether networks and performance is significantly mediated by social learning. In an age of MOOCs and learning analytics, such a model would allow educators, professional development leaders and academics to enhance learning and make informed decisions about designing learning strategies for an Internet-enabled society of the 21st century.

References


Enhancing Vocabulary Retention by Embedding L2 Target Words in L1 Stories: An Experiment with Chinese Adult e-Learners

Zi-Gang Ge
School of Network Education, Beijing University of Posts and Telecommunications, China // shouzhou11@126.com

(Submitted June 26, 2014; Revised November 4, 2014; Accepted January 29, 2015)

ABSTRACT
This study aims to investigate the effectiveness of a storytelling approach in Chinese adult e-learners’ vocabulary learning. Two classes of 60 students participated in the experiment, with 30 in the experimental group and 30 in the control group. The storytelling approach and the rote memorization approach were employed by the two groups respectively in one teaching session of 30 minutes. Two post-tests were administered to the two groups, with one completed immediately after the teaching and the other assigned three weeks later. A questionnaire survey was also assigned to the experimental group to obtain the learners’ perceptions of the storytelling approach. The results of the data analysis showed that the storytelling method was more effective than rote memorization in both short-term retention and long-term retention, though its effects tended to diminish a little bit with time. Possible suggestions to perfect the method were given in the paper.

Keywords
Vocabulary retention, Storytelling, Rote-memorization, E-learners

Introduction
Vocabulary items are often deemed as “building blocks” of a language (Amirian & Heshmatifar, 2013). Without the knowledge of a certain amount of vocabulary in the target language, learners cannot effectively learn to listen, speak, read, or write in the language (Nation, 2001). Fukkink et al. (2005) echoed this view by claiming that understanding will be greatly hampered without knowing the words’ meanings in a text.

Various methods have been suggested or tried for vocabulary acquisition. The most common way is to use a dictionary. Many scholars have examined the use of various dictionaries in vocabulary acquisition (Amirian & Heshmatifar, 2013; Chiu & Liu, 2013). In these studies, printed, digital and online dictionaries have been examined. Dictionaries, especially electronic and online dictionaries have been proved to enhance vocabulary acquisition and increase vocabulary retention span. But the dispute over dictionary use in vocabulary learning is always an unresolved issue, as some scholars (Knight, 1994; Schmitt, 2000; Erten & Tekin, 2007) claimed that learners needed to use more contextual or semantic clues to infer words’ meanings but not frequently look up the words in dictionaries.

Other than dictionaries, some scholars (Liao & Chen, 2012) examined the possibility of using lexical games or storytelling to facilitate vocabulary acquisition. In their study, Liao and Chen found that the games of Poker and Chinese chess could boost learners’ lexical growth and retention. Prince’s study (2012) tried to use the method of embedding target words in a story to improve learners’ vocabulary retention efficiency, and the results concluded that the story approach was effective in eliciting short-term memory of vocabulary.

Another vocabulary learning approach frequently discussed by scholars is the application of various vocabulary exercises. Lin and Hsu (2013) in their study compared the effectiveness of hierarchy vocabulary exercises and copying vocabulary exercises, and found that students doing the hierarchy vocabulary exercises outperformed those receiving copying exercise in both vocabulary gains and long-term retention. Another study (Hashemzadeh, 2012) proved that learners could recall more words by doing the normal type of blank-filling exercise.

No matter what methods are employed, they inevitably deal with the two stages in the process of vocabulary learning. The first stage is to get to know the word, or to memorize the word, and the second stage is to retain the word in memory as long as possible. Compared with the first stage, the second stage often presents a more difficult problem to language learners, as it is a very common phenomenon for learners to forget a word’s meaning over a certain time span. As there is no clear-cut distinction between the two stages, the above mentioned methods are also explored in many studies as to their effects on vocabulary retention over a longer period of time. But till now few
studies about vocabulary acquisition and retention have been done in an e-learning environment, which is the major theme of the present study. Another feature of the present study is that it tries to use an L1 (Chinese in this case) context to facilitate L2 (English in this case) vocabulary learning, which is different from many of the previous studies that tried to teach L2 vocabulary in an L2 context.

Review of literature

Vocabulary retention span

As mentioned in the previous part, the first stage in vocabulary learning is to memorize the target words. But the problem is that these words might not be correctly recalled over a long period of time. This problem obviously has to do with our memory (Gu, 2003). Memory, though a commonly used word in daily life, is more frequently explored in the psychological world. Parle et al. (2006) defined memory as the process of encoding, storing, and retrieving information.

From an information processing perspective, three types of memory can be identified: sensory memory, short-term memory and long-term memory (Carlson, 2010). Sensory memory is to tackle the sensory information which is perceived by sensory receptors (such as eyes and ears). An example of sensory memory is that after we have been exposed to an item, we probably will remember it at other times when we see it again. According to some scholars, this type of memory is very limited in capacity and cannot be prolonged via rehearsal (Sperling, 1963). Short-term memory deals with the information perceived in an active and readily available state. Like sensory memory, the capacity of short-term memory is also very limited, though the capacity can be increased by using the technique of chunking (Miller, 1956). Unlike the former two types, long-term memory is both larger in capacity and longer in duration. It is said that information held in long-term memory can remain indefinitely (Atkinson & Shiffrin, 1968).

From the above analysis, it is clear that information held in long-term memory is most likely to be retrieved over a long period of time, which is also a desired outcome in vocabulary learning. In other words, the ultimate goal of vocabulary learning is to retain the target words in our long-term memory, so that we can easily retrieve them when we want to use them.

Various ways have been examined to prolong learners’ vocabulary retention span. Laufer and Osimo (1991) investigated the use of “second-hand cloze” (which in the researchers’ words means “filling in the target items in a summary version of the original texts”) in vocabulary learning, and they found that this technique could improve long-term retention of words. They inferred that the effectiveness of this technique lied in the fact that it not only embodied some features of other memorization techniques, but also overcame their shortcomings. Rodriguez and Sadowki’s study (2000) compared four methods in vocabulary learning, namely rote rehearsal, context, keyword, and context/keyword. Rote rehearsal is more commonly referred to as rote memorization, which means repeating the target words until they can be recognized. The context method is to infer the meanings of the target words from the general situation that relates to them. The keyword method is to infer the target words’ meanings from certain words related to the target words. Their findings indicated that the context/keyword combined method would produce better learning outcome in long-term vocabulary retention. Parle et al. (2006) proposed that regular rehearsals were useful in transferring vocabulary knowledge from short-term memory to long-term memory. Gorjian et al. (2011), through their study, found that asynchronous computer-assisted language learning approaches would exert a significantly better impact on vocabulary retention over a longer time span, and they believed that these CALL approaches benefited both high and low achievers in long-term vocabulary retention.

In the study done by Nemati (2009), the author examined the impact of some memory vocabulary mnemonics on long-term retention and found that the memory strategies were more effective in inducing long-term retention. In the author’s mind, memory strategies belonged to the highest level of information processing and thus would produce long-term vocabulary retention. In her mind, being simply exposed to words without deeper levels of information processing could not enhance long-term vocabulary retention. Ghorbani and Riabi (2011) did a similar study about the impact of memory strategies on EFL learners’ vocabulary retention, and the results also showed that memory strategies could greatly enhance long-term vocabulary retention. In her follow-up study, Nemati (2013) investigated the impact of teaching vocabulary learning strategies on short-term and long-term retention, and found strategies
such as “grouping” (words with meanings related to the same topic) could significantly enhance both short-term and long-term retention of lexical items.

**Storytelling and vocabulary learning**

Among the various ways employed to boost vocabulary learning, storytelling is one of the oldest, as it is an ancient verbal activity used for information sharing, entertaining and many other purposes (Soleimani & Akbari, 2013). Gere et al. (2002) defined storytelling as the human activity of using oral or body languages to exchange information. McDrury and Alterio (2003) also hold a similar viewpoint about storytelling, and in their mind storytelling is a very important way for human beings to perceive the world.

Storytelling has been used in language teaching and learning for many years; for example, it is widely used as an effective approach to teach L1 and L2 languages to pre-school children across the world, which can be exemplified by the popularity of stories like Cinderella and Little Red Riding Hood.

When we turn our attention to vocabulary learning, we may find storytelling is also frequently used. Many studies have supported the claim that incidental vocabulary learning can be achieved by reading or listening to stories (Elley, 1991; Shu et al., 1995). Kirsch’s study (2012) focused on the use of storytelling to teach German lessons to primary students, and the results showed that the stories helped the learners to recall a considerable number of words. Prince’s study (2012) explored the effect of storytelling by embedding the target words in L2 sentences which revolved around certain topics. His study indicated that embedding target words in a story could enable learners to have better short-term vocabulary retention than when sentences are unrelated. In an attempt to see whether vocabulary acquisition from teacher explanation and repeated listening to stories could overcome the Matthew effect, Penno et al. (2002) found that children in the early years of school could learn vocabulary from listening to stories and that the learning outcome could be further enhanced by repeated story presentations and the supplementary explanation of difficult words from the teachers, though the methods did not overcome the Matthew effect.

As stories can be told by using different media, some research has been done to explore the different types of storytelling used in vocabulary learning. One study compared digital storytelling and paper story reading as to their impact on vocabulary retention rate, and the results showed that digital storytelling would increase learners’ vocabulary retention rate more than story reading only (Tütüniş & Şenel, 2013).

The usefulness of storytelling in vocabulary acquisition lies in the fact that stories can provide learners with a network of associations of the target words, or in other words, stories can provide contextual clues to language learners (Bowen & Marks, 1994; Penno et al., 2002). As Bowen and Marks (1994) noted, words would be often remembered when they appeared within the context of a story, because the new words were in rich networks of associations.

**The present study**

It is generally believed that L2 vocabulary is best learned when the learners can get much input of the target language (Krashen, 1989). And a common idea held by scholars is that the meaning of an L2 word can be best understood or even guessed out in an L2 context (Fischer, 1994; Nassaji, 2003; Huang & Eslami, 2013). But things are quite different in the Chinese e-learning environment. Most Chinese e-learners’ are of low-level English proficiency (Ge, 2011; Ge, 2012), so the context of too many L2 words might overwhelm them and discourage them. With this characteristic of learners in mind, the author of this paper tries to use the L1 context to help learners’ L2 vocabulary learning.

The present study aims to compare the impact of the storytelling approach and the rote memorization approach on adult e-learners’ vocabulary learning. The participants would take an immediate post-test and also a delayed post-test to examine their learning outcomes. A questionnaire survey would also be administered to those receiving the storytelling approach, with the aim to obtain their perceptions of the approach. The data obtained will be analyzed to address the following two research questions:
• Are there significant differences in the learning outcomes of the storytelling approach and the rote memorization approach?
• What attitudes did the participants have toward the storytelling approach?

Methods

Participants

60 participants were selected from an e-learning institution located in Beijing, China. They were all the first-year students majoring in computer sciences, ranging in age from 21 to 36 years old. They were randomly assigned to two groups, with 30 members in each group. Group A (the experimental group) was to receive the storytelling approach and Group B (the control group) would do rote memorizing in their learning. To make sure about the two groups' homogeneity at the beginning of the experiment, an online Oxford Placement Test (OPT) was administered (50 points in total). OPT, which is composed of 50 questions, is a standardized test usually used to test EFL learners’ English knowledge about grammar. The result of the test is shown in Table 1, and the data showed that the two groups had no significant differences in their existing English knowledge ($p = .962 > .05$). The mean scores of both groups were below 30 points, which indicated that these participants were at a very low English proficiency level.

Table 1. Comparison of the results on the Oxford Placement Test

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Std. deviation</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>30</td>
<td>12.17</td>
<td>4.00</td>
<td>25.00</td>
<td>5.038</td>
<td>0.048</td>
<td>58</td>
<td>.962</td>
</tr>
<tr>
<td>Group B</td>
<td>30</td>
<td>12.10</td>
<td>4.00</td>
<td>25.00</td>
<td>5.732</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Instruments

As the first-year students, these participants were obliged to learn the course of college English with the textbook of College English Book II published by Shanghai Foreign Language Education Press. So the target words assigned to learn in the experiment were chosen from the article of the first unit of this book. Ten words were chosen from the article. The number was considered appropriate for adult e-learners in this experiment, though some scholars pointed out that in class teaching the number of chunks of information should not exceed seven, or the short-term retention would not be effective (Moras, 2001). The reason why more than 7 items were chosen is that a larger amount of input might show more clearly the differences of the learning outcomes between the two groups.

For Group A students, the English words were embedded into a series of Chinese sentences (part of speech and the Chinese translations of the target words were given in brackets next to them), which were connected by a topic. However, the same target words were just listed along with their part of speech and Chinese translations for Group B students.

Procedure

Both classes attended a 30-minute online class session respectively through an SNS software called YY Messenger, which has the function of online synchronous teaching (as shown in Figure 1).

Prior to the experiment, a pre-test was administered to the two groups to examine the learners’ existing knowledge of the target words. A list of 15 words from Unit One of the book was presented through YY Messenger, and all of the subjects were required to report their Chinese translations of the words through emails in 10 minutes. Consequently, five words were discarded because they were familiar to some of the students in the two groups.

In the teaching session of Group A, the Chinese story containing all the target words was shown on the screen. The story is a coherent one and revolves around a certain topic (namely context) to make the ten target words related but not separated, as many studies approve that context will make retention easier (Bowen & Marks, 1994; Penno et al., 2002). The instructor would firstly read the story, and then all the students were given 10 minutes to read the story on
their own. In Group B, the procedure was almost the same as that of Group A, but the story was replaced with a list of the words along with their part of speech and Chinese meanings. And after the same instructor had read the target words, Group B participants were given 10 minutes to do a rote memorization of the words. Figure 2 shows the different forms of the teaching materials. The English version of the story used in Group A is given in Appendix A.

At the end of the teaching session, an immediate post-test was assigned to both groups. The post-test was in the same form of the pretest, but with only the 10 target words (10 points in total). The test time of the post-test was lengthened to 15 minutes, as the researcher presumed that the subjects would spend more time thinking over what they had just learned. Three weeks after the immediate post-test, at the beginning of another teaching session for each class a delayed post-test was administered through YY Messenger to both groups without notification in advance. The delayed post-test was completely the same as the immediate one. The two post-tests were used to assess subjects’ short-term and long-term retention of the target words respectively. After this teaching session, an online questionnaire survey was administered to Group A students, with the aim to obtain their perceptions of the storytelling approach.

Results

Data were collected from the two post-tests and the questionnaire survey, and all of them were shown and analyzed in the following part. Table 2 is an independent t-test of the immediate post-test.
Table 2 showed that the mean score of the immediate post-test in Group A was 7.5 and that of Group B was 4.9. The mean difference (7.5 - 4.9 = 2.6) was quite large compared with the total mark of 10. And the t-test showed that the difference was significant (p = 0 < .05) between the two groups, which implied that the storytelling approach had produced a much better vocabulary learning outcome than the traditional rote memorization approach. The small standard deviation in both groups indicated that the two different approaches had exerted a relatively even impact on subjects within each group.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>30</td>
<td>7.50</td>
<td>1.43</td>
<td>4.00</td>
<td>10.00</td>
<td>6.773</td>
<td>58</td>
<td>.000</td>
</tr>
<tr>
<td>Group B</td>
<td>30</td>
<td>4.90</td>
<td>1.53</td>
<td>3.00</td>
<td>10.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>6.20</td>
<td>1.97</td>
<td>3.00</td>
<td>10.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Another purpose of the study was to examine the long-term vocabulary retention by using the storytelling approach. The data of the delayed post-test was first processed with a descriptive analysis in SPSS, and then together with the immediate post-test was processed with an Analysis of Variance of Repeated Measures. Table 3-5 show the results.

Table 3. Descriptive analysis of the delayed post-test

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>30</td>
<td>5.800</td>
<td>1.186</td>
<td>4.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Group B</td>
<td>30</td>
<td>3.500</td>
<td>1.225</td>
<td>2.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>4.650</td>
<td>1.665</td>
<td>2.00</td>
<td>8.00</td>
</tr>
</tbody>
</table>

Compared with Table 2, Table 3 showed that both of the mean scores of the two groups had dropped in the delayed post-test. The maximum scores of the two groups also dropped. The decrease in these two sets of data indicated time lapse had some influence on the subjects’ vocabulary retention.

Since both of the two repeated factors (Group and Time) in this study have only two levels, the Mauchly’s test of Sphericity is definitely met (Field, 2008). The data shown in the following tables will show the impact of time and different teaching approaches on long-term vocabulary learning.

Table 4. Test of within-subjects effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>Sphericity Assumed</td>
<td>72.075</td>
<td>1</td>
<td>72.075</td>
<td>52.418</td>
</tr>
<tr>
<td>time * group</td>
<td>Sphericity Assumed</td>
<td>0.675</td>
<td>1</td>
<td>0.675</td>
<td>0.491</td>
</tr>
</tbody>
</table>

Table 4 showed that the two factors of Group and Time did not produce an interactive effect on the learning outcomes (p = .486 > .05), and there were significant differences in the immediate and delayed post-tests within each group. This may imply that whatever the approach is, a certain amount of vocabulary rehearsing is necessary for learners to obtain a better long-term retention rate.

Table 5. Test of between-subjects effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3531.675</td>
<td>1</td>
<td>3531.675</td>
<td>1543.029</td>
<td>.000</td>
</tr>
<tr>
<td>group</td>
<td>180.075</td>
<td>1</td>
<td>180.075</td>
<td>78.677</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>132.750</td>
<td>58</td>
<td>2.289</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5 showed the between-subjects effects of the factor “Group” (namely the two approaches). The p value (.00 < .05) indicated that the “Group” factor still produced significant differences in the vocabulary learning outcomes between the two groups three weeks later. Together with Table 3, we now can safely assert that the storytelling approach had a more facilitating impact on adult e-learners’ long-term retention of vocabulary.

In order to obtain subjects’ perceptions of the storytelling approach, a questionnaire survey based on a five-point Likert-scale (with one meaning Strongly Disagree and five meaning Strongly Agree) was administered to Group A students. Table 6 is the summary of the survey result.
Table 6. Questionnaire survey on subjects’ attitudes toward storytelling

<table>
<thead>
<tr>
<th>Survey questions</th>
<th>Q1: I like the story.</th>
<th>Q2: The topic of the story is familiar to me.</th>
<th>Q3: The story helps me remember the target words.</th>
<th>Q4: I used the story to help me complete the immediate post-test.</th>
<th>Q5: I used the story to help me complete the delayed post-test.</th>
<th>Q6: I will compose stories by myself in future study.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Mean</td>
<td>4.200</td>
<td>4.500</td>
<td>4.567</td>
<td>4.333</td>
<td>3.333</td>
<td>2.866</td>
</tr>
</tbody>
</table>

From Table 6, we can see that the storytelling method has been welcomed by the majority of Group A subjects, as is indicated by the mean score of Question 1. The success of the storytelling approach partly lies in the fact that the topic chosen was familiar to most students (indicated by the mean score of Question 2). The high rate of employing the story in the two post-tests showed that the intended measure had produced some lasting impact on the students’ vocabulary learning. But compared with that of Question 4, the mean score of Question 5 decreased, which indicated that as time went on, some problems might have arisen and consequently hindered the application of the story in vocabulary recall. By interviewing some Group A students, the researcher found that the reason might be that the story became unclear to some students after the three-week time span. As to Question 6, most students reported a negative attitude toward composing their own stories in future study. The researcher was told by some students that they thought they didn’t have the capability to make up stories as good as the one provided in the experiment.

Discussion

The results of the present study confirm the findings of some previous studies that storytelling is an effective method for vocabulary learning (Prince, 2012; Tütüniş & Şenel, 2013). There are two major differences between this study and the previous studies. The first is that the experiment reported in this paper was set in an e-learning environment, while the previous studies generally dealt with the traditional face-to-face teaching. The second difference is that this study tried to teach L2 vocabulary by embedding the target words in an L1 context, while many of other studies often employed L2 context as the measure.

It has been long preached that L2 can be best learned in L2 context, but the present study employed L1 as the medium for L2 vocabulary acquisition. So there comes the old issue of how mother tongue influences foreign language acquisition. In language acquisition, the phenomenon of language transfer is almost inevitable. Language transfer or linguistic interference means language learners applying knowledge from their previously acquired language to a new language (Odlin, 1989). Gass and Selinker (2001) categorized transfer into positive transfer and negative transfer. Positive transfer occurs when the relevant units or structures of the mother tongue and the target language are the same and linguistic interference can lead to correct language production. However, when language learners transfer units and structures that are not the same in the two languages, negative transfer will occur. So in language learning, negative transfer has to be avoided as much as possible to achieve a good learning result. On the lexical level, negative transfer often occurs. For example, many English words have the same meaning as that of Chinese words, but in fact many of these English words are used in different contexts. This phenomenon makes it easy for Chinese learners to make mistakes when using these English words. By embedding L2 words in L1 stories, the target words are clearly presented in a certain context, which is very familiar to learners, and thus learners will get to know the usage of the words more easily.

Another reason for using L1 context was deeply related to the English proficiency level of the subjects. As shown in the present study, most of the participants had a very low level of knowledge in English. If an L2 story context had been provided, it was very likely that many of the subjects would have been at a loss and discouraged when faced with a relatively long paragraph of English words no matter whether it was easy or not. As language context is often deemed as a very important medium for vocabulary learning (Webb, 2008), the unfamiliar L2 context may impede low-level learners’ learning process. To some extent, L1 context can be a much easier and more familiar tool for low-level L2 learners. So with the English proficiency level of the learners in mind, L1 stories will be more appropriate for these adult e-learners than L2 stories.

As mentioned in the section of literature review, stories can provide contextual clues for students, which are very useful for information memorization. Some people may argue that the context method and keyword method can also provide certain clues. But in most cases stories, especially dramatic stories, will elicit more interest and attention of
the learners than the normal kinds of context method or keyword method, and thus produce a better learning outcome.

The story used in this study was carefully constructed by the teacher and has exerted very good facilitating impact on students’ vocabulary retention. This finding contradicts the belief of some scholars that learning vocabulary from a specially constructed passage about a target word’s meaning is probably not as easy as from an authentic text (Beck et al., 1983; Herman et al., 1987). But data from the questionnaire survey of this study showed that composing stories was deemed as a tough task for most participants. One reason drawn from the interview with some of the students was that most of the low-achieving students did not have the capability to compose appropriate stories. In other words, the lack of ability or confidence will often curb low-achieving learners’ desire of applying the learning method actively. Students’ inertness in learning is very common among Chinese low-achieving e-learners, and often presents a headache to teachers.

The findings of the paper testify that storytelling will boost not only short-term vocabulary retention but also long-term retention. By comparing the results of the immediate and delayed post-tests, we may find that time really exerted a negative impact on vocabulary retention, no matter how good the learning strategy was. This finding echoes many other studies (Gorjian et al., 2011; Hashemzadeh, 2012; Tuan, 2012). This may imply that story rehearsal is also necessary to transform short-term retention into long-term retention. Story rehearsal may overcome the shortcomings of rote rehearsal but retain its effectiveness of transferring information from short-term memory to long-term memory (Dahlen & Caldwell-Harris, 2013). But on the other hand, compared with the more traditional method of rote rehearsal, story rehearsal is presumed to be much easier and more interesting, because stories will provide a network of associations of the target words, and hence make the process of rehearsing much simpler and more enjoyable. A possible method for substituting story rehearsal is to compose a series of stories which centers on a limited number of related topics. Let’s take the experiment in the present study as an example. All the vocabulary teaching of the textbook may revolve around the topic of the love affairs between Xiao Mei and Xiao Long. In this way the stories used in later vocabulary teaching sessions may naturally remind the learners of what they have gained in previous stories.

One more thing needs to be pointed out is that the delayed post-test was presented without early notification. Although some scholars (Peters, 2007; Keating, 2008) claimed that being informed of a vocabulary test did not necessarily mean a higher retention rate, it is very possible in the present study that if an early notification had been given, the students would have performed much better in the delayed post-test, since most Chinese students usually attach great importance to tests and thus are willing to do some kind of preparation before a test (vocabulary rehearsing in this case).

The follow-up experiment

To see whether the storytelling approach was effective for a different set of words, a follow-up experiment was done one month later after the first experiment to the same participants. The procedure was the same as that of the above-mentioned research, but this time the ten words were chosen from the article of Unit 3. The data are shown in Table 7-10.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>30</td>
<td>7.567</td>
<td>1.165</td>
<td>5.00</td>
<td>10.00</td>
<td>7.776</td>
<td>58</td>
<td>.000</td>
</tr>
<tr>
<td>Group B</td>
<td>30</td>
<td>5.033</td>
<td>1.351</td>
<td>3.00</td>
<td>10.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>6.300</td>
<td>1.788</td>
<td>3.00</td>
<td>10.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>30</td>
<td>6.167</td>
<td>1.147</td>
<td>4.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Group B</td>
<td>30</td>
<td>3.600</td>
<td>1.192</td>
<td>2.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>4.883</td>
<td>1.738</td>
<td>2.00</td>
<td>8.00</td>
</tr>
</tbody>
</table>
### Table 9. Test of within-subjects effects for the follow-up experiment

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>Sphericity assumed</td>
<td>60.208</td>
<td>1</td>
<td>60.208</td>
<td>91.217</td>
</tr>
<tr>
<td>time * group</td>
<td>Sphericity assumed</td>
<td>0.008</td>
<td>1</td>
<td>0.008</td>
<td>.013</td>
</tr>
</tbody>
</table>

### Table 10. Test of between-subjects effects for the follow-up experiment

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3752.008</td>
<td>1</td>
<td>3752.008</td>
<td>1631.104</td>
<td>.000</td>
</tr>
<tr>
<td>group</td>
<td>195.075</td>
<td>1</td>
<td>195.075</td>
<td>84.805</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>133.417</td>
<td>58</td>
<td>2.300</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The same analysis of the first experiment can be applied here, and we can see that the results of the follow-up experiment were almost the same as those of the first experiment. Group A students still outperformed Group B students on the two post-tests. Time still produced significant differences within each group, and the two teaching approaches also brought about significant differences between the two groups. With these two experiments, we can now more safely assert that the storytelling approach is more effective than the traditional rote-memorization approach in English vocabulary learning for adult e-learners.

### Conclusions

The experiment conducted by the study demonstrates the usefulness of the storytelling method in adult e-learners’ vocabulary learning. The most prominent feature of the study is that it used an L1 story context to teach L2 vocabulary, which is different from the more traditional approach of embedding L2 words in an L2 context. As the study indicates, compared with rote memorization, storytelling can bring about better learning outcomes in both the short-term vocabulary retention and the long-term retention, although its effects tend to diminish with time. Some possible ways may be used to curb the trend of diminishing. One is to do story rehearsal, which is presumed to be simpler and more enjoyable than rote memorization. Another way is to compose a series of stories which centers on a limited number of related topics, so that a later topic will reinforce the previous learning. On the other hand, the storytelling method seems to require more efforts from teachers. The survey results of the study showed that although the students enjoyed the teaching method, they themselves seemed unwilling to compose stories for themselves. This implies a big burden on teachers, if the method is to be applied over a longer period of time.

One possible limitation of the present study is that there were only a relatively limited number of target words to be learned in the experiment. As we know, college English teaching usually involves a big amount of vocabulary teaching. In one teaching session, normally more than 20 to 30 words will be taught to students. So can we embed so many target words in one story? Or even if we can do so, will the story be too long for students (especially low-achieving adult learners) to digest? The answers really need to be explored by future research.

### Acknowledgments

This work was supported by the Beijing Higher Education Young Elite Teacher Project under Grant number: YETP0471.

### References


Appendix A

The English version of the Chinese story (target words are underlined)

Xiao Long and Xiao Mei were lovers. Xiao Mei always smiled at Xiao Long, and in Xiao Long’s mind, the smile was a signal of love. Xiao Mei’s friends didn’t think their love was feasible, because Xiao Long was extremely poor and the two families did not match. Xiao Mei lived with her parents in a flat of 120 square meters, while Xiao Long’s family consisted of three generations and lived together in a flat of only 70 square meters. When Xiao Mei told her parents about her love with Xiao Long, her parents clearly expressed their opposition to their relationship. The atmosphere became tense immediately. Xiao Mei did not want to give up, so she drank a bottle of deadly poison one night. Luckily, her parents found her and sent her to the nearby hospital. With the help of the doctors, Xiao Mei survived. Now, although Xiao Mei’s parents originally did not approve their daughter’s love to Xiao Long, they had to give in now.
Engineering Students Learning Preferences in UNITEN: Comparative Study and Patterns of Learning Styles

Chen Kang Lee* and Manjit Singh Sidhu
Department of Graphics and Multimedia, College of Information Technology, Universiti Tenaga Nasional, Selangor, Malaysia // leeck82@gmail.com // manjit@uniten.edu.my

*Corresponding author

(Submitted July 8, 2014; Revised October 30, 2014; Accepted February 4, 2015)

ABSTRACT

Engineering educators have been increasingly taking the learning style theories into serious consideration as part of their efforts to enhance the teaching and learning in engineering. This paper presents a research study to investigate the learning style preference of the mechanical engineering students in Universiti Tenaga Nasional (UNITEN), Malaysia by means of the Personality and Learning Styles instruments. Descriptive statistical method was used to analyze the collected samples (n = 122). The findings of this study revealed that the preference style for engineering students are more towards spontaneous, pragmatic and concrete style of learning. This is consistent with the previous research findings on the learning preference for engineering students in UNITEN using Honey and Mumford Learning Style Questionnaire (LSQ) and Felder and Silverman Index of Learning Style (ILS) instruments. Therefore, it is important for the engineering educators to be aware of the students learning preferences by adjusting the teaching and learning strategies to unleash the learning potential of the students. Multi-approaches in teaching and learning are recommended to accommodate different students learning preferences while widening the students learning capabilities in engineering education.

Keywords

Engineering education, Learning styles, Learning preferences, Educational technology

Introduction

In the era of post-capitalist knowledge society, paradigm shift in engineering education is unavoidable in response to the rapid changes in the global market environment that emphasize on the innovation efforts for competitive advantage. Many practitioners in the engineering industry look seriously into this issue and did propose for new paradigms of engineering education in response to the changes in today’s increasingly knowledge driven environment. The details can be found in numerous research papers and reports (Chua, 2014; Froyd et al., 2012; Mistree et al., 2014; National Academies of Science, Engineering, and Medicine, 2007; National Science Board, 2007; Prados, 1998; Rajala, 2012; Rosen, 2007; Wince-Smith, 2005). In order to further enhance the quality of teaching and learning in engineering education, the students preference in learning is an important factor that should raise the attention of the education practitioners. This referred to the learning styles of the students. Learning styles are defined as “the characteristic cognitive, affective, and psychological behaviors that serve as relatively stable indicators of how learner perceives, interacts with, and responds to the learning environment”(Kolb, 1983, p. 4). As the engineering instructors, we should be aware that different students are comfortable with different learning styles (Felder & Brent, 2005). The understanding of students preferences may contribute to the adjustment of teaching strategies and the design of learning instructions that will better accommodate for students learning needs (Cavanagh & Coffin, 1994; Chen & Chiou, 2012; Graf, 2007; Graf et al., 2007; Noguera & Wageman, 2011; Pedrosa de Jesus et al., 2004).

Overview of learning styles

The research on learning styles has been active since four decades ago (Cassidy, 2004; Pedrosa de Jesus et al., 2004). There exist various definitions for learning styles. As refer by Campbell et al. (2003), learning styles is defined as a certain specified pattern of behavior according to which the individual approaches learning experience. While Felder and Spurlin (2005) defined learning styles as the different ways students take in and process information. Another popular definition for learning styles refers to individuals’ characteristics and preferred ways of gathering, organizing and thinking about information (Fleming, 2005). As noted by Kolb (1983), learning styles are not fixed personality traits but rather one’s adaptive orientation to learning. Felder and Spurlin (2005) shared similar view with Kolb by
which they stressed that “learning style profiles suggest behavioral tendencies rather than being infallible predictors of behavior” (p. 104). Many of the researchers did agree that individuals may tend to have a preference for one or two learning styles over others and the preferences can be affected by a student’s educational experience (Felder & Spurlin, 2005; Honey & Mumford, 1992; Kolb, 1983). Therefore, learning styles are “relatively stable but are not immutable” (Pedrosa de Jesus et al., 2004, p. 533). Throughout the learning process and based on different educational experience, the students may discover better way of learning and develop certain learning preferences. The learning style assessments can benefit both the instructors and students (Felder & Spurlin, 2005; Larkin-Hein & Budny, 2001). Thus, from the instructors’ perspective, the identification of the students learning styles may lead to gain better understanding of learners with different learning styles. This may contribute to assist the design of teaching and learning instructions to accommodate the different learning needs of students. While from the students’ perspective, better understanding of the learning styles may provide ideas on how the students might further improve on the less preferred styles and overcome the learning difficulties by developing the skills for balance approach in effective learning.

There are considerable amount of literature regarding the issue of learning styles (Cassidy, 2004; Coffield et al., 2004a; Coffield et al., 2004b; Deborah et al., 2012; Kolb et al., 2001; Riding & Rayner, 1998; Sadler-Smith, 1997; Sternberg & Grigorenko, 1997) and various instruments for learning styles measurements throughout all these years (Cornwell & Manfred, 1994; Dunn & Dunn, 1979; Felder & Silverman, 1988; Fleming, 2005; Gregorc, 1985; Honey & Mumford, 1992; Kolb, 1983; McCaulley, 2000). As argued by Hawk and Shah (2007), although various different instruments are available for learning styles measurement, the learning styles instruments or inventories vary in length, format and complexity, and no single instrument can capture all of the richness of the phenomena of learning styles.

Recently, engineering educators have been increasingly taking the learning style theories into serious consideration for classrooms teaching and learning (see, for example, Cagiltay, 2008; Felder, 1996; Felder & Brent, 2005; Felder & Silverman, 1988; Holvikivi, 2007; Lee & Manjit Sidhu, 2013a; Miskioglu & Wood, 2013; Rosati et al., 1988). In the literature, there are four widely accepted learning style models in engineering education context, which are the Myers-Briggs Type Indicator (MBTI), Kolb’s Learning Style Model, Felder-Silverman Model and Honey and Mumford Learning Styles Questionnaires (Manjit Sidhu, 2006; Ogot & Okudan, 2007). All these learning style instruments can be used to measure and identify the learners’ preferences.

**Kolb’s experiential learning model**

Kolb developed the learning style inventory (LSI) in 1976 and revised in 1985 (Tendy & Geiser, 1997). In Kolb’s model, students are classified as having a preference for (a) concrete experience or abstract conceptualization (how they take information in) and (b) active experimentation or reflective observation (how they process information) (Cornwell & Manfredo, 1994; Kolb, 1983; Stice, 1987). Figure 1 illustrates the learning styles and learning cycle based on Kolb’s model.

![Figure 1. Learning styles and learning cycle based on Kolb’s Model (Montgomery & Groat, 1998)](image-url)
The four types of learners in this classification (see Figure 1) scheme are:

- **Type 1** (concrete, reflective) – the diverger. Type 1 learners respond well to explanations of how course material relates to their experience, interests, and future careers. Their characteristic question is “Why?”.

- **Type 2** (abstract, reflective) – the assimilator. Type 2 learners respond to information presented in an organized, logical fashion and benefit if they are given time for reflection. Their characteristic question is “What?”.

- **Type 3** (abstract, active) – the converger. Type 3 learners respond to having opportunities to work actively on well-defined tasks and to learn by trial-and-error in an environment that allows them to fail safely. Their characteristic question is “How?”.

- **Type 4** (concrete, active) – the accommodator. Type 4 learners like applying course materials in new situation to solve real problems. Their characteristic question is “What if?”.

**Honey and Mumford’s Learning Styles Questionnaire (LSQ)**

Honey and Mumford defined learning styles as “a description of the attitudes and behaviours that determines an individual’s preferred way of learning” (Honey & Mumford, 2000, p. 6). Honey and Mumford proposed a model similar to Kolb’s in which individuals has a mixture of four learning styles, normally with a preference for one or two of the styles. Honey and Mumford’s learning styles questionnaire (LSQ) and Kolb’s learning styles inventory (LSI) are both diagnostic tests to help individuals identify their strengths, weaknesses, and development needs. The four learning styles, based on Kolb’s theory of stages in the learning cycle, are: activist, reflector, theorist, and pragmatist as shown in Table 1. Through the use of the LSQ instrument, instructors can gain a better understanding of individual learners’ attitudes, behaviors and learning processes (Armstrong et al., 2005; Jackson, 2002).

<table>
<thead>
<tr>
<th>Learning styles</th>
<th>Strengths</th>
<th>Weakness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activist</strong> (learn best when they are actively involve in new tasks)</td>
<td>Sociable, open-minded, welcome challenge, highly involved, prefers here-and-now</td>
<td>Bored by implementation details and the longer term, always seeks the limelight</td>
</tr>
<tr>
<td><strong>Reflector</strong> (learn best through review and reflect)</td>
<td>Good listener, tolerant, sees different perspectives, postpones judgement, cautious</td>
<td>Takes a back seat in meetings, low profile, distant</td>
</tr>
<tr>
<td><strong>Theorist</strong> (learn best when they can relate new information to concept &amp; theory)</td>
<td>Integrates observations with theory, rationale, objective, analytical</td>
<td>Perfectionist, detached, impatient with subjective and intuitive thinking</td>
</tr>
<tr>
<td><strong>Pragmatist</strong> (learn best when they see relevance of real life issues)</td>
<td>Experimenter, quick to adopt and try out new ideas, practival, down-to-earth</td>
<td>Impatient with theory, impatient with open-ended discussion</td>
</tr>
</tbody>
</table>

**The Myers-Briggs Type Indicator**

People are classified on the Myers-Briggs Type Indicator (MBTI) according to their preferences on four scales derived from Jung’s Theory of Psychological Types (McCaulley, 2000; Pittenger, 2005).

- **Extraverts** (try things out, focus on the outer world of people) or **Introverts** (think things through, focus on the inner world of ideas).

- **Sensors** (practical, detailed-oriented, focus on facts and procedures) or **Intuitors** (imaginative, concept-oriented, focus on meanings and possibilities).

- **Thinkers** (skeptical, tend to make decisions based on logic and rules) or **Feelers** (appreciative, tend to make decisions based on personal and humanistic considerations).

- **Judgers** (set and follow agendas, seek closure even with incomplete data) or **Perceivers** (adapt to changing circumstances, postpone reaching closure to obtain more data).
As discussed by Felder and Brent (2005), most engineering instruction is oriented toward introverts (lecturing and individual assignments rather than active class involvement and cooperative learning), intuitors (emphasis on science and math fundamentals rather than engineering applications and operations), thinkers (emphasis on objective analysis rather than interpersonal considerations in decision-making), and judgers (emphasis on following the syllabus and meeting assignment deadlines rather than on exploration of ideas and creative problem solving).

Felder-Silverman Index of Learning Styles (ILS)

The Felder-Silverman Learning Style Model classifies students along four dimensions: sensing/intuitive, visual/verbal, active/reflective and sequential/global as shown in Table 2 (Felder, 1993; Felder & Silverman, 1988). Previous studies (mainly undergraduate students in engineering domain) that adopted the Felder-Silverman ILS are summarized as shown in Table 3.

| Table 2. The four dimensions of Felder and Silverman’s learning styles |
|---------------------------------|---------------------------------|
| Learning Styles | Descriptions |
| Sensory/Intuitive | Sensors prefer facts, data, experimentation, sights and sounds, physical sensations are careful and patient with details, but may be slow. Intuitions prefer concepts, principles and theories, memories, thoughts, insights and may be quick but careless. |
| Visual/Verbal | Visual learners prefer pictures, diagrams, charts, movies, demonstrations and exhibitions. Verbal learners prefer words, discussions, explanations, written and spoken explanations, formulas and equations. |
| Active/Reflective | Active learners learn by doing and participating through engagement in physical activity or discussion. Reflective learners learn by thinking or pondering through introspection. |
| Sequential/Global | Sequential learners take things logically step by step and will be partially effective with understanding. Global learners must see the whole picture for any of it to make sense and are completely ineffective until they suddenly understand the entire subject. |

<p>| Table 3. Reported learning preferences (adapted from Felder and Spurlin, 2005) |
|---------------------------------|-----------------|----------------|---------------|-----------------|---------------|-------|
| Sampled population | A | S | Vs | Sq | N | Reference |
| Iowa State, Materials Engr. | 63% | 67% | 85% | 58% | 129 | Constant (1997) |
| Michigan Tech, Env. Engr. | 56% | 63% | 74% | 53% | 83 | Paterson (1999) |
| Oxford Brookes Univ., Business | 64% | 70% | 68% | 64% | 63 | Vita (2001) |
| British students | 85% | 86% | 52% | 76% | 21 | |
| International students | 52% | 62% | 76% | 52% | 42 | |
| Ryerson Univ., Elec. Engr. | |
| Students (2000) | 53% | 66% | 86% | 72% | 87 | Zywno&amp;Waalen (2001) |
| Students (2001) | 60% | 66% | 89% | 59% | 119 | Zywno (2002) |
| Students (2002) | 63% | 63% | 89% | 58% | 132 | Zywno (2003) |
| Faculty | 38% | 42% | 94% | 35% | 48 | |
| Tulane, Engr. | |
| Second-Year Students | 62% | 60% | 88% | 48% | 245 | Livesay et al. (2002) |
| First-Year Students | 56% | 46% | 83% | 56% | 192 | Dee et al. (2003) |
| Universities in Belo Horizonte (Brazil) | |
| Sciences | 65% | 81% | 79% | 67% | 214 | Lopez (2002) |
| Humanities | 52% | 62% | 39% | 62% | 235 | |
| Univ. of Limerick, Mfg. Engr. | 70% | 78% | 91% | 58% | 167 | Seery et al. (2003) |
| Univ. of Michigan, Chem. Engr. | 67% | 57% | 69% | 71% | 143 | Montgomery (1995) |
| Univ. of Puerto Rico-Mayaguez | |
| Biology (Semester 1) | 65% | 77% | 74% | 83% | 39 | Buxeda&amp; Moore (1999) |
| Biology (Semester 2) | 51% | 69% | 66% | 85% | 37 | Buxeda&amp; Moore (1999) |
| Biology (Semester 3) | 56% | 78% | 77% | 74% | 32 | Buxeda&amp; Moore (1999) |
| Elect. &amp; Comp. Engr. | 47% | 61% | 82% | 67% | ? | Buxeda et al. (2001) |
| Univ. of Sao Paulo, Engr. | 60% | 74% | 79% | 50% | 351 | Kuri&amp;Truzzi (2002) |
| Civil Engr. | 69% | 86% | 76% | 54% | 110 | |
| Elec. Engr. | 57% | 68% | 80% | 51% | 91 | |</p>
<table>
<thead>
<tr>
<th>Institution</th>
<th>Study Year</th>
<th>Enrollment</th>
<th>Degree</th>
<th>Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indust. Engr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univ. of Technology Kingston, Jamaica</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univ. of Western Ontario, Engr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First year engr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourth year engr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engr. faculty</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brunel Univ. UK, IS &amp; Computing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universidad de las Americas, Puebla</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>California Polytechnic State Univ.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auckland Univ., Chemical &amp; Materials Engr. course</td>
<td>34%</td>
<td>69%</td>
<td>81%</td>
<td>72%</td>
</tr>
<tr>
<td>Engr. course</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utah State Univ., Mechanical &amp; Aerospace Engr. course</td>
<td>41%</td>
<td>89%</td>
<td>89%</td>
<td>69%</td>
</tr>
<tr>
<td>Beijing Forestry Univ., Civil Engr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Personality and Learning Styles Questionnaire**

Personality and Learning Styles Questionnaire is a set of instrument that measures individual’s learning styles. The questionnaire has been designed by Richard Ogden in UK, a professional psychologist based on reliable and valid psychometric principles (Ogden, 2007). According to Richard, it is primarily intended for those in an academic environment, although the content will be useful for anyone interested in understanding more about learning and their own, special personality preferences.

The learning style model explores three key areas, highlighting how the student may prefer to go about learning things or approaching tasks based on (i) approach to learning (Structured or Spontaneous), (ii) focus on learning (Pragmatic or Conceptual), (iii) transfer of learning (Concrete or Fluid). The detail descriptions for the three key areas of learning preferences is listed in Table 4.

**Table 4. The detail descriptions of three key areas of learning preferences**

<table>
<thead>
<tr>
<th>Areas</th>
<th>Preference style</th>
<th>Structured</th>
<th>Spontaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach to learning</td>
<td></td>
<td>Structured learners are more likely to:</td>
<td>Spontaneous learners are more likely to:</td>
</tr>
<tr>
<td>(To what extent the learners need structure and organization during learning?)</td>
<td></td>
<td>• Like well organized environments and therefore might feel uncomfortable in more ambiguous situations.</td>
<td>• Learn through trial-and-error</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Prefer their learning to be well structured and formally planned.</td>
<td>• Be content with lecturers to give them “loose” and brief guidelines and they will be happy to get on with their assignments</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Are good at following step-by-step procedures.</td>
<td>• Be happy with less structured approaches to learning.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Be more likely to maintain focus and avoid distractions.</td>
<td>• Need lots of variety in their day and may get bored with routine.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Prefer to stick with tried-and-tested approaches and methods.</td>
<td>• Get stuck in; can be impatient with instructions or briefings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Get uncomfortable if things are left too “loose” or they do not know what is coming up.</td>
<td>• Enjoy spontaneity and are not worried that they don’t know what is coming up next.</td>
</tr>
<tr>
<td>Focus on learning</td>
<td></td>
<td>Pragmatic</td>
<td>Conceptual</td>
</tr>
<tr>
<td>(How interested are the learners in the underlying concepts and workings?)</td>
<td></td>
<td>Pragmatic learners are more likely to:</td>
<td>Conceptual learners are more likely to:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Focus on practical aspects e.g. how useful, and how can they apply the learnt skills to something?</td>
<td>• Enjoy understanding how things work from a theoretical perspective</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Tend to be focused on concrete, more immediate benefits of learning.</td>
<td>• They are more likely to enjoy complex, theoretical thinking about subjects such as Psychology.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Don’t necessarily see the point of being</td>
<td>• Spend time thinking about concepts and</td>
</tr>
</tbody>
</table>
absorbed into the theory or spending time on conceptual discussions.
• Believe in keeping things neat and simple.
• Make their minds up quickly, think on their feet.
• Prefer hands-on practical jobs, perhaps more suited to vocational education.
• Be more focused with the task at hand and not likely to be side-tracked by conceptual details.
• Be seen as “down-to-earth,” having “common sense” and good at getting things done.

• Taking it to a deeper level of understanding, perhaps to appreciate wider possibilities and related subjects or information.
• Appreciate the logic and rationale behind proposed procedure.
• Be more curious about how the world around them “works,” more likely to ask “why?” or “how?” in their mind.
• Carefully weigh things up and therefore on occasions may find it harder to be decisive about complex matters, less likely to take a clear stances towards matters.
• Get occasionally engrossed in the things that personally interest them and may be lose sight of the practical goals or tasks at hand.
• Very high scorers can seem to have their “head in the clouds” at times by more pragmatic people.

<table>
<thead>
<tr>
<th>Concrete</th>
<th>Fluid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concrete learners are more likely to:</td>
<td>Fluid learners are more likely to:</td>
</tr>
<tr>
<td>• Take their learning literally i.e., this learnt skill is used for this specific situation.</td>
<td>• Enjoy tackling several things at the same time.</td>
</tr>
<tr>
<td>• Find it more difficult to adapt what they have learnt to other similar situations.</td>
<td>• Like exploring the links or connections between things.</td>
</tr>
<tr>
<td>• Prefer following clear instructions and to be offered or given solutions.</td>
<td>• Automatically consider widening the original application after learning “what else can I do with this knowledge?”</td>
</tr>
<tr>
<td>• Need to concentrate on one thing at a time, working through information in a step-by-step fashion.</td>
<td>• See if it is possible to transfer and adapt learning from one situation to other very different situation.</td>
</tr>
<tr>
<td>• Have less need to review and explore what they can do with what they have learnt.</td>
<td>• Boost their learning by drawing on their own previous experiences, perhaps from totally different areas.</td>
</tr>
<tr>
<td></td>
<td>• Be more able to adapt to changing situations.</td>
</tr>
</tbody>
</table>

Furthermore, as per Ogden (2007), the student learning preference can also be represented using “Hemispheric Map” Diagram (see Figure 2). It is a simpler way of representing where the students preferences are, and also shows the student which side of brain the student may prefer to use when processing information. The left hemisphere preference and right hemisphere preference both shown different and contrast characteristics as stated in Table 5.

![Figure 2. Sample “Hemispheric Map” diagram indicated a particular student’s learning preference and the preference side of brain when processing information](image)
Table 5. The characteristics both for left hemisphere preference and right hemisphere preference

<table>
<thead>
<tr>
<th>Left hemisphere preference</th>
<th>Right hemisphere preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likes written information, responds well to spoken instructions.</td>
<td>Likes to see things in action, and responds better to demonstrations (prefer demonstrations).</td>
</tr>
<tr>
<td>Solves problems in a step-by-step, logical manner.</td>
<td>Solves problems by using “gut feel” and hunches</td>
</tr>
<tr>
<td>Looks for differences and when things don’t “fit.”</td>
<td>Looks for patterns and links</td>
</tr>
<tr>
<td>Plans and structures.</td>
<td>Looks for relationships and similarities</td>
</tr>
<tr>
<td>Prefers established, more objective types of information e.g., science.</td>
<td>Prefers more subjective, diffuse or elusive information, e.g., art, politics.</td>
</tr>
<tr>
<td>Likes asking questions, likes things with a clear answer or predictable outcome</td>
<td>Happy with ambiguity, explores matters without need for a clear answer.</td>
</tr>
<tr>
<td>Less open or valuing of feelings</td>
<td>Responds more to feelings and open with emotions.</td>
</tr>
<tr>
<td>Splits things up - makes distinctions</td>
<td>Jumps in, makes things up as go along, chunks things together</td>
</tr>
<tr>
<td>Looks for cause and effect</td>
<td>Interested more in expression than cause and effect.</td>
</tr>
</tbody>
</table>

Each individual learner may be able to identify their learning preference towards using the left side or the right side of the brain based on the report generated after the questionnaire was submitted through the online form. The report would provide specific advice to help the learners to balance his/her approach in learning and to learn more effectively.

**Methodology**

This research study aimed for two objectives. Firstly is to investigate the learning preference for engineering students in UNITEN by means of the Personality and Learning Styles instruments. Secondly is to compare the current research finding with the previous learning styles research on engineering students in UNITEN.

In this study, the personality and learning styles questionnaire was adopted. This is because (i) it is free and easily available online with prompt response on the student’s learning styles preference, (ii) it is primarily intended for those in an academic environment and may suit to engineering education and (iii) provides customized and constructive feedback for each individual regarding their learning behavior and preferences. The current findings will be compared with two of the previous research in learning styles (Lee & Manjit Sidhu, 2013a; Manjit Sidhu, 2006) as discussed in the next section. The personality and learning styles questionnaire is comprehensive and is about 12-15 pages long. It is available online without any charges. In the questionnaire, students are asked to respond themselves against the 76 statements using a five (5) point Likert Scale. Once the student submits the questionnaire online, an instant report will be generated and provides suggestions on how the student can utilize his/her mind and adapt their behavior to learn more effectively.

In August 2012, five sections of the 3rd year mechanical engineering students from Universiti Tenaga Nasional (UNITEN) were invited to participate in this research study and a total of 122 samples were collected through online questionnaire. The response rate of the survey achieved 81.3% due to the active participation from the students. Two mechanical academic staff were involved in assisting the students throughout the process. Short briefing was provided regarding the purpose of this research and the procedures needed to complete the questionnaire before the students started to fill in the questionnaire. Each student took approximately 15-20 minutes to complete the online questionnaire. The reports generated were collected for analysis purpose using descriptive statistical method.

Two weeks before the actual survey, two of the academic staff and three of the final year students were invited to try on the pilot questionnaire. The pilot study served two purposes, firstly is to determine the durability for completion of the questionnaire and secondly is to identify any potential problem(s) with the questionnaire design (layout and readability). Feedbacks were collected at the end of the pilot study. The questionnaire design (layout and readability) was found to be professional (simple and clear to understand). The duration for the questionnaire completion was recommended to be 15 to 20 minutes. The data was collected and the details were analyzed as referred in the following section.
Data analysis and results

Through the reports generated, the students (respondents, $n = 122$) may be classified as adopting a particular learning style preference, based on the score obtained on individual scale using the personality and learning style questionnaire. According to Ogden (2007), the learning model for each student comprises of three key areas, each split down the middle into two opposite styles (Figure 3). The student’s response may fall in either side of the key area and this indicated a “type” or preference in learning as shown in Figure 3. If the student is not scoring much out of the middle band for a particular area (the higher the scale scored, the stronger the preference), this showed that the student does not have a strong preference one way or the other in that area, which means the student is likely to take a balanced approach.

Figure 3. Sample graph indicated a student’s learning preference according to three key areas

As referred to Ogden (2007) and discussed in previous section, the model explores three key areas, highlighting how the student may prefer to go about learning things or approaching tasks. Based on the data collected from 122 students, the learning preference scale score and its distributions are compiled as shown in Table 6.

Table 6. The learning preference scale score of the engineering students ($n = 122$)

<table>
<thead>
<tr>
<th>Key areas</th>
<th>Scale</th>
<th>Scale</th>
<th>Scale</th>
<th>Scale</th>
<th>Scale</th>
<th>Scale</th>
<th>Scale</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach</td>
<td>structured</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Focus</td>
<td>pragmatic</td>
<td>9</td>
<td>19</td>
<td>51</td>
<td>26</td>
<td>10</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Transfer</td>
<td>concrete</td>
<td>2</td>
<td>13</td>
<td>43</td>
<td>29</td>
<td>28</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

The graph shown in Figure 4 further illustrates the distribution of students’ learning preference scale score that was derived from data in Table 6. According to the student’s learning preference in three key areas, there will be eight possible combination of learning preferences. A detailed distribution of the engineering student’s preferences according to the eight combination of learning preference is compiled and listed as shown in Table 7.

From the 122 respondents (see Figure 4 and Table 7), it can be clearly identified that 77.87% of the engineering students have the learning preference on spontaneous + pragmatic + concrete style. This is followed by 10.66% of the engineering students having the learning preference on spontaneous + pragmatic + fluid style. While 4.1% of the students have the learning preference on the structured + pragmatic + concrete style and spontaneous + conceptual + concrete style. The remaining 3.3% of the students prefer the learning in structured + conceptual + concrete style.

As discussed in the previous section, the student learning preference can also be represent using “Hemispheric Map” diagram (see Figure 2). It is based on the same scales and data as shown in the graph (see Figure 3) previously, but presents data in an alternative way. The eight possible combinations of learning preferences according to the 3 key areas can be further categorized into either the left hemispheric preference or right hemispheric preference. Table 8 shows the detailed categorization of students’ scoring and learning preferences according to the hemispheric map. It is interesting to know that more than 80% of the students have the left preference side of brain when processing information. Only 14.76% of the students response indicated for right preference side of brain when processing information.
Table 7. Detailed distribution of UNITEN engineering student’s learning preferences (3 key areas)

<table>
<thead>
<tr>
<th>Learning preferences</th>
<th>Sample (n)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structured + Pragmatic + Fluid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Structured + Pragmatic + Concrete</td>
<td>5</td>
<td>4.1</td>
</tr>
<tr>
<td>Structured + Conceptual + Concrete</td>
<td>4</td>
<td>3.3</td>
</tr>
<tr>
<td>Structured + Conceptual + Fluid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Spontaneous + Conceptual + Fluid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Spontaneous + Pragmatic + Fluid</td>
<td>13</td>
<td>10.66</td>
</tr>
<tr>
<td>Spontaneous + Pragmatic + Concrete</td>
<td>95</td>
<td>77.87</td>
</tr>
<tr>
<td>Spontaneous + Conceptual + Concrete</td>
<td>5</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>122</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 8. Categorization of students’ learning preference according to the hemispheric map

<table>
<thead>
<tr>
<th>Left hemisphere preference</th>
<th>Percentage (%)</th>
<th>Right hemisphere preference</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structured + Pragmatic + Fluid</td>
<td>0</td>
<td>Structured + Conceptual + Fluid</td>
<td>0</td>
</tr>
<tr>
<td>Structured + Pragmatic + Concrete</td>
<td>4.1</td>
<td>Spontaneous + Conceptual + Fluid</td>
<td>0</td>
</tr>
<tr>
<td>Spontaneous + Pragmatic + Concrete</td>
<td>77.87</td>
<td>Spontaneous + Pragmatic + Fluid</td>
<td>10.66</td>
</tr>
<tr>
<td>Structured + Conceptual + Concrete</td>
<td>3.3</td>
<td>Spontaneous + Conceptual + Concrete</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>85.27</td>
<td></td>
<td>14.76</td>
</tr>
</tbody>
</table>

Discussion

Through the research findings, it can be identified that more than 75% of the engineering students prefer the spontaneous + pragmatic + concrete style and a preference for using the left side of the brain. Spontaneous style learners get the best out of learning when they are allowed to get stuck quickly and try things for themselves. Furthermore, they prefer to learn through the “trial and error” process with sufficient “doing” activities. They will easily get bored with routine and will need lot of variety learning activities to keep them focused. Meanwhile, pragmatic style learners are more focused on the practical, tangible and immediate benefits of learning things. They learn more with “hands-on” exercises, easily get bored with learning about theories or concepts that are complex or less relevant and prefer to keep things simple and easy-to-understand.
The strong preference towards spontaneous style and pragmatic style of learning are consistent with the findings in the previous research (Lee & Manjit Sidhu, 2013a). In the previous findings (see Table 9), engineering students showed a very strong preference in activist learning style using the Honey and Mumford’s LSQ. As known, activist learners are those learners that learn best when they are actively involve in new tasks. They are highly involved in tasks, preferred new challenge and likely to learn through the trial and error process. Furthermore, in another research findings using the Felder and Silverman ILS (see Table 9), engineering students also showed higher preference in active dimension as compared to the reflective dimension using Felder and Silverman’s learning styles questionnaire. This is also consistent whereby the research findings for engineering students in other universities reported the same (higher preference towards active dimension - see Table 3).

As defined by Felder and Silverman (1988), active learners learn by doing and participating through engagement in physical activity or discussion. Thus, Martínez Cartas (2012) recommended the instructors to provide practical troubleshooting methods or drill exercises to provide practice in order to accommodate the learning preference for active learners. This clearly provide evidence that the traditional “chalk and talk,” passive and one way delivery of the teaching and learning process for the engineering subjects would not fit well with the students’ learning preference. More variety of learning activities that actively involved the students in performing new tasks should be introduced, for example: practical troubleshooting methods and active experimentation.

Table 9. Comparison of three research studies on UNITEN engineering students learning preferences

<table>
<thead>
<tr>
<th>Learning style instruments</th>
<th>The Felder-Silverman ILS</th>
<th>Personality and learning styles questionnaire</th>
<th>Honey and Mumford’s LSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey conducted</td>
<td>(Manjit Sidhu, 2009)</td>
<td>August 2012</td>
<td>(Lee &amp; Manjit Sidhu, 2013a)</td>
</tr>
<tr>
<td>Respondents (n)</td>
<td>n = 60</td>
<td>n = 122</td>
<td>n = 104</td>
</tr>
<tr>
<td>Findings (Learning Preferences)</td>
<td>85% geared towards sensing learners</td>
<td>(1) More than 80% of the respondents have the preference to use the left side of the brain for thinking and learning.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100% geared towards visual learners</td>
<td>(2) Left hemisphere preference:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>90% geared towards active learners</td>
<td>(i) solves problems in a step-by-step, logical manner</td>
<td></td>
</tr>
<tr>
<td></td>
<td>86.7% geared towards sequential learners</td>
<td>(ii) prefers established, more objective types of information</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(iii) Likes things with a clear answer or predictable outcome</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(iv) Less open or valuing of feelings</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(v) Looks for cause and effect</td>
<td></td>
</tr>
</tbody>
</table>

With the advancement in ICT, many technologies can be utilized to aid in the formal teaching and learning of engineering education (Fernandez et al., 2011; Yueh et al., 2014). For example, TAPS (Technology Assisted Problem Solving) packages (a type of computer aided learning software tool) were developed and used by students to assist their learning in engineering problem solving. The development of engineering simulation software tool such as the Desktop Virtual Reality TAPS package (see Table 10) would help the students (with spontaneous and pragmatic learning style) to experience the operations process of the engineering mechanism through the interaction with 2-D and 3-D animations. The problem solving mechanism in the engineering software tool provides timely feedback and assessment directly related to learner’s interactions. This will lead to exploratory approach in learning and better visualization of the engineering concepts especially those problems that involves dynamic motion. Other researches also supports that visualization skill is important in engineering and science (Ault & John, 2010; Pedrosa et al., 2014).

As for the concrete style learners, they are good at applying their learning to clearly define problems or questions. They prefer to follow a step-by-step approach to learn a well-defined task or a clear, straightforward subject area. This finding is consistent with the previous research study that engineering students have higher preferences towards sequential dimension in Felder and Silverman learning style model (see Table 9).
Table 10. Recommended ICT tools to support the learning styles of engineering students

<table>
<thead>
<tr>
<th>Style of learning</th>
<th>Characteristic</th>
<th>ICT</th>
<th>General outcome and benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spontaneous</td>
<td>-prefer to learn through the “trial and error” process with sufficient “doing” activities -easily get bored with routine and will need lot of learning activities to keep them focused.</td>
<td>3-D Desktop Virtual Reality TAPS package -employed 2-D graphics, 3-D geometric model, colors, sound and “tweening” technique to produce animation</td>
<td>-Allow student to experiment simulated problem-solving in 3-D environment for better interaction and visualization -With simulations, students are able to take a more active role in learning. -Dynamic representations enable more efficient communication of complex engineering concepts -The problem solving environment provides timely feedback and assessment directly related to user’s interactions. -The stereoscopic images and animated models could create the great interest and enthusiasm in learning.</td>
</tr>
<tr>
<td>Pragmatic</td>
<td>-prefer to learn more with “hands-on” exercises -easily get bored with learning about theories or concepts that are complex or less relevant -prefer to keep things simple and easy-to-understand</td>
<td>Coach based TAPS package -employed simple expert system rules to coach user through the sequential order in engineering problem solving task - employed 2-D graphics, colors, sound and “tweening” technique to produce digital learning environment</td>
<td>-Coaching is provided to enhance the user’s problem solving experience while performing complex tasks in 2-D environment -The problem solving environment provides timely feedback and assessment directly related to the user’s interactions. -The feedback helps increase the user’s ability to reason and analyse the problem solving environment.</td>
</tr>
<tr>
<td>Concrete</td>
<td>-prefer to follow a step-by-step approach to learn a well-defined task or a clear, straightforward subject area.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As described by Felder and Silverman (1988), sequential learners preferred linear thinking process rather than holistic thinking process and usually will learn through small incremental steps. This indicated that the engineering students have the preferences to learn and solve the engineering problems through the logical step-by-step approach. In order to maximize the learning experience for better knowledge absorption, the teaching and learning strategies should emphasize more on the systematic and sequential step-by-step approach for engineering problem solving.

Currently, it can be clearly identified that the teaching and learning in engineering for problem solving did emphasize on the sequential step-by-step approach. However, it was found that there was insufficient justification for each of the step involved throughout the problem solving process especially those sample questions provided in the engineering textbooks. Although many of the engineering instructors did perform verbal explanations regarding the steps involved by using the limited sample problems provided in the textbook, many of the students (especially those slow learners in term of engineering domain) still face the difficulties in gaining the understanding regarding the steps involved and the justification for it. Thus, it is suggested that coach based interactive Technology Assisted Problem Solving Packages (TAPS) can be used to aid in the students learning process through the step-by-step guidance procedure (Manjit Sidhu, 2006, see Table 10). This may enhance the students understanding regarding the “what,” “why” and “how” for each of the step involved for the engineering problem solving according to their own learning time and pace. More details regarding the design and development of TAPS packages can be referred to (Manjit Sidhu, 2006).

From the research findings, it can be identified that more than 80% of the students (see Table 8) have the preference for using the left side of the brain. For the left brain dominant learners, the students will be more comfortable to learn and solve problems in a step-by-step, logical manner. This is useful when dealing with complex engineering
problems whereby the students will normally break the complex problem into different parts and solve the problem part by part in a sequential order till the final outcome is achieved. Through this, the students may be able to see the logical flow of the questions and this is an effective way of solving the engineering problem. The details of the engineering problem solving approach can be referred to previous research (Lee & Manjit Sidhu, 2013b).

In order to strike a balanced learning approach, the engineering instructors may put in some efforts to guide and train the students to exercise their right brain by relating the thinking sense in a wider picture. This means that the students should be trained to think how to relate a specific problem to the main topic in learning and how does it fit to the learning outcomes and learning objectives. At the same time, how this particular problem solving may relate to the real life applications and how does it benefit/contribute to the welfare of the society should also be part of the concern for the students (as the open ended questions for thought) besides the traditional learning in developing the problem solving skills. So, in the long run and for developing balanced learning skills, students should be well equipped with solving the engineering problems both through the bottom-up approach (exercise the use of left brain) and the top-down approach (exercise the use of right brain). As stressed by Franzoni and Assar (2009), “whole brain” learning is known to be a far more effective way to learn.

In order to exercise the right brain, the students may try to think out of the box when solving the engineering problems by providing the “what-if” alternative solution. This may train the students to start valuing the problem by providing not only the standard solution but also some alternative creative ideas for it. Regarding this issue, the engineering instructors are encouraged to create the opportunity in the teaching space to allow students to think out of the box and propose new ideas. Ideas are more important and students should be provided with the opportunity and space to raise their ideas. Creative ideas are the main source for radical innovation that may lead to potential paradigm shift in the near future. The learning activities such as the brainstorming session in a small group, ideas presentation should be encourage and embedded in the class especially for engineering design subjects. This provide the opportunity for the students to express their thoughts, opinions and feeling towards engineering design and at the same time cultivate their interest to appreciate the previous design and the current design. This is consistent with the discussion on the new paradigm of engineering education as can be found in Duderstadt (2008).

**Conclusions**

Through the research findings, it can be identified that engineering students in UNITEN shared a common learning style preferences that they preferred to learn in sequential, logical way with various hands-on practical activities. They are “open-minded” and prefer to try out new forms of learning activities. However, the current teaching and learning approach may not fully accommodate the learning preference of the students. The efforts on how to shape a balance learning approach by utilizing the current learning preference and strengthen the less preference way of learning is a great challenge faced by the engineering instructors. Carefully designed and well structured ICT software tools that match the learning styles of engineering students could strengthen their problem solving skills. The lessons from the findings of this research reminded engineering educators that we need to provide multi-approach in teaching and learning in order to accommodate different students learning preferences while widening the students learning capabilities. Wider learning capabilities will contribute to shape a professional engineer not only to equip with the technical know-how but also with broadened perspective in serving the human nation for long term sustainability growth.

**References**


280


Development of an Inquiry-Based Learning Support System Based on an Intelligent Knowledge Exploration Approach

Ji-Wei Wu¹, Judy C. R. Tseng²* and Gwo-Jen Hwang³

¹Institute of Education, Jiangsu Normal University, Xuzhou, Jiangsu, China // ²Department of Computer Science and Information Engineering, Chung Hua University, Hsinchu, Taiwan // ³Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan // smallwu315@gmail.com // judycrt@chu.edu.tw // gjhwang.academic@gmail.com

*Corresponding author

(Submitted September 11, 2014; Revised December 16, 2014; Accepted January 4, 2015)

ABSTRACT
Inquiry-Based Learning (IBL) is an effective approach for promoting active learning. When inquiry-based learning is incorporated into instruction, teachers provide guiding questions for students to actively explore the required knowledge in order to solve the problems. Although the World Wide Web (WWW) is a rich knowledge resource for students to explore using search engines, it is hard for students to filter out required knowledge from the unorganized web pages and the huge number of search results. In this paper, an inquiry-based learning support system with two advanced knowledge exploration modules, namely the Q&A module and the Segmented Supplemental Material (SSM) module, is proposed and developed to assist students’ inquiry-based learning. This study also evaluates the effects of the proposed system on students’ learning achievement and learning efficiency with the various supports provided by the proposed system. Experiments were undertaken on a “Computer Science” course for first year undergraduate students. The experimental results show that with the support of the two advanced modules, both the students’ learning achievement and their learning efficiency were enhanced.

Keywords
Inquiry-based learning, IBL, Knowledge exploration, Web-based learning, Intelligent tutoring systems

Introduction
Inquiry-Based Learning (IBL), contrary to traditional pedagogy, is a rather effective approach for promoting active learning. Students are engaged in the learning activities, and the focus is shifted from teacher-led to student-centered learning (Pedaste & Sarapuu, 2006). When IBL is conducted as the teaching/learning strategy, students need to extensively explore the necessary information or knowledge in their problem solving process. Thanks to the rapid growth of the Internet, the World Wide Web (WWW) makes available various resources which are able to provide students with the required knowledge during their learning process. However, the students may not have sufficient abilities to discover and filter out suitable knowledge from the unorganized web pages and the huge number of search results. Even if they do have such abilities, it might cost them a great amount of effort and time as it is necessary to read through numerous search results to find the required knowledge. Furthermore, teachers cannot know students’ learning situation during their learning process.

From the literature, it can be observed that guidance in learning core concepts is beneficial to students’ learning (Dufresne, Gerace, Hardiman, & Mestre, 1992; Sandoval & Reiser, 2004; Nelson, 2013). Also, it has been suggested that students with a more constructivist epistemological framework perform better in inquiry contexts (Linn & Songer, 1993; Tobin, Tippins, & Hook, 1995; Windschitl & Andre, 1998). For these reasons, an Inquiry-Based Learning (IBL) support system with two advanced knowledge exploration modules, namely the Q&A module and the Segmented Supplemental Material (SSM) module, is proposed and developed in this paper. The Q&A module provides students with related historical questions asked by students and associated answers given by the teacher. A mechanism which is able to instantly find existing answers to students’ questions was developed. Moreover, students can ask for assistance from the teacher through the Q&A module. The SSM module applies an efficient text segmentation algorithm (Wu, Tseng, & Tsai, 2011) to divide the knowledge documents provided by the teacher into topical segments. For example, assume that a knowledge document related to the topic Network is provided by the teacher; the process of the text segmentation is to divide the document into topically coherent segments, each of which is related to a subtopic, such as Local Area Networks (LAN) and Wide Area Networks (WAN). A search
mechanism is provided to rapidly acquire the most relevant parts (i.e., topical segments) of the knowledge documents for students’ questions. Both of these advanced knowledge exploration modules eliminate the need for students to filter out the required knowledge from a huge number of web pages by themselves.

Moreover, students’ profiles and learning history are automatically recorded by the proposed system. This will help the teacher to further analyze and explore whether and why conducting IBL will improve students’ learning achievement and learning efficiency.

To evaluate the effects of the proposed system in students’ inquiry-based learning process, a field experiment was undertaken on the freshman course “Computer Science.” An inquiry-based learning scenario was designed, and the students’ learning achievement and learning efficiency with the various supports provided by the proposed system were investigated. The experimental results reveal that providing the two advanced knowledge exploration modules is helpful to students in terms of improving both their learning achievement and learning efficiency.

Literature review

Inquiry-Based Learning

Inquiry-Based Learning (IBL) is a student-centered pedagogy that focuses on questioning, critical thinking, and problem solving (Liu, Lee, Chang, Trapahgan, & Horton, 2006). It has been demonstrated that IBL can engage students in learning. The basic idea of IBL is the question, which may be established by students, teachers or by negotiation among them (Levy, Aiyegbayo, & Little, 2009). Besides asking meaningful questions, finding answers, which engages students in meaning making, is also involved in IBL activities (Liu et al., 2006; Levstik & Burton, 1997). During the process of IBL, the information needed to solve the questions would not have been previously covered explicitly in lectures or readings, although it would normally build on previously known materials (Prince & Felder, 2007).

In contrast to traditional pedagogy, IBL focuses on students’ active learning. Their role is changed from passive listeners to active explorers, while the role of the teacher is changed from lesson instructor to problem-solving guide. Moreover, thanks to the rapid growth of the Internet, instead of traditional textbooks, knowledge resources from the WWW have become a mainstream element in learning activities. It is beneficial to teachers to conduct IBL, and students can learn knowledge anywhere and at any time.

IBL has been widely adopted by educators since the 1980s; successful cases have been reported for various subjects and student levels (Chiang, Yang, & Hwang, 2014; Liu et al., 2006; Shih, Chuang, & Hwang, 2010; Wang, 2010). In recent years, several researchers have been devoted to investigating methods of web-based education to encourage student-centered learning (Hwang, Kuo, Chen, & Ho, 2014; Wu, Hwang, Kuo, & Huang, 2014). Those studies include qualitative and quantitative approaches, and system development and evaluation. Pedaste and Sarapuu (2006) developed a support system for IBL for problem-solving tasks related to a virtual hike. Oliver (2008) explored the utility and efficacy of the application of a Web-based tool to promote learner engagement among first year undergraduate students. Kim and Yao (2010) developed a prototype of a web-based learning support system, called OTHI (Online Treasure Hunt for Inquiry based learning). Raes, Schellens, De Wever, and Vanderhoven (2011) investigated the impact of different modes of scaffolding on students learning science through a web-based collaborative inquiry project in authentic classroom settings. Sun and Looi (2013) developed a web-based science learning environment, called WiMVT (Web-based inquirer with Modeling and Visualization Technology).

From the literature, it has been demonstrated that IBL is effective in promoting students’ learning performance. However, some existing IBL support systems provide students only with teaching materials during their IBL activities. In this case, the teaching materials may not always be sufficient for solving the students’ learning problems. Also, the rich knowledge resources provided by the Internet are unavailable. Although some existing IBL support systems provide students with resources on the Internet, the students may not have adequate abilities to discover and filter out suitable knowledge, and even if they do, it will cost them great effort and time as it is necessary to read through the vast number of search results to find what they require. Furthermore, teachers do not
know students’ learning situation during their learning process. It is thus essential to develop a versatile learning assistant system to support students with advanced knowledge exploration modules when Inquiry-Based Learning (IBL) is conducted.

**Epistemic scaffold**

Similar to the scaffolds built to perform construction at construction sites, an epistemic scaffold is a useful object built as a support structure for knowledge production. An epistemic scaffold is built through a process of building up a structure of arguments and evidence to substantiate the use of the model to do knowledge production work (Nelson, 2013).

In the past, epistemic scaffolds have been successfully applied in various studies. To help novice physics students carry out expert-like (hierarchically structured) problem analyses, Dufresne, Gerace, Hardiman and Mestre (1992) provided questions sequenced to be consistent with expert-like strategies for a problem-oriented learning task. Several studies have reported the effectiveness of using epistemic scripts to support learning, in particular, improving novices’ ability to solve problems (Dufresne, Gerace, Hardiman, & Mestre, 1992; Sandoval & Reiser, 2004; Nelson, 2013). For example, Sandoval and Reiser (2004) employed an epistemic support in conducting a technology-based inquiry activity. From the findings, they suggested that epistemic tools could play a unique role in supporting students’ inquiry. To better understand how scientists create and contest claims about the utility of animal models, Nelson (2013) introduced the metaphor of an epistemic scaffold. The metaphor supports animal behavioral models as experimental platforms for producing knowledge. Also, the metaphor helps reframe various methodological disputes.

From the literature, it is observed that guidance while learning core concepts is beneficial to students’ learning (Dufresne, Gerace, Hardiman, & Mestre, 1992; Sandoval & Reiser, 2004; Nelson, 2013). Also, it is suggested that students with a more constructivist epistemological framework perform better in inquiry contexts (Linn & Songer, 1993; Tobin, Tippins, & Hook, 1995; Windschitl & Andre, 1998). In this study, an inquiry-based learning support system with two advanced knowledge exploration modules, namely the Q&A module and the Segmented Supplemental Material (SSM) module, has been developed based on the two epistemic scaffolds. From the experimental results, it is inferred that the two mechanisms can benefit students in terms of improving their IBL performances via guiding them to find core concepts and providing the structured knowledge related to the target problems.

**An Inquiry-Based Learning (IBL) support system with advanced knowledge exploration modules**

In order to make the process of students’ knowledge exploration more efficient, an IBL support system with advanced knowledge exploration modules is proposed and developed in this paper. The architecture of the proposed system is shown in Figure 1.

In the proposed system, three knowledge exploration modules, namely a Web module, a Q&A module, and a Segmented Supplemental Material (SSM) module, are included. Although web search is readily available through popular search engines, such as the Google search engine, we include the Web module in our system to record students’ learning history of Internet searches for performance analysis (Hung, Hwang, Lin, Wu, & Su, 2013; Hwang, Tsai, Tsai, & Tseng, 2008). The Q&A module and the SSM module are advanced knowledge exploration modules which are newly proposed to provide selected and organized knowledge for students to explore.

The Web module contains a web searching component and a learning history database. Once a question is submitted by a student, the related web pages will be retrieved through the web searching component and provided to the student for reference.

The Q&A module contains a Q&A searching component, a Q&A database, and a teacher assistance component. Once a question is submitted by a student, related historical questions will be searched for and the associated answers
are retrieved from the Q&A database through the Q&A searching component. If the retrieved knowledge is not satisfactory, the student can ask for assistance from the teacher through the teacher assistance component. Once the student’s request for the teacher’s assistance is submitted, the teacher assistance component will automatically notify the teacher by email. The teacher can answer the student’s question through the teacher assistance component at any time. After the teacher has answered the student’s question, the answer will be emailed to the student and the solved question and corresponding answer will be added to the Q&A database for future knowledge exploration.

![Figure 1. System architecture of the inquiry-based learning support system](image)

The SSM module contains an SSM searching component, an SSM database, and an SSM management component. Through the SSM management component, teachers can provide documents of supplemental materials at any time. Those documents will be divided into topical segments, each of which contains only the sentences related to one specific topic, before they are stored in the SSM database. Once a question is submitted by a student, the related SSMs will be retrieved through the SSM searching component from the SSM database and provided to the student as a reference. The search results of the SSM searching component are topical segments, rather than the original documents, of supplemental materials that the teacher provides. In such a design, the burden of filtering out required knowledge from a long document will be reduced.

During the IBL activities, students’ learning history is automatically recorded in the learning history database. In this database, each student’s login and logout times are recorded. Also, students’ learning process, such as posed questions, browsed web pages, browsed Q&As, browsed SSMs, and browsing time, is recorded. The aim of recording students’ learning history is to assist teachers with further information on students’ situations during their IBL process. Teachers can evaluate students’ performance during their study. Also, the factors which affect students’ learning achievement can be identified by analyzing the learning history records.

**Web searching component**

The design of the web searching component is shown in Figure 2. The knowledge source of the web searching component is web pages on the Internet. The workflow of the web searching component is described below: First, the submitted question is set as a parameter and transmitted to the Google search engine. Then, after the Google searching process, the results, including titles and abstracts of related web pages, are retrieved and displayed on the user interface (as shown in Figure 3). The student can further explore the details of a web page by clicking the link contained in the search results. The student can also bookmark the web page, and provide feedback information about whether the web page is helpful in solving his/her question (as shown in Figure 4).
Figure 2. The design of the web searching component

Figure 3. Knowledge exploration through the web searching component

Figure 4. Presenting the information of a web page
Q&A searching component

The design of the Q&A searching component is shown in Figure 5. The knowledge source of the Q&A searching component is the Q&A database which consists of questions students have previously posed and corresponding answers provided by the teacher. Each of the questions in the Q&A database is represented as a Characteristic Vector (CV):

\[ CV(Q_i) = \{ (k_1, w_{i1}), (k_2, w_{i2}), \ldots, (k_j, w_{ij}), \ldots, (k_n, w_{in}) \} \]  

(1)

where \( Q_i \) represents question \( i \); \( k_j \) represents keyword \( j \); and \( w_{ij} \) represents the weight of \( k_j \) in \( Q_i \). The weight of each keyword is calculated by TF×IDF (Salton & McGill, 1983):

\[ w_{ij} = tf_{ij} \times \log \left( \frac{N}{df_j} \right) \]  

(2)

where \( tf_{ij} \) represents the term frequency of \( k_j \) in \( Q_i \); \( N \) represents the number of questions in the Q&A database; and \( df_j \) represents the number of questions containing \( k_j \). The weight \( w_{ij} \) represents the importance of keyword \( j \) in \( Q_i \). The higher term frequency (\( tf_{ij} \)) and lower document frequency (\( df_j \)), which make higher \( w_{ij} \), indicates higher importance of \( k_j \) in \( Q_i \).

The workflow of the Q&A searching component is described as follows: First, the student submits a question in the form of a natural language sentence or a group of keywords through the user interface. If the question is in the form of a natural language sentence, the sentence will be preprocessed to extract the contained keywords (Wu, Tseng, & Tsai, 2011). Then, the submitted question is represented as a CV. The CV is used to compare with the CVs saved in the Q&A database for matching relevant questions. To compare the CV of the submitted question and that of the questions saved in the Q&A database, the inner product similarity measure is adopted:

\[ s_i = \sum_{j=1}^{n} (w_j \times w_{ij}) \]  

(3)

where \( s_i \) represents the similarity between the submitted question and question \( i \) in the Q&A database (i.e., historical question \( i \)); \( w_j \) represents the weight of keyword \( j \) in the submitted question; and \( w_{ij} \) represents the weight of \( k_j \) in \( Q_i \). A higher \( s_i \) indicates higher similarity between the submitted question and historical question \( i \). The historical questions with non-zero similarity are sorted in descending order and are returned to the student for reference (as shown in Figure 6). The student can further explore a related historical question and its corresponding answer by...
clicking the link contained in the search results. When the student reads the information, he/she can provide feedback information about whether the historical Q&A is helpful in solving his/her question (as shown in Figure 7).

Teacher assistance component

The design of the teacher assistance component is shown in Figure 8. If the student is not satisfied with the retrieved knowledge materials from the Q&A database, the SSM database, or the Internet, he/she can ask for assistance from the teacher through this component (as shown in Figure 9). When a question is posed by a student, a notification will be automatically sent to the teacher asking for assistance.

After receiving the notification, the teacher will reply to the student’s question through this component (as shown in Figure 10). Then, the newly solved question and the corresponding answer will be added to the Q&A database for future knowledge exploration. Through this component, the solved questions will contribute to the Q&A database and make knowledge reusing more efficient.
SSM management component

The design of the SSM management component is shown in Figure 11. The workflow of the SSM management component is described as below: First, documents of supplemental materials are uploaded by the teacher (as shown in Figure 12). Then, each of the documents is divided into several topical segments by a text segmentation algorithm. The aim of dividing the supplemental materials in this way is to reduce the students’ burden of filtering out the required information from the documents, which is time consuming. The topical segments, named Segmented Supplemental Materials (SSM), can help students obtain the most relevant information directly, instead of spending a considerable amount of time exploring the whole document of supplemental material. This approach makes students’ knowledge exploration more efficient.
In this paper, an efficient text segmentation algorithm from our previous work (Wu, Tseng, & Tsai, 2011), named TSHAC, is adopted. TSHAC is a Hierarchical Agglomerative Clustering (HAC) (Willett, 1988) based text segmentation algorithm which segments a text into topical segments according to the following steps:

1. **Text preprocessing**: Each of the sentences in the text is preprocessed by keyword extraction, and is represented as a CV.

2. **Sentence-similarity matrix construction**: The similarities between each sentence and its neighbour are firstly computed by the cosine similarity measure (Wu, Tseng, & Tsai, 2011):

   \[
   \cosine(s_i, s_j) = \frac{\sum_i w_{ik} w_{jk}}{\sqrt{\sum_i w_{ik}^2} \sqrt{\sum_j w_{jk}^2}}
   \]  

   where \(s_i\) represents sentence \(i\) (i.e., the CV of sentence \(i\)); and \(w_{ik}\) represents the weight of keyword \(k\) in sentence \(i\).

   The cosine similarity measure is used to compute the cosine of the angle between the CVs of sentences \(i\) and \(j\). A higher cosine similarity indicates higher similarity between sentences \(i\) and \(j\). Then, the sentence gaps with zero similarity are regarded as dividing points, and the text is divided into several groups. The number of sentences included in the biggest group is named block size. After this, a sentence-similarity matrix is constructed. Each of the elements is the cosine similarity between two sentences. To make the process of text segmentation more efficient, only the similarities alongside the main diagonal (with block size) are computed.

3. **Topic boundary identification**: To identify topic boundaries, a fitness function is designed to evaluate a potential segmentation of a text with the sentence-similarity matrix. Three factors, the dissimilarity between two consecutive segments that respectively precede and succeed the potential topic boundary, the standard deviation of the segment size, and the number of sentences contained in segments, are considered in the fitness function:

   \[
   f = \frac{\sum_{i=1}^{\text{seg}\times} \text{sim}_{\text{inner}}^j - \text{sim}_{\text{outer}}^j}{\text{seg} \times \text{seg}_{\text{inner}}} - \text{STDEV} - \frac{\text{seg}_{\text{inner}}}{\text{seg}}
   \]  

   where \(\text{sim}_{\text{inner}}^j\) represents the average sentence similarity within two consecutive segments that respectively precede and succeed the potential topic boundary \(j\); \(\text{sim}_{\text{outer}}^j\) represents the average sentence similarity between two
consecutive segments that respectively precede and succeed the potential topic boundary \( j \); \( STDEV \) represents the standard deviation of the segment size; and \#\text{seglen} \leq 2 \) represents the number of segments containing less than or equal to two sentences. Then, HAC is applied to find the best segmentation which achieves the highest fitness value: each of the divided groups generated in Step 2 is firstly regarded as an independent cluster. The similarities between each cluster and its neighbor are computed by considering the lexical cohesion and the standard deviation of the cluster size:

\[
Sim(c_i, c_j) = \frac{2 \times \sum_{s \in c_i, s' \in c_j} \cos \theta(s_m, s_n)}{(|c_i| + |c_j|)^2} - STDEV
\]

where \( c_i \) represents cluster \( i \); \( s_m \) represents sentence \( m \); and \( STDEV \) represents the standard deviation of the cluster size. According to the similarities, the nearest pair of adjacent clusters is merged. The procedures of similarity computation and cluster merging are repeated until only one cluster is left. After this, a hierarchical cluster tree is constructed. The fitness function is used to cut the hierarchical cluster tree to achieve the highest fitness value. Finally, the input groups are divided into several clusters, each of which is regarded as a topical segment.

After topical segment generation, each of the topical segments is saved in the SSM database for future knowledge exploration (as shown in Figure 13). The SSMs will nourish the SSM database and will help students to efficiently explore the related knowledge.

SSM searching component

The design of the SSM searching component is shown in Figure 14. The searching process of the SSM searching component is identical to that of the Q&A searching component except that the knowledge sources are different (as shown in Figures 15-16). The knowledge source of the SSM searching component is the SSM database rather than the Q&A database.
While the student searches for required knowledge through the proposed system, the student’s learning history, including browsed knowledge materials, start time, end time, etc. will be automatically recorded in the learning history database.

The proposed system is developed on the Windows platform using Microsoft Visual Basic and an SQL Server: The user interface and knowledge exploration modules were developed using Microsoft Visual Basic; the databases were stored in the Microsoft SQL Server.

**Field experiment**

In order to test whether the proposed system is helpful in terms of improving students’ learning achievement and learning efficiency when IBL is conducted, a field experiment was performed on the freshman course “Computer Science.” This study explored the effects of the proposed system on students’ learning achievement and learning efficiency with the various supports of the proposed system.

**Experiment design**

The experimental procedure in this study consisted of the following steps of a quasi-experimental design:
Step 1: Design the IBL activities

- Concept map construction: A concept map is a graphical representation of a network where each node represents a concept of the learning domain, and links represent relationships between concepts (Novak, 1998; Tseng, Sue, Su, Weng, & Tsai, 2007). In this study, a concept map, constructed by the teacher according to his/her experience of giving lessons, is used to make sure all the guided problems and supplemental materials are sufficient to cover the learning scope of the IBL activities.

- Guided problems design: Guided problems are designed based on the knowledge concepts within the concept map. Students are guided to participate in the learning activities, including analyzing problems and collecting related knowledge materials by knowledge exploration.

- Supplemental materials uploading: Well-prepared supplemental materials are uploaded through the SSM management component before the IBL activities are conducted. Each of the uploaded supplemental materials is divided into topical segments and saved in the SSM database.

Step 2: Conducting learning activities

A tutorial for the system interfaces is provided before the learning activities are conducted. A concept map is then provided to the students to help them conduct the following two IBL activities:

- In-class learning: Some guided problems are given to the students, and the IBL activity is conducted in-class for 50 minutes.

- After-class learning: After the in-class learning is finished, more guided problems are given to the students for conducting the IBL activity after-class at their convenience within one week.

Step 3: Evaluation

- In-class learning achievement evaluation: After the in-class IBL activity, the students were asked to complete an in-class test, including problems designed based on the same knowledge concepts as that of the guided problems, in order to evaluate their learning achievement. That is, students’ learning achievement was measured by the test score.

- After-class learning achievement evaluation: After the after-class IBL activity, the students took an after-class test, including problems designed based on the same knowledge concepts as that of the guided problems, to evaluate their learning achievement. Identical to the in-class learning achievement evaluation, the students’ learning achievement was measured by their test scores.

- After-class learning efficiency evaluation: The time the students spend on finishing the after-class IBL activity could differ. Hence, students’ learning time is used for evaluating their learning efficiency. Students’ learning efficiency was measured by the ratio of achievement (test score) to study time.

Before the IBL activities, the learning and test scope was estimated by the teacher, and the learning activity was tested by a pilot run to make sure that students could definitely finish learning within the predefined learning time with only basic assistance (i.e., only the web searching). Hence, it is inferred that, after learning, the three groups of students had the ability to operate the system and learned the concepts incorporated in the guided problems. The major difference among the three groups could be due to the ways of organizing the learning concepts and the mastering of the concepts and skills. That is, providing more advanced modules might enable the students to learn in a more efficient way, so that they had more time to master the concepts and skills; moreover, the better epistemic scaffolds could benefit the students more in terms of understanding the relationships between the concepts, and hence they could gain more from the learning activity.
Participants and the experimental process

In this study, 90 first year undergraduate students from three classes, 30 students in each class, were invited to participate in the experiment. The participants were divided into three groups by class: a control group, experimental group 1, and experimental group 2. During the experiment, the main difference between the three groups was the number of knowledge exploration modules provided. In experimental group 2, full system support was provided to the students, including all three modules for exploring knowledge from the Q&A database, the SSM database, and the web pages. In experimental group 1, both the modules for exploring knowledge from the Q&A database and the web pages were provided, while in the control group, only the Web module which enables knowledge exploration from the web pages was provided.

In this study, the course “Computer Science” was used as the content subject. The learning activities were focused on the unit “The Internet,” and were conducted after the mid-term examination. Before the beginning of the learning activities, the concept map, guided problems, and supplemental materials were designed and prepared by the teacher. For in-class learning, 14 knowledge concepts within the concept map and 20 guided problems were provided to the students. For after-class learning, 12 more knowledge concepts within the concept map and 35 more guided problems were provided to the students after the in-class learning had been completed.

The whole process of the experiment is shown as Figure 17. At first, the three groups of students were taught to operate the proposed system (15 minutes for each group). Then the in-class learning activity started and lasted for 50 minutes. After the students finished the in-class learning activity, post-test 1 (i.e., the in-class learning achievement evaluation) was conducted for 25 minutes. Subsequently, the after-class learning began and took one week for each group of students. After the after-class learning activity had been completed, the students were asked to take post-test 2 (i.e., the after-class learning achievement evaluation) for 25 minutes.

The learning scenarios of the field experiment are shown as Figure 18. After the whole process of learning was completed, the learning achievement and learning efficiency of the three groups of students were analyzed and compared based on the scores of the mid-term examination (i.e., the pre-test), post-test 1, and post-test 2, as well as the learning time of the after-class learning.
Investigation targets and research hypotheses

In addition to web searching, the proposed system provides two advanced knowledge exploration modules (i.e., the Q&A module and the SSM module) to further support inquiry-based learning activities. To assist students in solving their questions during IBL activities, related historical questions and segmented supplemental materials are provided to students for reference. Although providing students with more information may be beneficial to their learning, it may lead to the problem of information overload (Chen, Kinshuk, Wei, & Chen, 2008; Huang, Liu, Lee, & Huang, 2012). Information overload refers to the situation in which students are confronted with a mass of information and need to decide whether certain information about a topic is worth retaining (Chen, Kinshuk, Wei, & Chen, 2008; Huang, Liu, Lee, & Huang, 2012). Information overload is a main hindering factor for efficient and effective learning in learning and teaching scenarios (Wolpers, Najjar, Verbert, & Duval, 2007).

To explore the effects of the proposed system on students’ learning achievement and learning efficiency, three targets were investigated: The first concerns the students’ in-class learning achievement, the second concerns their after-class learning achievement, while the third concerns their after-class learning efficiency. Since the time spent on learning was the same for each student in the in-class learning, in-class learning efficiency was not investigated in our study. It was expected that providing advanced knowledge exploration modules would improve both the students’ learning achievement and their learning efficiency, and that the more advanced modules that were provided, the greater the benefit that would be obtained. Hence, the research hypotheses were defined as follows:

Investigation target 1: In-class learning achievement

H1.1: In in-class learning, students in experimental group 1 have better learning achievement than those in the control group.
H1.2: In in-class learning, students in experimental group 2 have better learning achievement than those in the control group.
H1.3: In in-class learning, students in experimental group 2 have better learning achievement than those in experimental group 1.

Investigation target 2: After-class learning achievement

H2.1: In after-class learning, students in experimental group 1 have better learning achievement than those in the control group.
H2.2: In after-class learning, students in experimental group 2 have better learning achievement than those in the control group.
H2.3: In after-class learning, students in experimental group 2 have better learning achievement than those in experimental group 1.
Investigation target 3: After-class learning efficiency

If the validation of H2.1 results in no significant difference, a hypothesis (H3.1) is proposed: In after-class learning, students in experimental group 1 spent less learning time than those in the control group.

If the validation of H2.2 results in no significant difference, a hypothesis (H3.2) is proposed: In after-class learning, students in experimental group 2 spent less learning time than those in the control group.

If the validation of H2.3 results in no significant difference, a hypothesis (H3.3) is proposed: In after-class learning, students in experimental group 2 spent less learning time than those in experimental group 1.

Regarding investigation target 3, the after-class learning efficiency of any two groups can be compared only when their after-class learning achievements have no significant difference. In such a case, those students who spent less learning time have better learning efficiency. Hence, in investigation target 3, each of the hypotheses (i.e., H3.1, H3.2, and H3.3) is proposed only when there is no significant difference in after-class learning achievement between students of the two considered groups.

Instruments

In this study, one pre-test and two post-tests (i.e., post-test 1 and post-test 2) were adopted as the tools to evaluate the students’ learning achievement. Also, students’ learning times, which were recorded in their learning history, were adopted to evaluate the students’ learning efficiency. Both the pre-test and post-tests were designed by the teacher and were conducted before and after the learning activities, respectively. The pre-test was the mid-term examination of the course “Computer Science.” It consisted of 24 single-choice items, 4 multiple-choice items, and 7 question-and-answer items with a full score of 100. It was used to check whether the three groups of students had equivalent basic knowledge for the subsequent learning activities. Post-test 1 consisted of 16 single-choice items and 5 question-and-answer items with a full score of 100, while post-test 2 consisted of 17 single-choice items and 2 question-and-answer items with a full score of 100.

Methods

Before the students’ learning achievement and learning efficiency were evaluated, the results of the pre-test were analyzed to confirm whether the three groups of students had equivalent prior knowledge for taking the experiment. ANOVA (Analysis of Variance) was adopted to evaluate whether the three groups of students had significant differences in their pre-test scores. According to the ANOVA result, significant difference ($F = 14.75, p < 0.001$) between the three groups of students was found. The results indicate that the three groups had an unequal level of prior knowledge for taking the experiment.

As stated in the previous literature (Kuehl, 1999; Mitnik, Recabarren, Nussbaum, & Soto, 2009), ANCOVA (Analysis of Covariance) allows the comparison of treatments that show differences in their covariate pre-test values. In order to eliminate the effect of the non-equivalent prior knowledge, ANCOVA was performed to analyze the results of post-test 1 and post-test 2, by using the pre-test scores as a covariate, for evaluations of in-class learning achievement and after-class learning achievement, respectively.

Before ANCOVA, the homogeneity of variance assumption and the homogeneity of regression coefficients were tested. For in-class learning, the Levene’s test for equality of variances was not significant ($F = 0.86, p > 0.05$). Hence, the homogeneity of variance assumption was not violated. Also, the result ($F = 0.75, p > 0.05$) indicated that the assumption of homogeneity of regression coefficients was not violated. Similar results were obtained in after-class learning. The Levene’s test for equality of variances was not significant ($F = 0.51, p > 0.05$), which indicated the homogeneity of variance assumption was not violated. Also, the results ($F = 0.71, p > 0.05$) indicated that the assumption of homogeneity of regression coefficients was not violated.
**Results and discussion**

**Evaluation of in-class learning achievement**

The results of the ANCOVA are given in Table 1 and show that there is significant difference \( (F = 11.61, p < 0.001) \) among the learning achievements of the three groups of students.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>Std. err</th>
<th>( F(2, 87) )</th>
<th>Pairwise comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-test 1</td>
<td>Experimental group 2 (A)</td>
<td>30</td>
<td>64.07</td>
<td>9.71</td>
<td>63.87</td>
<td>2.00</td>
<td>11.61***</td>
<td>(A) &gt; (B)</td>
</tr>
<tr>
<td></td>
<td>Experimental group 1 (B)</td>
<td>30</td>
<td>57.53</td>
<td>8.44</td>
<td>57.59</td>
<td>1.83</td>
<td></td>
<td>(A) &gt; (C)</td>
</tr>
<tr>
<td></td>
<td>Control group (C)</td>
<td>30</td>
<td>49.73</td>
<td>11.25</td>
<td>49.88</td>
<td>1.91</td>
<td></td>
<td>(B) &gt; (C)</td>
</tr>
</tbody>
</table>

*Note.*** \( p < .001.\)

This indicates that the number of knowledge exploration modules provided affected the students’ learning achievement. According to the post-hoc analysis, the learning achievement of experimental group 2 (adjusted mean = 63.87) was the best. Moreover, the learning achievement of experimental group 1 (adjusted mean = 57.59) is better than that of the control group (adjusted mean = 49.88). Hence, the research hypotheses H1.1, H1.2, and H1.3 were supported.

**Evaluation of after-class learning achievement**

The results of the ANCOVA are given in Table 2 and show that there is significant difference \( (F = 8.51, p < 0.001) \) among the learning achievements of the three groups of students.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>Std. err</th>
<th>( F(2, 87) )</th>
<th>Pairwise comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-test 2</td>
<td>Experimental group 2 (A)</td>
<td>30</td>
<td>58.10</td>
<td>14.66</td>
<td>55.35</td>
<td>2.58</td>
<td>8.51***</td>
<td>(A) &gt; (B)</td>
</tr>
<tr>
<td></td>
<td>Experimental group 1 (B)</td>
<td>30</td>
<td>40.83</td>
<td>10.44</td>
<td>41.64</td>
<td>2.37</td>
<td></td>
<td>(A) &gt; (C)</td>
</tr>
<tr>
<td></td>
<td>Control group (C)</td>
<td>30</td>
<td>40.00</td>
<td>14.20</td>
<td>41.95</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.*** \( p < .001.\)

According to the post-hoc analysis of post-test 2, it shows that the learning achievement of experimental group 2 (adjusted mean = 55.35) is better than that of the control group (adjusted mean = 41.95) and the learning achievement of experimental group 2 is also better than that of experimental group 1 (adjusted mean = 41.64). However, there is no significant difference between the learning achievements of experiment group 1 and the control group. Hence, the research hypotheses H2.2 and H2.3 were supported, while hypothesis H2.1 was not.

**Evaluation of after-class learning efficiency**

As described in the Experiment Design and Investigation Targets and Research Hypotheses sections, the time students spent during their after-class IBL activity could differ. If there is no significant difference between the after-class learning achievements of two considered groups of students, learning efficiencies can be further investigated by analyzing the times the students spent on after-class learning. As shown in Table 2, the after-class learning achievements of the control group and experimental group 1 did not differ significantly, and thus the research hypothesis H2.1 is not supported. Hence, the research hypothesis H3.1 should be tested. However, since hypotheses H2.2 and H2.3 are supported, H3.2 and H3.3 are not tested in this subsection.

To further explore the students’ learning efficiency, a \( t \)-test was performed to evaluate whether there was a significant difference in the after-class learning time between the control group and experimental group 1. Firstly, a \( t \)-test was performed to evaluate whether the control group and experimental group 1 had a significant difference in the pre-test scores. The results show that the mean and standard deviation of the pre-test are 46.87 and 14.99 for experimental
group 1, and 41.90 and 15.80 for the control group. From the t-test results \((t = 1.25, p > 0.05)\), these two groups did not differ significantly. Based on this statistical evidence (i.e., they did not differ by more than chance), we assumed that the two groups had similar prior knowledge.

Secondly, the after-class learning time of the two groups of students was analyzed by t-test. From Table 3, the t-test results \((t = -2.12, p < 0.05)\) showed that the two groups had a significant difference in after-class learning time. In the control group, the students spent much more time (33.50 minutes in average) on the learning activity than experimental group 1 did (7.67 minutes in average). However, the learning achievement of the two groups, as shown in Table 2, is not significantly different. Therefore, the learning efficiency of experimental group 1 was better than that of the control group. That is, research hypothesis H3.1 was supported.

<table>
<thead>
<tr>
<th>Group</th>
<th>(N)</th>
<th>Mean (minute)</th>
<th>(SD)</th>
<th>(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group 1</td>
<td>30</td>
<td>7.67</td>
<td>21.51</td>
<td>-2.12*</td>
</tr>
<tr>
<td>Control group</td>
<td>30</td>
<td>33.50</td>
<td>63.08</td>
<td></td>
</tr>
</tbody>
</table>

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

The main reason why experimental group 1 had better learning efficiency than the control group might be because of the Q&A module that was made available to experimental group 1. The analytical results show that the students of experimental group 1, in which both the Q&A module and Web module were provided, reach the same learning achievement as those of the control group, in which only the Web module is provided, in a much shorter time. That is, the students learned more efficiently with both the Q&A and Web modules than those students who learned with only the Web module.

**Discussion and conclusions**

As described in the Literature Review section, the guidance regarding core concepts and a more constructivist epistemological framework are beneficial to students’ learning (Dufresne, Gerace, Hardiman, & Mestre, 1992; Linn & Songer, 1993; Tobin, Tippins, & Hook, 1995; Windschitl & Andre, 1998; Sandoval & Reiser, 2004; Nelson, 2013). With this in mind, an Inquiry-Based Learning (IBL) support system was proposed and developed in this study. In addition to web searching, which is readily available through search engines, the proposed system provides two advanced knowledge exploration modules: the Q&A module and the SSM module, to further support inquiry-based learning activities.

Although more knowledge, which can also be obtained from the Web, was provided to the students for reference, the proposed system eliminates the need for students to screen out irrelevant knowledge from a huge number of web pages by themselves. Hence, the problem of information overload can be avoided and the benefit of the advanced knowledge exploration modules can be obtained. The experimental results show that providing the two advanced knowledge exploration modules is helpful to students for improving both their learning achievement and their learning efficiency, revealing the effectiveness of conducting IBL with the proposed system.

It is also found that, for in-class learning, providing advanced knowledge exploration modules, either only the Q&A module or both the Q&A and the SSM modules, improves students’ learning achievement. Moreover, the more knowledge exploration modules that are provided, the better the learning achievement that is achieved. Such results could be attributed to the provision of well-organized materials that meet the need of students, as indicated by several previous studies (Denton, Madden, Roberts, & Rowe, 2008; Wu, Hwang, Milrad, Ke, & Huang, 2012). In comparison with those poorly structured materials that might contain inconsistent information, the advanced knowledge exploration modules provide an epistemological framework which supports the students to accurately and efficiently resolve their learning problems.

On the other hand, for after-class learning, the use of the SSM module improves students’ learning achievement, implying that the selected segments provided by the SSM module meet the students’ requirements. Moreover, it is found that the Q&A module improves the students’ learning efficiency for after-class learning. It is inferred that the answers provided by the teacher through the Q&A module can shorten the time needed for the students to acquire the knowledge for solving their problems. Such a benefit of the Q&A module can be more apparently observed in after-
class learning, which allows students to have more flexible time to learn. To sum up, the two advanced knowledge exploration modules, the Q&A module and the SSM module, proposed in this paper are helpful for students in their inquiry-based learning process.

Although the results revealed that the proposed system is beneficial to students’ learning, there still exists a limitation in the experiment design. The lack of random assignment of participants may limit the internal validity of the design, and thus limit the strength of causal claims that can be made. In addition to the statistical evidence, there were several reasons for us to assume that the two groups of students had similar prior knowledge before the experiment. First, they took the same entrance exam and got similar scores before entering the university; second, the students in Taiwan took the same computer courses in elementary and high school; third, the students had taken the same computer courses taught by the same teachers after entering the university.

In the future, we will apply more text mining techniques, such as synonym identification and personalized searching, to further enhance the knowledge exploration functions of the proposed system. Moreover, we will promote the use of our proposed system in inquiry-based learning, and carry out a broader investigation of the feedback of both teachers and students regarding the use of the modules.

References


A Role-Play Game to Facilitate the Development of Students’ Reflective Internet Skills

Wilfried Admiraal
Leiden University Graduate School of Teaching, Leiden University, Leiden, The Netherlands // w.f.admiraal@iclon.leidenuniv.nl

(Submitted October 4, 2013; Revised May 27, 2014; Accepted November 18, 2014)

ABSTRACT

Although adolescents are currently the most frequent users of the Internet, many youngsters still have difficulties with a critical, reflective, and responsible use of the Internet. A study was carried out on teaching with a digital role-play game to increase students’ reflective Internet skills. In this game, students had to promote a fictional celebrity. The six game levels cumulated in complexity and scope, and students played different roles in each of them (e.g., manager, marketing manager, and journalist). In a one-group pre-test/post-test design, the implementation of the role-play game in four secondary school classes was evaluated using a questionnaire on students’ reflective Internet skills. On both pre-test and post-test, girls generally out-performed boys on reflective Internet skills. Repeated measures analyses showed that this gender difference was significantly reduced by teaching with the role-play game. Boys significantly increased their reflective Internet skills, but girls did not. Implications for teaching with this role-play game are presented along with indications how the role-play game could be redesigned to be effective for all students.

Keywords

Media literacy, Reflective internet skills, Role-play game, Secondary education, Gender differences

Introduction

Even though they are members of a digital generation, many youngsters are digitally illiterate because they still have great difficulties with critical, reflective, and responsible use of the Internet (e.g., van Deursen & van Diepen, 2013; Kuiper, Volman, & Terwel, 2005; Squire, Devane, & Durga, 2008; Walraven, Brand-Gruwel, & Boshuizen, 2009). Students need to develop their ability to determine the accuracy of online resources and the credibility of other users, to know their online audience and how to present themselves accordingly, and to understand social norms and online communication rules (cf. Litt, 2013). In the curricula of Dutch primary and secondary education, little attention is paid to the development of this kind of reflective Internet skills of students. Students have to develop these skills implicitly, during their free time or school time while completing school assignments. Teaching students these reflective Internet skills more explicitly might facilitate the development of these skills. In this paper, we present the results of an evaluation of teaching with a role-play game to promote students’ reflective Internet skills.

Reflective internet skills

With the rapid development of information and communication technologies and online communication technologies in particular, adolescents are currently the most frequent users of the Internet. They spend more time online than adults do, and they use the Internet more for social interaction. Although studies from the 1990s showed that adolescents’ online communication increased their online relationships with strangers (e.g., Nie, 2001), recent studies provide evidence that adolescents use the Internet more to maintain their existing friendships, become less concerned how others perceive them, and feel fewer inhibitions in disclosing intimate information (cf. Valkenburg & Peter, 2009). Obviously, when media use changes, its outcomes may change. Using the Internet for more and more intimate communication demands more and different Internet skills of the user.

In a recent review of measurements of Internet skills, Litt (2013) concluded that most studies focused on information and technical skills and that more attention is needed to study communicative and socio-emotional skills, such as the ability to determine the credibility of other users, to know one’s online audience and how to present oneself accordingly, and to understand the social norms of sites and services. Various concepts have been used to describe these reflective Internet skills, of which media literacy is widely used in Dutch primary and secondary education. Buckingham, Banaji, Burn, Carr, Cranmer, & Willett (2005) defines media literacy as “the ability to access,
understand and create communications in a variety of contexts” (p. 9). This definition implies that media-literate students know what information is available and how they can retrieve it; use and process information in a wide range of ways; make judgments about the authenticity, honesty, relevance, or accuracy of information; and create new products, and present, share and transfer information in suitable forms. The latter is also explored by Litt (2012) as the notion of an “imagined audience”— a person’s mental conceptualization of the people with whom he or she is communicating.

Studies on Internet skills of secondary school students in the Netherlands (van Deursen & van Diepen, 2013; van Deursen, van Dijk, & Peters, 2011; Kuiper, Volman, & Terwel, 2009; Walraven et al., 2009) show that students possess technical skills to use the Internet, but that they have problems with, among other things, accurate search queries, critical evaluation of Internet sources, and awareness of their online audience. Many students appeared to be inconsistent, impulsive, and impatient Internet users. These findings align with conclusions about reflective Internet skills of secondary school students in other countries (e.g., Ba, Tally, & Tsikalas, 2002; Gui & Argentin, 2011; Kim & Lee, 2013). Apparently, frequently use of the Internet does not automatically increase a reflective use of the Internet. Adolescents need more guidance on how to use the Internet in a more reflective way, either in a schooling context or during pastime.

More explicit attention for how to use the Internet in secondary schools might facilitate the development of reflective Internet skills of students. However, Internet skills are not a standard component of the Dutch curriculum of secondary schools and therefore few attempts are made in schools to improve them. These reflective Internet skills are not explicitly taught because teachers are not familiar with the concept of media literacy (Moore, 2002) or teaching reflective Internet skills does not fit easily in the school curriculum (Meabon Bartow, 2014; Walton & Archer, 2004). Teaching reflective Internet skills is not about shielding young people from the influence of the media but about enabling them to make informed decisions and to develop their understanding of and participation in the media that surrounds them (Buckingham et al., 2005; Buckingham, 2007). In teaching reflective Internet skills, students need to define their information needs, formulate questions and know how to locate, evaluate and use information from many online sources. Probert (2009) argued that they need to be aware also of issues relating to the ethical use of information, such as copyright and plagiarism.

From earlier research (e.g., Coiro, 2003; Dalton & Proctor, 2008; Kuiper et al., 2009; Walraven et al., 2009), the following design principles can be derived to support teaching reflective Internet skills in secondary schools: (1) the use of authentic tasks (tasks that are related to students’ everyday life; (2) teaching with group discussions to overcome school practice in which students work quite isolated behind their computer and students’ tendency to view the Internet as offering ready-made facts and figures; and (3) support for teachers on how to connect with the Internet experiences of students. Internet games form one such authentic environment for students to practice and develop their reflective Internet skills.

**Teaching with games**

Teaching with games can form an effective learning environment for teaching reflective Internet skills. Literature reviews on game-based learning showed both short-term and long-term cognitive effects for students (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Perrotta, Featherstone, Aston, & Houghton, 2013; Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013; Wouters & van Oostendorp, 2013). In a meta-analysis on the role of instructional support in game-based learning, Wouters and van Oostendorp (2013) analyzed findings of 29 studies on effects of game-based learning on knowledge and cognitive skills. These authors distinguished various types of instructional support of games which were found to be effective in terms of increasing knowledge and cognitive skills: (1) games in which ideas, characters, topics, and messages could be personalized (2) games with oral instead of written explanations and (3) games with corrective feedback, modeling problem-solving processes and stimulating students’ reflection and collaborative work. Neither narrative elements nor the interactivity of a game were sufficient to explain differences in students’ knowledge gain and growth of cognitive skills.

Recently, three review studies were published on the use of games in teaching and learning. Firstly, Wu, Hsiao, Wu, Lin, and Huang (2012) published a meta-analysis of studies on game-based learning from the period 1971–2009. The focus of their analysis of the 91 selected studies was how the studies on game-based learning were explicitly informed by learning theories. One of their conclusions is that only a few studies were informed by learning theories.
and the ones that were based their learning principles underlying teaching with games on constructivist ideas and experiential learning theory. The latter finding aligns with the types of instructional support that Wouters and van Oostendorp (2013) found triggered cognitive effects of game-based learning.

Secondly, Connolly et al. (2012) reviewed literature on teaching and learning with either serious digital games or digital games for entertainment. They examined 129 studies, from the period 2004–2008, distinguishing different game genres. Simulations and puzzle games were mostly used for learning, and action games, role-play games, and generic games were mostly used for entertainment. Relatively few papers in the review were classified as serious games (games used for education).

Thirdly, Wouters et al. (2013) conducted a meta-analysis of cognitive and motivational effects of digital serious games, with an extensive overlap of studies from the meta-analysis on instructional support by Wouters and van Oostendorp (2013). Although the period they studied covered 1990–2012, most studies were published more recently (2007–2012). On the basis of 39 studies, the authors conclude that the use of serious games was effective in terms of learning (knowledge and cognitive skills) and retention, but it was not more motivating. These effects were found to be stronger when the game was used in multiple sessions rather than in one, when it was played in group rather than individually, and when it was supported by other instructional methods that preferably mixed active and passive learning.

The literature mentioned above suggests that teaching with games has the potential to increase students’ knowledge and improve their learning activities. Particularly, the meta-analysis of Wouters and van Oostendorp (2013) recommend that the use of games be accompanied by instructional support and a variety of pedagogical activities. However, the use of games in teaching might not be equally effective for all students, depending on their learning and technology preference. Although differences between boys and girls in time spent playing games seem to decrease (Homer, Hayward, Frye, & Plass, 2012), digital game-playing has emerged as a predominately male pastime (cf. Bryce & Rutter, 2003). Boys not only spend more time playing games than girls, they also prefer to play different kinds of games. Boys tend to prefer action games (first-person shooters, fighting, and sports games), while girls show a preference for playing simulations (virtual-world and virtual-life games) and puzzle games. No gender differences are found with most other genres, such as role-play games and adventure games (Bonanno & Kommers, 2005; Hamlen, 2011; Homer et al., 2012).

These differences in preferences for games and gaming between boys and girls seem to be related to a different motivation to play games. Boys tend to be more motivated to outdo other players (performance-only or performance combined with mastery-achievement goals), while girls are less motivated by such goals. Gender is strongly correlated with gaming frequency: very few girls were frequent gamers and few boys were non-gamers (Heeter, Lee, Medler, & Magerko, 2011). Therefore, findings on gender differences in effects of teaching with games are mixed and depend on game genre and the way it is used in class (Bonanno & Kommers, 2005; Hamlen, 2011; Homer et al., 2012).

A study will be reported on the evaluation of teaching with a role-play game to promote reflective Internet skills of lower grades secondary school students. In this study, competition between pairs was a major game activity in teaching with the role-play game. Therefore, two hypotheses were formulated about the effects of the use of the role-play game:

- Teaching with the role-play game increases students’ reflective Internet skills
- Boys increase their reflective Internet skills more than girls

**A role-play game on reflective internet skills**

A digital role-play game (SplitsZ!) was set up to develop students’ reflective Internet skills. An online workspace was used, similar to Habbo Hotel (http://www.habbo.com/). The mission of the game was to promote a fictional celebrity using towers with billboards and monitors (see Figure 1 for a screenshot of one of the towers). Each tower was owned by a small group of students (pairs or triads created by the teacher to be mixed-gender groups) who constructed their towers with clips, images, slogans, other textual sources. In each of the six levels, these pairs of students played different roles (e.g., manager, marketing manager, and journalist), which were represented by
avatars. The more popular a celebrity was, the higher the popularity scores of that group. Students were not required to play the game at home or work on debates and reflections outside the lessons.

Figure 1. Screenshot of one of the towers in SplitsZ!

Lesson series

The intervention included a series of 12 to 18 lessons (in six to eight weeks) in which the six levels of the game were played. The assignments in each level were prepared by the teacher and completed by students in class. Each level was played within one lesson. In the next one or two lessons, students reflected on the assignments of the particular game level, and the teacher organized a class debate. After the teachers briefly introduced the purpose of the game, assignments, and game activities, students played the game in pairs or triads, which were created by the teachers. The role of the teachers was limited to a general introduction, supporting students with technical problems and elaborating on what the requirements of the assignments were. All content was included in the game, polls, and game assignments.

Game levels

The six levels of the game increased in complexity and scope. In level 1, students started to develop a celebrity profile (layout, character, and characteristics) and completed assignments on the influence of online media on how a person is perceived. In level 2, students had to decide how to introduce their celebrity in the virtual world and completed assignments on online branding, advertising, and social networking. Students used Hyves to promote their celebrity. In level 3, students took the role of a media reporter in order to evaluate the celebrities of other students in
public and completed assignments on hypes, trends, privacy, and serious editing. In level 4, students produced a celebrity performance (clips, slideshow, etc.) and reflected on famous performances and advertisements. In level 5, students reworked the media campaign of level 2 in order to generate hype around their celebrity. Finally, in level 6, students took the role of journalist and wrote in pairs or triads a critical review about one of the campaigns of their peers, which were set up in level 5.

Assignments

The game assignments and polls aimed to trigger the students’ awareness of how to present themselves and others online, how to assess online sources critically, and what the influence of the Internet can be on their behavior and on society. Completed assignments were included in a web journal, which students presented at the end of each level (third lesson). Teachers used 20–24 polls (four per level) to test students’ reflective Internet skills. These polls asked students whether they did or did not agree with several statements about a particular media campaign, Internet photo series, or some Internet gossip. Examples of polls are about a photo series of Sarkozy (“what has been photo shopped?”), paintings of Napoleon (“which painting was commissioned by Napoleon?”) or a gossip about a Dutch politician. All polls started with an instruction on what kind of reflective skills were expected. Most students split up roles and activities instead of collaboratively working on the assignments and polls.

Methods

Designs

Using a one-group pre-test/post-test design, we collected additional quantitative data of one class from four secondary schools in the Western part of the Netherlands (with 80 Grade 7 students (of whom 59% were female). These students played in gender-mixed pairs or triads that were determined by the individual school teachers. The data of one boy was removed after outlier analyses.

Data

Students’ self-assessment of their reflective Internet skills was measured with a 15-item questionnaire indicating behavior about and attitude towards consciously downloading information (images, clips, and text) from the Internet (seven items, Cronbach’s $\alpha = 0.71$) and uploading information on the Internet (eight items, Cronbach’s $\alpha = 0.73$). This questionnaire was developed specifically for this study and was based on the concept of digital judgment and the ability to acquire, process, and produce digital information (Hatlevik & Christophersen, 2013). Most items were formulated in statements and students had to indicate how applicable the statements were on their own situation (1 = not at all to 5 = to a large extent). An example item for downloading information is “I check the accuracy of Internet sources I use for homework assignments” and for uploading information: “I think carefully about who sees my profile on the Internet.” The students also reported their daily use of technology, such as social software, communication software, games, and virtual worlds (all items on a five-point scale with 1 = low frequency and 5 = high frequency), as well as their gender and age.

Analyses

In order to examine the effects on students’ reflective skills, repeated measures analyses were used with the pre- and post-test scores on both indicators of reflective Internet skills as within-subjects variables and gender as between-subjects factor. In subsequent analyses, students’ age and technology use were added as covariates. Analyses were performed on student level as pairs and triads were all mixed-gender groups and students were allowed to switch groups. As the hypotheses were one-sided, a significance level of $\alpha = .10$ was used.
Results

The results of the repeated measures analyses are summarized in Table 1. For reflective Internet skills, in terms of consciously downloading information from the Internet, we found a significant effect of time and of time by gender. Students generally scored higher on the post-test than on the pre-test ($F(1,67) = 3.28; p = .075; \eta^2 = .05$), with an interaction effect for gender ($F(1,67) = 3.82; p = .055; \eta^2 = .05$). Boys ($t(29) = −2.07; p = 0.05$) and not girls ($t(38) = 0.14; p = 0.89$) showed a small increase in their scores on consciously downloading information. We found a similar, yet stronger, outcome with consciously uploading information to the Internet. Students had increased significantly at the post-test ($F(1,67) = 6.79; p = .011; \eta^2 = .09$), with an interaction effect for gender ($F(1,67) = 5.04; p = .028; \eta^2 = .07$). Boys increased their scores significantly ($t(29) = −2.90; p = 0.01$), and girls did not, staying at their high level of the pre-test ($t(38) = 0.30; p = 0.76$).

Table 1. Repeated measures analyses with gender for downloading and uploading online information

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th></th>
<th>Post-test</th>
<th></th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Downloading information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>3.10</td>
<td>1.04</td>
<td>3.45</td>
<td>0.72</td>
<td>30</td>
</tr>
<tr>
<td>Girls</td>
<td>3.86</td>
<td>0.66</td>
<td>3.84</td>
<td>0.75</td>
<td>39</td>
</tr>
<tr>
<td>All</td>
<td>3.53</td>
<td>0.92</td>
<td>3.67</td>
<td>0.76</td>
<td>69</td>
</tr>
<tr>
<td>Uploading information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>2.36</td>
<td>0.69</td>
<td>2.82</td>
<td>0.77</td>
<td>30</td>
</tr>
<tr>
<td>Girls</td>
<td>3.25</td>
<td>0.72</td>
<td>3.28</td>
<td>0.78</td>
<td>39</td>
</tr>
<tr>
<td>All</td>
<td>2.86</td>
<td>0.83</td>
<td>3.08</td>
<td>0.80</td>
<td>69</td>
</tr>
</tbody>
</table>

After controlling for students’ ages and their daily use of technology, both interaction effects with gender were somewhat stronger (with $\eta^2 = .08$ for downloading information and $\eta^2 = .11$ for uploading information). So, we found a medium-sized positive effect of game-based learning on reflective Internet skills of boys. The differences between boys and girls in downloading and uploading information from the Internet are significant, even at the post-test ($t(1,67) = 2.20; p = .02$ for downloading and ($t(1,67) = 2.20; p = .02$ for uploading information). This means that, although we found an effect of the role-play game on reflective skills of boys, boys still scored significantly lower than girls after the game was played.

Discussion and conclusion

A significant gender effect was found of the digital role-play game, which meant that boys, and not girls, increased their reflective Internet skills, as indicated by their awareness of Internet use for both downloading information and uploading information to the Internet. However, even after playing the game, boys still scored significantly lower on both measures than girls.

An implication of this study might be to play the role-play game individually in class and in pairs or small groups outside class. Although playing in pairs was meant to trigger discussion and interaction behind the computer, students split up tasks and activities instead. The role-play is designed to be played with collaboration and competition as a part of the game, not in front of the computer. This would also imply that the game can be (partly) played at home, which would also offer better possibilities for including parents in game activities and in the communication about the game. Parent tutoring appeared to be a fruitful way to create additional learning opportunities for children (Daly III & Kupzyk, 2012; Erion, 2006).

Most technology issues (e.g., slow-moving avatars, limited server space, and unclear error messages) were solved in the next release of the game. We still need to make the game more gender-inclusive. Although the competitive part, which is understood to attract boys more than girls, was not fully implemented because most students did not find parts of the software on how they could increase their popularity scores, boys increased their reflective skills more than girls did. In an earlier study on a secret trail game (Admiraal, Huizenga, Akkerman, & ten Dam, 2011), we found that competitive game activities positively mediated the learning outcomes for boys and, to a lesser extent, for girls. However, in contrast to the current study, both boys and girls showed positive learning outcomes. This might
be caused by the fact that other game activities (i.e., collecting data and navigating through the game) mediated gender effects on learning outcomes as well. The findings of the literature on gender-inclusive design of games (e.g., Dickey, 2006; Kafai, 1994; van Eck, 2006) can provide insight into how to design a gender-inclusive game, although a teacher can reinforce or weaken gender-inclusiveness by stimulating, adding, or deleting particular game activities. Games do not teach students. What is important is how games are used in class.

Keeping in mind that our conclusions on gender differences in game-based learning were limited because they were based on the investigation of a particular educational game, we would like to close with implications for game-based learning and gender-inclusive game design. In line with the viewpoint of Gansmo (2009) about a gender-inclusive use of technology in education, we would advocate to see game-based learning as a variety of technologies and approaches. Using multiple hardware and software as well as varying themes, characters and layout of games could make game-based learning a fruitful experience for both boys and girls. The use of variety and multiplicity also holds for how games are used in education. Assignments that solely ask for activity, strategy, or creativity might benefit either boys or girls, whereas problem-solving tasks or game tasks that require student exploration with various student cognitive activities can create meaningful learning opportunities for both boys and girls.

References


Using Mixed-Modality Learning Strategies via e-Learning for Second Language Vocabulary Acquisition

Fang-Chuan Ou Yang1 and Wen-Chi Vivian Wu2*

1Department of Information Management, Chien Hsin University of Science and Technology, Chung-Li, Taiwan // 2Department of English Language, Literature and Linguistics, Providence University, Taichung, Taiwan // ouyang@uch.edu.tw // wcwu@pu.edu.tw

*Corresponding author

(Submitted November 12, 2013; Revised June 2, 2014; Accepted October 13, 2014)

ABSTRACT

This study demonstrated an e-learning system, MyEVA, based on a mixed-modality vocabulary strategy in assisting learners of English as a second language (L2 learners) to improve their vocabulary. To explore the learning effectiveness of MyEVA, the study compared four vocabulary-learning techniques, MyEVA in preference mode, MyEVA in basic mode, an Internet dictionary, and a traditional paper-based dictionary. The findings indicate that the mixed-modality with preference-mode setting stimulated the greatest vocabulary acquisition and the best retention for L2 students. The more proficient students in the study selected a preferred strategy that they employed consistently. Less proficient students did not select a preferred strategy and improved less than their more proficient fellow students. Researchers also provided recommendations for language teachers, curriculum designers, and in particular, the developers of English e-learning systems.

Keywords

Vocabulary-learning strategy, CALL (computer-assisted language learning), Serial learning, Linear learning, Holistic learning, e-Learning

Research background

For students who study English as a second language (L2), vocabulary learning is always a primary concern and plays a key role. Schmitt (2000) claimed that L2 students require 2,000 words to maintain conversations. College students need to know at least 4,000 words to understand English textbooks (Hu & Nation, 2000). However, many college students in Taiwan know fewer than 2,000 words, and therefore, have impediments to comprehending English textbooks (Huang, 2004). Vocabulary knowledge and size influence capability in all four of the language skills, listening, speaking, reading, and writing. Thus, how to develop good learning strategies, methods, and tools is an important issue (Liu et al., 2010; López, 2010).

A vocabulary learning strategy (VLS) should be part of an overall instructional design assisting learners to discover meaning as well as memorize and internalize vocabulary (Nation, 2001). Abundant research shows that a constructivist instructional design emphasizing student-centered active learning is superior to older lecture/memorization models (Brandon & All, 2010; Bruning, Schraw, & Ronning, 1999).

Rubin (1981) said that use of VLSs can help L2 learners memorize and comprehend words. Oxford (1990) concluded that a strong VLS can help learners recall words more effectively. Learners, therefore, should use adaptive learning strategies (Oxford & Crookall, 1990). Many studies have examined how VLSs assist learning and how to select the best VLS for learners of different ages and genders (Fan, 2003; Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999; Schmitt, 1997).

In previous research, however, no VLS has been preferred consistently by learners, and individuals have selected different VLSs based on their learning preferences and feelings about effectiveness (Chacón-Beltrán, Abello-Contesse, & del Mar Torreblanca-López, 2010; Hacquebord & Stellingwerf, 2007; Levine & Reves, 1998; Nation, 2001, 2004; Schmitt, 1997, 2000). Skehan (1991) claimed that learner styles and strategies may be the mediators between learner characteristics and learning outcomes. Tight (2010) examined three conditions (style matching, style mismatching, and mixed modality) to explore inferences about styles and strategies for second-language acquisition. Tight indicated that mixed-modality instruction stimulated the greatest learning and the best retention, which echoed Chun and Plass (1996a, 1996b), who found that words or phrases are remembered better when presented with multiple media.
The authors of the current study believe that Tight’s findings provided insights for development of a VLS and that the effectiveness of mixed-modality vocabulary learning should be further explored. However, there is little academic literature about e-learning vocabulary acquisition strategies. In particular, investigation into mixed-modality EFL vocabulary e-learning is noticeably lacking. This study addressed this gap by exploring the effectiveness of a mixed-modality vocabulary-learning strategy via e-learning for L2 English undergraduates in Taiwan. The significance of this paper is not simply in examination of mixed-modality learning, but more importantly is in the development of affordances of technology that can be used to design specific instructional technology that address specific educational outcome goals.

The researchers developed an integrated vocabulary-learning system named My English Vocabulary Assistant (MyEVA) based on mixed-modality vocabulary strategies. The MyEVA e-learning system was implemented using the Rich Internet Application capability of Adobe AIR (2008) and made available to L2 undergraduates from the Chien Hsin University of Science and Technology in Taiwan. MyEVA included eight VLSs initially based on a strategy classification by Schmitt (1997), adjusted to be suitable for undergraduates in Taiwan. In addition, the authors intended to observe further whether different navigational pathways employed in mixed-modality vocabulary learning would impact the usage preference and learning outcomes. For this reason, MyEVA was designed with two navigational pathways that learners could use while studying vocabulary: (1) basic mode: The system was freely open for students to learn target words and use diverse VLSs, and (2) preference mode: The students were asked to set their preferred learning strategy, and the system showed that strategy by default when they navigated to a word. To evaluate the effectiveness of MyEVA, learning outcomes were observed for four learning techniques: MyEVA in basic mode, MyEVA in preference mode, an Internet dictionary (Yahoo Dictionary), and a traditional paper-based dictionary.

The researchers conducted the study to answer following research questions: (1) Did MyEVA (in either basic mode/preference mode) improve the learners’ vocabulary capability? (2) Which learning technique achieved the best vocabulary learning outcomes for L2 learners with different levels of vocabulary proficiency? (3) Did individuals with different levels of vocabulary proficiency have differences in preferred learning strategies?

### Composition of the mixed-modality vocabulary learning strategies

Many studies have defined the content and classifications of vocabulary learning strategies (Gu, 2003; Gu & Johnson, 1996; Nation, 1990; Nation, 2001; Schmitt, 1997; Schmitt, 2000). The VLSs used in the MyEVA were developed according to Schmitt’s structure (Schmitt, 1997; Schmitt, 2000). Schmitt classified VLSs into two categories: discovery and consolidation. **Discovery** consists of a determination strategy and a social strategy. **Consolidation** comprises a social strategy, a memory strategy, a cognitive/metacognitive strategy, pictures/imagery, related/unrelated words, grouping, and the word’s orthographical and phonological forms, as shown in Table 1.

Schmitt listed a total of 58 different vocabulary-learning strategies derived from the typology in Table 1, ranging from the traditional learning strategies (e.g., looking up vocabulary in a paper-based dictionary) to social strategies (e.g., collaborative learning with friends). Schmitt (2005) also found that the memory and cognitive strategies under the consolidation category had the best learning effectiveness for vocabulary learning.

<table>
<thead>
<tr>
<th>Category of VLS</th>
<th>VLS</th>
<th>Core concept of VLS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discovery</strong></td>
<td>Determination strategy</td>
<td>The way learners understand the definition of new words. For example, guessing from context, using reference materials, or asking someone else.</td>
</tr>
<tr>
<td></td>
<td>Social strategy</td>
<td>Discovering new words by employing the social strategy of asking someone who knows. Teachers are often in this position.</td>
</tr>
<tr>
<td><strong>Consolidation</strong></td>
<td>Social strategy</td>
<td>Developing language capability from social activities. For example, discussion, communication, and cooperation with other people.</td>
</tr>
<tr>
<td></td>
<td>Memory strategy</td>
<td>Storing and recalling information. Skills used in this strategy include induction, classification, correlation, and imagery.</td>
</tr>
<tr>
<td></td>
<td>Cognitive/metacognitive strategies</td>
<td>Learners thinking about their own learning. For example, trial and error, self-evaluation, and self-management.</td>
</tr>
<tr>
<td></td>
<td>Learning new word by studying them with pictures depicting their meanings instead of text definition.</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Related/unrelated words</strong></td>
<td>Linking a word to another word which the learner already knows or has no sense of relationships.</td>
<td></td>
</tr>
<tr>
<td><strong>Grouping</strong></td>
<td>Organizing words into groups by word meaning category.</td>
<td></td>
</tr>
<tr>
<td><strong>Orthographical and phonological form</strong></td>
<td>Focusing on the target word’s orthographical or phonological form to facilitate recall.</td>
<td></td>
</tr>
</tbody>
</table>

To develop the mixed-modality for MyEVA, the researchers selected target VLSs to offer to students according to the following process. First, candidate VLSs were selected from the consolidation category’s learning strategies in Schmitt’s typology. Second, the social strategy was excluded because the intent of the study was to use mixed-modality vocabulary e-learning strategies in self-learning situations. Third, vocabulary learning strategies that could not be implemented in an e-learning system were rejected. One special vocabulary learning strategy not in Schmitt’s list, “Chinese-assonance strategy,” was selected because it is popular in language schools in Taiwan (English cram schools). Ultimately, the following eight VLSs (see Figure 1) were selected for implementation in MyEVA:

**Word-card strategy:** Word cards are a basic VLS, providing complete information on a word, through an e-dictionary or web dictionary. The description includes the word, K.K. phonics, word type (part of speech), word meaning/explanation in the native language, an example sentence, pronunciation sound, and an illustration. Figure 1(a) shows an example of the word-card strategy.

**Flashcard strategy:** Flashcards are a set of cards that include vocabulary information on either or both sides for private study. The learner often writes a word on one side of the card and the word explanation on the reverse side. Learners often use flashcards to memorize words quickly in their spare time. The flashcards in MyEVA were designed as a card-style component and provided simplified explanations for the target word. Figure 1(b) shows an example of the flashcard strategy.

**Chinese-assonance strategy:** The Chinese-assonance strategy is a special vocabulary learning method that uses similar-sounding phrases or sentences in the Chinese language to associate with English words. In general, the assonance strategy gives a Chinese sound similar to the English word and an associated interesting explanation. L2 learners can easily remember the similar sound via the explanation written in their native language and directly remember the target word. Figure 1(c) shows an example of assonance strategy written in a Chinese sentence.

**Synonym strategy:** Synonyms are different words with identical or similar meanings (e.g., silent and quiet, tranquil and serene). The synonym strategy provides a reference link that lists words grouped together according to their similarity of meaning. The word-card strategy provides complete word information (e.g., K.K. phonics, word type, and example sentence) related to synonyms, via a reference link. Figure 1(e) shows an example of the synonym strategy.

**Antonym strategy:** Antonyms are words with opposite meanings. The antonym strategy provides learning effects similar to those in the synonym strategy. Figure 1(d) shows an example of the antonym strategy.

**Imagery strategy:** Visual imagery, used to represent words, is an effective strategy for memory storage. Actually, some words are easier to visualize than others. The imagery strategy works best with words that correspond to concrete objects, but may not work well with words that denote abstract concepts because such concepts can be hard to depict. The use of the imagery strategy in MyEVA mainly employed an illustration to depict one to four concrete objects, via specific scenes or vivid cartoons. Figure 1(f) shows an example of the imagery strategy.

**Grouping strategy:** Research suggests that our mental lexicon is highly organized and efficient, and that items related semantically are stored together. Word frequency is another factor that affects storage, because the items used in memory most frequently are easier to retrieve (Hincks, 2003). The grouping strategy facilitates the learning process by grouping vocabulary items in semantic fields, such as topics (e.g., leisure and fun), categories (e.g., types of food), and scenes (e.g., city parts). Figure 1(g) shows an example of the grouping strategy in MyEVA.

**Clipping strategy:** Clipping is a word-formation process that breaks words into several small parts. The clipping strategy in MyEVA consisted mainly of three types: back-clipping (e.g., con[trol]), fore-clipping (e.g., tele[phone]),
and middle-clipping (e.g., inhabit). Words that have any types of clipping may provide a morphological clue to the meaning of the word or serve as an anchor for readers to remember the target word. Figure 1(h) shows an example of the imagery strategy.

Figure 1. The screenshot of eight selected vocabulary-learning strategies in MyEVA

Using the mixed-modality vocabulary-learning strategy in MyEVA

To integrate a mixed-modality vocabulary learning strategy in an e-learning system, it is critical to provide suitable navigational pathways for learners in each VLS. Students can differ in their learning technique in terms of the serial
versus the holistic approach (Ausburn & Ausburn, 1978). Serial learners prefer information to be introduced gradually and work in a linear manner, whereas holistic learners tend to address a learning task as an integrated big picture and then explore the details as their grasp of the overall picture takes shape. Two navigational pathways of the mixed-modality vocabulary learning strategy in MyEVA were developed based on the concepts of serial and holistic approaches, as shown in the Figure 2:

**Basic mode (with basic navigational pathway):** In the basic mode, students could freely select any of the available VLSs for study of target words in MyEVA. The displayed sequence of VLSs was arranged via the difficulty level (Word card → Flashcard → Chinese assonance → Antonym → Synonym → Imagery → Grouping → Clipping). The system guided students viewing the VLSs in a linear manner (i.e., starting the learning with the word-card strategy and followed by the Flashcard strategy, Chinese-assonance strategy, etc.). A screenshot of the basic mode in MyEVA is shown in Figure 2(a).

![Basic mode](image)

**Preference mode (with preference navigational pathway):** The vocabulary learning materials using the preference mode were essentially the same as the basis mode, except that the preference mode asked learners to pre-configure a preferred learning strategy before actual learning. The learning process prior to selecting a preference mode was: (1) The learner quickly viewed each available VLS for target. (2) The VLS exploration phase, the learner was required

![Preference mode](image)

*Figure 2. The screenshot of basic mode and preference mode in MyEVA.*

Preference mode (with preference navigational pathway): The vocabulary learning materials using the preference mode were essentially the same as the basis mode, except that the preference mode asked learners to pre-configure a preferred learning strategy before actual learning. The learning process prior to selecting a preference mode was: (1) The learner quickly viewed each available VLS for target. (2) The VLS exploration phase, the learner was required
to select one VLS for each word, which was considered the best memorizing strategy for that word. (3) Once the preference configurations for all target words were complete, MyEVA automatically displayed the preferred VLS every time the learner studied the word. A screenshot of the preference mode in MyEVA is shown in Figure 2(b).

In the development of learning materials for MyEVA, this study focused on TOEIC (Test of English for International Communication) vocabulary, which is highly regarded in educational and industrial institutes. MyEVA included 3,569 TOEIC words. Each target word had at least two VLS options.

**Research design**

The focus of this study was to determine whether mixed-modality VLSs applied on an e-learning system had a significant effect on L2 vocabulary acquisition. Therefore, the study compared the effectiveness of four devices: A: MyEVA in basic mode, B: MyEVA in preference mode, C: an Internet dictionary (Yahoo Dictionary), and D: a traditional paper-based dictionary. During the course of the study, each student used each of the four techniques, with a different group of words in each technique. The order in which each student would use the four techniques was determined randomly. Details of the research design are shown below:

### Selection of target words

This study used a selecting policy for target words based on Folse (2006). The main concern in selecting the target words for the research was that they be unknown to the participants. Because there were four groups of words and each participant experienced tests in each of the four groups, the researchers selected 36 words from the system database (comprising a total of 3,569 TOEIC words). Each group included nine words, six of which were observed target words and three were confusing words. The selection protocol for the target words was as follows:

*From the Brown Corpus, words that are used the least were selected:* In this study, the target words needed to be words that had not been learned previously by the participants in order to ensure that determination of learning effectiveness was the result of the four learning situations. Thus, according to the Brown Corpus (retrieved from http://nltk.googlecode.com/svn/trunk/nltk_data/index.xml), the researchers selected words that were used less frequently.

*Target words were nouns, verbs, and adjectives:* Different words may be memorized differently. Laufer (1990) suggested that nouns are the easiest to learn and adverbs are the most difficult. Verbs and adjectives are in the middle. Words selected in this study were mostly nouns, verbs, and adjectives. Adverbs were not selected in order to avoid particularly difficult words and because adverbs usually have the fewest potential vocabulary-learning strategies.

*Each word was independent and was not part of a phrase:* Some words are used only in specific phrases and thus were eliminated; otherwise, the special grammatical memory of the participants would influence the results.

*Words were dispersed in four groups, and the first letter of each word was based on alphabetical order:* In this study, the first letters of words in the four groups were fixed (a, i, g, h, a, v, a, i) to reduce the chances of students guessing memorized words, thus increasing the validity and reliability of the test.

*Length of each word was controlled:* When a word is long, it is more difficult to memorize. Thus, this study selected words of similar length. For instance, the number of letters in *hazard* in the first group, *harvest* in the second group, *hammer* in the third group, and *harmony* in the fourth group was six, seven, six, and seven, respectively.

*Setting of confusing words:* It is easier for learners to memorize shorter words. Among the verbs, nouns, and adjectives of the nine words in each group, the researcher respectively selected one shorter word as a confusing word. The purpose was to easily observe the learning effectiveness of target words (Folse, 2006). The words that were finally selected are shown in Table 2:
Table 2. Selected words used in the study

<table>
<thead>
<tr>
<th>Group</th>
<th>Target word (v)</th>
<th>Target word (v)</th>
<th>Interfered word (v)</th>
<th>Interfered word (v)</th>
<th>Target word (n)</th>
<th>Target word (a)</th>
<th>Target word (a)</th>
<th>Interfered word (a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Accredit</td>
<td>Inhabit</td>
<td>Glean</td>
<td>Gutter</td>
<td>Hazard</td>
<td>Anonymity</td>
<td>Visual</td>
<td>Advisable</td>
</tr>
<tr>
<td>2</td>
<td>Abstain</td>
<td>Inherit</td>
<td>Graze</td>
<td>Grant</td>
<td>Harvest</td>
<td>Ambition</td>
<td>Voluntary</td>
<td>Antenatal</td>
</tr>
<tr>
<td>3</td>
<td>Accuse</td>
<td>Intervene</td>
<td>Greet</td>
<td>Grasp</td>
<td>Hammer</td>
<td>Alternative</td>
<td>Volatile</td>
<td>Amiable</td>
</tr>
<tr>
<td>4</td>
<td>Annoy</td>
<td>Initiate</td>
<td>Grind</td>
<td>Gravel</td>
<td>Harmony</td>
<td>Amateur</td>
<td>Variable</td>
<td>Ambiguous</td>
</tr>
</tbody>
</table>

Pilot test: Before formal data collection, this study conducted a pilot test, intended to test the selection of target words and ensure that all target words were unfamiliar words for the participants, in order to ensure the content validity of the test. The process of the pilot test was the same as the formal one; however, fewer samples were employed.

Pre-test: Retention of learning is an important consideration in all instructional design. The researchers, therefore, used a pre- and post-test methodology to measure whether the students remembered the vocabulary after the completion of the learning activities. The vocabulary knowledge scale (VKS) was used in pre-testing and post-testing to evaluate the participants’ vocabulary capabilities. The VKS is a five-point self-reported scale developed by Wesche and Paribakht (1996), which allows L2 learners to indicate how well they know items of vocabulary. The VKS is usually used to measure small gains in knowledge to compare the effectiveness of different vocabulary instructional techniques (Schmitt, 2000). This study used a modified three-level VKS to detect even partial gains in the degrees of the participants’ vocabulary knowledge. Thirty-six words were included in the VKS for the pre-test. The participants were asked to complete the pre-test within 40 minutes. An example of the VKS is as shown in Figure 3. If MyEVA were fully integrated into a curriculum, language learners would use it outside the classroom over a longer period of time, and also use the learned words in other learning contexts, leading to reinforcement and repeated use of the vocabulary until mastery is achieved. As such, the authors conducted only one post-test, one week after the learning tasks ended, and did not attempt to measure subsequent memory decline.

Abstain

1. I don’t know this word. ☐

2. I know this word. The word meaning is __________________.

(Please write down a synonym or translate this word to Chinese.)

3. I know this word and I can use this word to make a sentence.

(Please use this word to write a simple sentence in English.)

(*You have to finish question 2 if you answer question 3.)

Figure 3. The modified three-level VKS

Preparation and word learning: During word learning, each participant had to use each of the four learning techniques with each in the four groups of words. To avoid the recency effect (Tversky & Kahneman, 1981), the effect of the learners’ memory of the latest teaching materials, this study arranged and combined four groups of words (1, 2, 3, and 4) with four VLS techniques (A, B, C, and D). There were 576 possible results. For instance, participants in (3C, 1A, 2B, and 4D) would learn the third groups of words via the Internet dictionary, the first group of words via the basic MyEVA mode, the second group of words via the MyEVA in preference mode, and the fourth group of words via traditional paper-based dictionary. In order to avoid the effect on the results of the participants’ time spent using the Internet dictionary and the traditional paper-based dictionary, the researcher constructed one webpage with hyperlinks to all target words to the Yahoo dictionary, including 36 words (four word groups). Thus, participants could directly look up the words on the webpage.

Regarding the traditional paper-based dictionary, the researcher first distributed mini-dictionaries selected from the English-Chinese Dictionary (Ji, 1993) published in Taiwan, including the 36 words (four groups of words) in the...
study. There were manuals for learning. Formal data collection was conducted in a computer classroom. Each participant had one computer, with earphones and an Internet connection. At the beginning of the study, the researcher taught participants how to learn with MyEVA (including basic mode and preference mode) and how to learn using the Yahoo dictionary and the traditional mini-dictionary. After this instruction, the researcher distributed one sheet to each participant and indicated the participants’ ID, learning technique, and words in the four learning phases. The researcher then guided the participants to learn according to the situations arranged. Each learning phase lasted for 20 minutes.

**Post-test:** Post-testing was conducted one week after the completion of the learning activities. The one-week period was selected for two reasons: (a) to measure actual retention of words learned several days after the learning activity to avoid the often-encountered phenomenon of students retaining knowledge only until a test or task is complete, and (b) to reduce the possibility of the student remembering words from the pre-test. The study concluded near the end of the academic semester for the students, who then left the university and would not have been available for a post-test conducted after a longer interval. In order to probe the participants’ preference for the four learning techniques, one item was added in post-testing, and the learners were asked to develop a ranking for the four learning techniques. The procedure is as shown in Table 3.

<table>
<thead>
<tr>
<th>Week 1</th>
<th>Description</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>Hand out the VKS paper for measuring the participants’ vocabulary capability.</td>
<td>40 minutes</td>
</tr>
<tr>
<td>Week 2</td>
<td>Instruction for using MyEVA with basic mode, preference mode, Yahoo dictionary, and mini-dictionary.</td>
<td>60 minutes</td>
</tr>
<tr>
<td>Word learning (with randomly assigned sequence)</td>
<td>Learning 9 words via MyEVA with basic mode.</td>
<td>20 minutes</td>
</tr>
<tr>
<td></td>
<td>Learning 9 words via MyEVA with preference mode (VLS exploring phase was limited to 5 minutes).</td>
<td>20 minutes</td>
</tr>
<tr>
<td></td>
<td>Learning 9 words via Internet dictionary.</td>
<td>20 minutes</td>
</tr>
<tr>
<td></td>
<td>Learning 9 words via traditional paper-based dictionary.</td>
<td>20 minutes</td>
</tr>
<tr>
<td>Post-test</td>
<td>Hand out the VKS paper for measuring the participants’ vocabulary capability.</td>
<td>40 minutes</td>
</tr>
</tbody>
</table>

**Participants**

This study was conducted at a private institute of science and technology in northern Taiwan. Nine undergraduates participated in the pilot test, selected as a convenience sample. The VKS mean score in pilot testing was 1.22 (the maximum score for the 24 target words was 48), which confirmed that the words used in the study were unfamiliar to the participants. Ninety-three undergraduates participated in the formal phase of the study. They were non-English majors who enrolled in an elective course of multimedia design, taught by the first researcher. Because the course was an elective, the participants ranged from second-year to fourth-year students, i.e., sophomores, juniors, and seniors. They were 20–23 years old, including 42 males and 51 females. The participants were volunteers who wished to improve their English, invited to participate by the instructor. Participation was not part of their grade in the course.

At the university at which this study was conducted, and for most university graduates of non-English departments in Taiwan, the graduation standard is a TOEIC score of 350. Among the 93 participants, 39 had passed the threshold (TOEIC ≥ 50) and 54 had not. The researchers placed 39 participants were in a “good learner group” and the other 54 participants into a “poor learner group.”

**Findings and discussion**

The data analysis led to the following answers to the research questions:
Research Question 1: Did MyEVA (in either basic mode/preference mode) improve the learners’ vocabulary capability?

The findings of the study show that the VKS scores from the pre-testing and post-testing were 1.06 and 16.61, respectively. A paired sample t-test showed significant differences (at the 95% confidence level) between the pre-test and post-test (t = 18.27, p < .001). Therefore, the students did retain the vocabulary learned via MyEVA for longer than the duration of the learning activities, and MyEVA did improve the vocabulary capability of the participants in the study over the course of the study. This outcome is not surprising, given the wealth of studies in the academic literature showing the benefits of computer-assisted language learning.

Research Question 2: Which learning technique achieved the best vocabulary-learning outcome for L2 learners with different levels of vocabulary proficiency?

When all learners were considered, use of the preference mode resulted in the best learning outcomes (M = 4.83), followed by the basic mode (M = 4.43), Internet dictionary (M = 4.33), and traditional dictionary (M = 3.00). The result was similar to research findings of Tight (2010) that mixed-modality vocabulary learning strategies enhance word-learning effectiveness for EFL students. The differences of VKS scores between the pre-test and post-test are as shown in Figure 4.

However, when the different levels of vocabulary proficiency of the participants were broken out, the preference and basic modes had different effects on learning effectiveness. For the good learner group, the results of the VKS scores, listed from highest to lowest were preference mode (M = 6.08), the basic mode (M = 4.69), Internet dictionary (M = 4.38), and traditional dictionary (M = 2.77). The poor learner group, however, experienced different learning outcomes compared to the good learner group. The Preference mode produced the greatest learning outcome (M = 4.39), followed by the Internet dictionary (M = 4.10), basic mode (M = 3.94), and traditional dictionary (M = 3.23).

In other words, the good learners achieved the best results with the preference mode by a notable margin over the basic mode, which held only a slight advantage over the Internet dictionary, with the traditional dictionary a distant fourth. For the poor learners, on the other hand, the preference mode was only slightly better than the Internet dictionary, followed closely by the basic mode, with the traditional dictionary, again, in a distant fourth place.
To determine the significance of these findings among the groups, a one-way ANOVA was calculated, indicating that differences did exist among the four groups \((F = 6.972, p < .001)\). In order to further determine whether there was significant difference in learning effectiveness among the four strategies, the researchers used paired t-tests. The t-test results are as shown in Table 4. The results indicated that for the good learners, the preference mode performed significantly better than the other three learning techniques (result #1, result #4, result #5, \(p < .001\)). There were no significant differences between the basic mode and Internet dictionary for participants in the good learner group (result #2, \(p = .352\)). For the poor learner group, there were no significant differences among the preference mode, basic mode, and the Internet dictionary (result #1, result #2, result #4); although all produced significantly better results than the traditional dictionary (result #3, result #5, result #6).

In this study, basic mode and preference mode of MyEVA represented two different navigational pathways: serial and holistic. Both modes implemented the mixed-modality learning strategy. The results of pre-test and post-test scores in Figure 4 show that the preference mode obtained a greater learning outcome than the basic mode. The results of paired t-test of preference mode and basic mode in Table 4 (result #1) showed that for the good learners, the preference mode was significantly better than the basic mode; however, for the poor learners, there were no significant differences between preference mode and basic mode. These findings indicate that good learners achieved the best retention using the holistic navigational pathway (preference mode), while the poor learners’ learning outcomes had no significant differences among distinct navigational pathways.

<table>
<thead>
<tr>
<th>Groups</th>
<th>All learners</th>
<th>Good learners</th>
<th>Poor learners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((N = 93))</td>
<td>((N = 39))</td>
<td>((N = 54))</td>
</tr>
<tr>
<td></td>
<td>(p)-value</td>
<td>(p)-value</td>
<td>(p)-value</td>
</tr>
<tr>
<td>1. Preference mode vs. Basic mode (MyEVA Mobile with different navigational pathways)</td>
<td>.108</td>
<td>.000***</td>
<td>.127</td>
</tr>
<tr>
<td>2. Basic mode vs. Internet dictionary</td>
<td>.280</td>
<td>.352</td>
<td>.407</td>
</tr>
<tr>
<td>3. Basic mode vs. Traditional dictionary</td>
<td>.000***</td>
<td>.000***</td>
<td>.028*</td>
</tr>
<tr>
<td>4. Preference mode vs. Internet dictionary</td>
<td>.005**</td>
<td>.000***</td>
<td>.862</td>
</tr>
<tr>
<td>5. Preference mode vs. Traditional dictionary</td>
<td>.000***</td>
<td>.000***</td>
<td>.000***</td>
</tr>
<tr>
<td>6. Internet dictionary vs. Traditional dictionary</td>
<td>.005**</td>
<td>.000***</td>
<td>.000***</td>
</tr>
</tbody>
</table>

Note. *\(p < .05\).
**\(p < .01\).
***\(p < .001\).

The good learner group, by definition, had higher prior English knowledge and ability, and in EFL learning, ability, confidence, and motivation are all positively interlinked (Wu, Marek & Yen, 2012). Thus, their higher ability resulted in more self-confidence and higher motivation, compared to the poor students, resulting in higher learning effectiveness when using the mixed-modality vocabulary-learning strategy of the preference mode.

The poor learner group’s use of the mixed-modality vocabulary learning strategy was not significantly better than when using the Internet dictionary (result #2 and result #4). Mixed use of VLSs did not result in higher learning effectiveness, and the students made similar progress with a single VLS, such as a word card from the Internet dictionary. The result demonstrated that, for the poor learner group, without the capability and habit of using the mixed-modality vocabulary-learning strategy, they would not have higher learning effectiveness.

It is notable that the traditional dictionary produced the lowest results for both the good and poor learner groups. This finding is not surprising, given the academic literature, which shows that students greatly prefer multimedia instructional materials, compared to conventional print materials (Lu, 2008; Nation, 2001). In other words, most students in Taiwan are used to digital and multimedia information sources. The learning effectiveness of these sources could not be enhanced by the unexciting traditional dictionary.

**Research Question 3: Did individuals with different levels of vocabulary proficiency have differences in preferred learning strategies?**

The post-test results were used to observe the preferences of the participants for the four learning techniques. In order to make this determination, each participant’s first preference for a learning technique was awarded four
points, the second preference received three points, the third received two points, and the fourth was given one point. The calculation result is as shown in Table 5.

<table>
<thead>
<tr>
<th>Learning technique</th>
<th>All learners (N = 93)</th>
<th>Good learners (N = 39)</th>
<th>Poor learners (N = 39)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre. rank</td>
<td>Score</td>
<td>Pre. rank</td>
</tr>
<tr>
<td>Basic mode</td>
<td>2</td>
<td>268</td>
<td>2</td>
</tr>
<tr>
<td>Preference mode</td>
<td>1</td>
<td>299</td>
<td>1</td>
</tr>
<tr>
<td>Internet dictionary</td>
<td>3</td>
<td>256</td>
<td>3</td>
</tr>
<tr>
<td>Traditional dictionary</td>
<td>4</td>
<td>107</td>
<td>4</td>
</tr>
</tbody>
</table>

*Note. Score = rank1 * 4 + rank2 * 3 + rank3 * 2 + rank4 * 1.*

The preference mode was the preferred mode, when all students were considered. For the good learners’ group, the preference mode was followed by the basic mode, the Internet dictionary, and the traditional dictionary. For the poor learners, however, the preference mode was followed by the Internet dictionary, then the basic mode, and finally the traditional dictionary.

According to previous research, when learners use a suitable learning strategy, it leads to higher learning effectiveness (Leaver, Ehrman, & Shekhtman, 2005). It is not surprising that the participants in the good learners’ group selected a preferred strategy for learning after first trying and comprehending all available strategies for learning the target words (VLS exploring phase). The basic mode listed all available strategies according to difficulty, and learners freely learned the words by any learning strategy (Hacquebord & Stellingwerf, 2007; Hsiao & Oxford, 2002). The findings, as well as observation by the researchers, showed that the poor learners did not select a preferred strategy but rather tended to use what they perceived to be the easiest strategy (such as word cards or flashcards).

In Taiwan, the Internet is popular and most students are familiar with learning and looking up vocabulary words by Internet dictionary. To some degree, the MyEVA basic mode and the Internet dictionary both allowed learners to learn by freely browsing. The basic mode has more VLSs, but the poor learners had lower self-efficacy than the good learners and, as a result, saw merit in what they considered to be the more familiar, and therefore easier, learning technique, the Internet dictionary. Nevertheless, the preference mode offered focused learning and memorization, and was useful and preferred by both good and poor learner groups.

**Conclusions and recommendations**

This study demonstrated a mixed-modality vocabulary learning system for L2 undergraduates, and compared its learning outcomes with an Internet dictionary and a traditional paper-based dictionary. The results showed that the mixed-modality VLS, with the preference strategy setting (preference mode), stimulated the greatest vocabulary retention. Findings also indicated that participants with different prior knowledge had distinctly different learning outcomes. The mixed-modality VLS in basic mode had an effect that was similar to the Internet dictionary for lower proficiency vocabulary learners. Higher proficiency vocabulary learners made use of the preference strategy, and achieved more effective vocabulary retention; however, the mixed-modality vocabulary learning strategy produced approximately the same learning outcomes as the Internet dictionary, with which the students were familiar.

The researchers believe the results of this study will provide insight for language teachers, curriculum designers, and in particular, the system developers of English e-learning systems for L2 learners and offer the following recommendations:

- E-learning techniques, such as MyEVA, can play a valuable role in EFL vocabulary learning because students believe that studying via e-learning is convenient and even preferred, compared to traditional classroom drill and memorization.
Practitioners must remember that students have different learning styles and that e-learning should support these learning styles with multiple vocabulary learning strategies (VLSs). The mixed-modality application in the current study allowed students to select a preferred mode or VLS. The modes employed in this study represent a starting place for other EFL vocabulary applications, i.e., word cards, flashcards, Chinese assonance, antonyms, synonyms, imagery, grouping, and clipping. In particular, more study would be valuable to analyze how each of these mixed-modality vocabulary e-learning strategies aligned with various learning styles to yield the most desirable learning outcomes.

Although digital learning can occur outside the classroom via various technologies, motivation is still a key to effective learning and must therefore be addressed in the classroom. In this study, the poor learners never selected a preferred VLS, doing whatever they perceived as easiest or the least work necessary to complete the requirements of the assignment to use MyEVA. There is a broad body of academic literature addressing motivation that can be tapped (such as Dörnyei, 2014). The overall instructional design of any curriculum needs to motivate poor learners as well as serve the interests of better learners.

This was a relatively short-term test using a mixed-modality VLS via e-learning. In order to pave the way for long-term use in a sustainable instructional technology design (Kennedy & Levy, 2009), more and longer testing should be performed in the future. Although the current study used an application suitable for a desktop, laptop, or notebook computer, an application similar to MyEVA could also be made available in the form of a smart phone or tablet computer app, to allow further flexibility in when and where the student uses the application. Future systems similar to MyEVA should also have a variety of ways to browse vocabulary words in order to achieve the best learning outcome. More ways of browsing would allow students to better make use of their own learning style, and might, therefore, help motivate less ambitious learners.

Our understanding of the process of developing instructional technology to support long-term curriculum needs is that (1) affordances of various technological options must be determined, typically with relatively short tests, (2) based on the wide range of available affordances, technologies are selected to address specific educational outcomes are incorporated into the curriculum, i.e., as a long-term implementation, and (3) evaluation takes place over multiple semesters or academic years (Kennedy & Levy, 2009; Colpaert, 2010; Dörnyei, 2014). This study resides in step one of this process, determining affordances, and lays the groundwork for the next step.

The findings of this study indicate that the use of mixed-modality VLSs via e-learning can improve the vocabulary learning outcomes of EFL learners. Different learning pathways, as offered in this study, can also influence the learning outcomes. The final influence on the outcome is the frequency with which the learner uses the technique. The researchers hope that CALL researchers and system developers wishing to design a mixed-modality VLS system can benefit from the experiences and the recommendations in this study.

Acknowledgements

The authors would like to thank the NSC of the Republic of China, Taiwan, for financially supporting this research under contract No. NSC 101-2511-S-231-001.

References


What Connected Educators Do Differently
(Book review)

Reviewer:
Christopher Bowen
Ph.D. candidate, University of Tennessee, USA
cbowen11@vols.utk.edu

Textbook Details:
What Connected Educators Do Differently
Written by Todd Whitaker, Jimmy Casas, and Jeffrey Zoul

In order to strengthen their skills, education professionals are expected to participate in professional development experiences. These opportunities allow educators to collaborate and create solutions to increase the success of their students. Yet, most teachers claim professional development is not tailored to their individual needs (Marzano, Waters, & McNulty, 2005). In What Connected Educators Do Differently, education professionals Todd Whitaker, Jimmy Casas, and Jeffrey Zoul, use their combined 90 years of educational experience to present their solution: personalized professional development through networking with educators using social media. With the pressures of accountability and high-stakes standardized tests, professional development usually focuses on state and federal mandates. Educators typically are not asked for input on how to strengthen their own pedagogy. Within a school, the needs of individual teachers can range from technology integration to classroom management to instructional strategies. When educators have a voice in aligning professional development goals to their personal needs, they feel more confident about using their new skills to lead their students to success (Marzano et al., 2005).

Guiding the novice

The strengths of this book are numerous. A standout feature is the practicality and personalized writing style, which I believe, gives the reader confidence in the insights being offered. Using relevant advice and personal accounts, Whitaker, Casas, and Zoul guide readers in an authentic manner through an interpersonal (and intrapersonal) journey toward becoming a connected educator. Each chapter, labeled as a key connector, is an action step that connected educators choose to take to advance their professional growth – it’s what connected educators do differently. Through the linking of key connectors, the authors paint a picture of what it is like to be a connected educator and how it can transform personal and professional lives.

To assist in creating (or continuing) a personal and professional learning network (P²LN), each chapter ends with three sections. The first section introduces five connected educators, their Twitter handles, and their personal advice related to the chapter. Providing five online resources recommended by members of the authors’ connected communities, the second section assists the reader with the implementation of the chapter’s theme. The final section presents five additional challenges to becoming a connected educator. The authors’ inclusion of stories from members of their own P²LNs encourages educators to begin this personal and professional journey of risk, change, service, and growth. I found these sections to be extremely helpful.

When should you connect?

Whitaker, Casas, and Zoul welcome readers with an ancient proverb: “The best time to plant a tree is twenty years ago. The second best time is now” (p. xix). This proverb sets the stage by encouraging veteran and apprentice teachers alike that now is the time to become a connected educator. Even with the authors’ collective experiences, including K-12 teaching and administrative roles, presenting at professional conferences, and authoring over thirty books, they admit they are far from being “digital natives.” This vulnerable revelation set my mind at ease by providing a powerful example of how establishing personal and professional connections outside of the walls of their schools propelled Whitaker, Casas, and Zoul toward excellence.
Why connect?

Because there’s little information available on personalized professional development, most educators are hesitant about initiating their own growth as educators. This factor, combined with the majority of the workday is spent with students, results in limited peer-to-peer interactions that unfortunately affects student learning. With their multiple responsibilities, educators typically do not have time in their daily schedules to learn (individually or collaboratively) something new about their practices. These factors, along with professional isolation, result in low self-efficacy, reduced performance, personal burnout, and high turnover rate. Whitaker, Casas, and Zoul believe P2LNs provide individualized collaborative learning with “job-alike” peers to educators desiring to extend their connections beyond their resident colleagues. Individual educators know what areas to strengthen in their pedagogy; forming a connected community allows educators to choose what they want to learn, when they want to learn, and how they learn. It also provides personalized support, flexible scheduling and a comfortable environment. Finally, connected educators use social media (i.e., Twitter, Facebook, etc.) to build school communities where administrators, teachers, staff, and students share their “school story” with parents, community members, and even policy makers.

How do I connect?

Gently treading the line of an outright Twitter endorsement, Whitaker, Casas, and Zoul explain through practical step-by-step instructions how building an initial P2LN using Twitter is a personal investment, but one that leads to multiple benefits. From advice on creating a Twitter profile to building a network, the authors support readers throughout this personal journey toward becoming connected. However, the authors do not avoid mentioning the personal investment of time and effort required to be a connected educator; the authors even use the word “invest” in the first chapter’s title. The old adage, it is better to give than to receive, is revived within the virtual world with advice on giving and taking. Connected educators give back to their P2LN by promoting people and their ideas, responding to a message, question, or call, and by connecting members with people they know. Similarly, they take ideas, suggestions, and resources, reach out for support from members of their P2LN, and gather additional members to add to their P2LN.

Because I was not a connected educator prior to reading this book, I followed the authors’ instructions and advice, and I found them to be extremely user-friendly. I created my Twitter profile and began developing my own P2LN by searching Twitter hashtags, following connected educators suggested throughout this book, and attending Twitterchats recommended by the authors and their connected communities. In fact, three of the first five people in my P2LN were Whitaker, Casas, and Zoul!

Set apart from the rest

Whitaker, Casas, and Zoul consistently challenge our commitment to become connected through the use of statements that illustrate how connected educators are set apart from their colleagues. Some of the statements are even reprinted in callout boxes within the text, including the following:

- “We have paddled our way through this vast sea we call teaching and learning on our own for way too long. The time is now for us to take responsibility for owning our professional learning by networking with individuals who share our passion and our desire to be the change!” (p. 13).

- “Therefore, [connected] educators do not find time to give back (As there is no additional time to “find” in the 24 hours we are allotted each day); they make time to give back” (p. 48).

- “No matter how well versed we are on any topic, together we can always know even more. The good news is that there are literally thousands of educators ready and willing to lend their expertise if asked, typically swiftly and absolutely free of charge” (p. 50).

- “Connected educators embrace their vulnerability and extend their learning outside their comfort zone in order to make a broader and greater impact on their school communities and across global communities everywhere” (p. 62).
“At times, instead of turning to someone within our current organization for advice, it behooves us to seek out someone from our learning network who can offer an outside perspective into the situation” (p. 85).

Things to consider

Even though Whitaker, Casas, and Zoul’s book is a valuable guide for the connected educator, it would be a challenge for some educators to “communicate with purpose” (p. 30). While reading the key connector about communication, I thought about my students from low socioeconomic households and their limited access to technology. One shouldn’t assume that every family in a Title I school district would have continuous access to social media. This limitation would make it difficult to interact with all school stakeholders using social media such as Twitter, Facebook, Instagram, or blogs, as highly encouraged by the authors.

Because the authors did not mention the possibility of stakeholders being left out of community-building due to a technological void, I decided to voice my concerns through direct messages (i.e., private tweet) to members of my recently formed P2LN. I received feedback and suggestions from educators all over the country who have faced similar issues within their school communities.

Conclusion

Overall, What Connected Educators Do Differently is a useful and encouraging manual of instructions, advice, and personal accounts of the rights and responsibilities of being a connected educator. Whether a veteran educator, administrator, first-year teacher, or studying to be in the field of education, this book provides the information needed to create a personal and professional learning network.

References


The Chicago Handbook of University Technology Transfer and Academic Entrepreneurship
(Book Review)

Reviewer:
Hadi Shaheen
Management Consultant – Strategy and Innovation
Hadi.Shaheen@gmail.com

Textbook Details:
The Chicago Handbook of University Technology Transfer and Academic Entrepreneurship
Edited by Albert N. Link, Donald S. Siegel, and Mike Wright
2015, 297 pages, Published by The University of Chicago Press

Prior to the 1980’s Patent and Trademark Law Amendments Act, also known as the Bayh-Dole Act (Stevens, 2004), the US federal government retained all ownership rights of patents, inventions and discoveries that were created at universities through government-sponsored research. In a bipartisan attempt to nurture a staggering economy, the Bayh-Dole Act was enacted and the institutional ownership model of federally funded research-based inventions was set to motion. The groundbreaking act gave universities the right to own inventions and negotiate licensing terms.

Universities and research institutions have since then been endeavoring the commercialization tracks, navigating through a magnitude of technology transfer processes, and establishing licensing units. This has led to research interest from various disciplines to try to understand the academic entrepreneurship phenomena, and identify required constituents that could further shape successful commercialization projects.

In this book, Link, Siegel and Write present us with an arrangement of articles written by some of the leading thinkers on the subject of academic entrepreneurship and technology transfer. The information presented is based on solid evidence, rich in data and comes at a time where research institutes face an increasing pressure to commercialize and license their innovations.

The authors present technology transfer officers and researchers with compelling insights and evidence-based findings on how to develop a shared understanding of implications and policies that could foster academic entrepreneurship. The book is a much-needed resource to guide university technology transfer stakeholders in the innovation ecosystem to better plan and strategize their work.

The structure of information in the book and selection of articles from multi-disciplinary thinkers, offer readers intriguing ideas and applicable tools for implementation. Readers immediately start to learn about the best organizational practices through understanding key characteristics of successful academic startups. Moreover, entrepreneurs’ understanding of technology transfer processes is identified as key component for successful commercialization activities.

Fiona Murray and Julian Kolen (authors of Chapter 4) shift our perspectives and invite us to wear the entrepreneur’s hat and explore the three dimensions of technology transfer: national, local, and individual. Andrew Nelson and Tom Byers (authors of Chapter 5) help us initiate effective technology transfer activities by identifying and examining internal resources and competencies essential for overcoming commercialization challenges. According to the same authors, entrepreneurial education is an important factor, which may help accelerate innovation and “improve the effectiveness of university spin-outs” (Nelson & Byers, 2015). Albert Link and John Scott (authors of Chapter 6) present to us an examination of the constituents of innovation ecosystems, and the role innovation research parks play in stimulating academic entrepreneurship.

The implications of policies on commercialization and innovation are also valuable components of this book. The book presents us with a thorough examination of the international academic entrepreneurship phenomena and the consequences of patent policies on technology transfer in various parts of the world. It reveals a compelling study on critical determinants of successful university ventures and startups. The information presented offers extensive
guidance to university administrations and technology transfer officers for establishing entrepreneurial universities. For example, according to the findings, two of the key success factors that help build an entrepreneurial university are: inviting more entrepreneurial researchers to join the faculty, and rewarding academic entrepreneurs who are already engaged in innovative endeavors at the university.

The Chicago Handbook of University Technology Transfer and Academic Entrepreneurship is definitely a real treat and an insightful resource for anyone involved in academic entrepreneurship or in the innovation ecosystems.

References


The Chicago Handbook of University Technology Transfer and Academic Entrepreneurship
(Book Review)

Reviewer:
Hadi Shaheen
Management Consultant – Strategy and Innovation
Hadi.Shaheen@gmail.com

Textbook Details:
The Chicago Handbook of University Technology Transfer and Academic Entrepreneurship
Edited by Albert N. Link, Donald S. Siegel, and Mike Wright
2015, 297 pages, Published by The University of Chicago Press

Prior to the 1980’s Patent and Trademark Law Amendments Act, also known as the Bayh-Dole Act (Stevens, 2004), the US federal government retained all ownership rights of patents, inventions and discoveries that were created at universities through government-sponsored research. In a bipartisan attempt to nurture a staggering economy, the Bayh-Dole Act was enacted and the institutional ownership model of federally funded research-based inventions was set to motion. The groundbreaking act gave universities the right to own inventions and negotiate licensing terms.

Universities and research institutions have since then been endeavoring the commercialization tracks, navigating through a magnitude of technology transfer processes, and establishing licensing units. This has led to research interest from various disciplines to try to understand the academic entrepreneurship phenomena, and identify required constituents that could further shape successful commercialization projects.

In this book, Link, Siegel and Write present us with an arrangement of articles written by some of the leading thinkers on the subject of academic entrepreneurship and technology transfer. The information presented is based on solid evidence, rich in data and comes at a time where research institutes face an increasing pressure to commercialize and license their innovations.

The authors present technology transfer officers and researchers with compelling insights and evidence-based findings on how to develop a shared understanding of implications and policies that could foster academic entrepreneurship. The book is a much-needed resource to guide university technology transfer stakeholders in the innovation ecosystem to better plan and strategize their work.

The structure of information in the book and selection of articles from multi-disciplinary thinkers, offer readers intriguing ideas and applicable tools for implementation. Readers immediately start to learn about the best organizational practices through understanding key characteristics of successful academic startups. Moreover, entrepreneurs’ understanding of technology transfer processes is identified as key component for successful commercialization activities.

Fiona Murray and Julian Kolen (authors of Chapter 4) shift our perspectives and invite us to wear the entrepreneur’s hat and explore the three dimensions of technology transfer: national, local, and individual. Andrew Nelson and Tom Byers (authors of Chapter 5) help us initiate effective technology transfer activities by identifying and examining internal resources and competencies essential for overcoming commercialization challenges. According to the same authors, entrepreneurial education is an important factor, which may help accelerate innovation and “improve the effectiveness of university spin-outs” (Nelson & Byers, 2015). Albert Link and John Scott (authors of Chapter 6) present to us an examination of the constituents of innovation ecosystems, and the role innovation research parks play in stimulating academic entrepreneurship.

The implications of policies on commercialization and innovation are also valuable components of this book. The book presents us with a thorough examination of the international academic entrepreneurship phenomena and the consequences of patent policies on technology transfer in various parts of the world. It reveals a compelling study on critical determinants of successful university ventures and startups. The information presented offers extensive
guidance to university administrations and technology transfer officers for establishing entrepreneurial universities. For example, according to the findings, two of the key success factors that help build an entrepreneurial university are: inviting more entrepreneurial researchers to join the faculty, and rewarding academic entrepreneurs who are already engaged in innovative endeavors at the university.

The Chicago Handbook of University Technology Transfer and Academic Entrepreneurship is definitely a real treat and an insightful resource for anyone involved in academic entrepreneurship or in the innovation ecosystems.

References


Music and the Making of Modern Science
(Book review)

Reviewers:
Rébecca Guillot
Private piano and music theory teacher
Undergraduate student
School of Computing and Information Systems
Athabasca University, Alberta, Canada
rebecca.guillot@outlook.com

Isabelle Guillot
Undergraduate student
School of Computing and Information Systems
Athabasca University, Alberta, Canada
iguillot@sympatico.ca

Textbook Details:
Music and the Making of Modern Science
Written by Peter Pesic
2014, 360 pages, Published by the Massachusetts Institute of Technology
ISBN 978-0-262-02727-4

Has music been shaped by and emerged from science or is it the other way around? In his book “Music and the Making of Modern Science,” Peter Pesic argues that music has been instrumental in providing insightful avenues of solutions to several scientists as they were struggling to unfold mysteries throughout the course of history. Holding a doctorate in Physics from Stanford University, tutor and musician-in-residence at St. John’s College, Santa Fe as well as author of several books published by the MIT Press, Pesic possesses a unique background in which physics, history, and music are singularly intermingled, thus highly qualifying him to write this book.

In a very knowledgeable and brilliant manner, he addresses an academic audience with several thoroughly explained evidences based on numerous references which show how music does not have a secondary place in the development of science, but rather a forefront, even a leading position in the advancement of science discoveries. While convinced that music preceded science, the author does not make it an absolute, but rather chose specific “cases in which music led the way” (Pesic, 2014, p. 5). Though this book does not cover a complete account of scientific discoveries initiated by music, the cases judiciously chosen by the author allow the reader to follow a central thread where the discoveries through music by savants and their unanswered questions are taken up by the next generation of researchers who are answering the questions and building upon the findings of their predecessors which bring to the surface new questions, and this cycle continues as they walk on the path of advancement.

Reviewed from a music teacher and music student perspective, we read this book pondering how its content could help such readers expand their understanding on the real impact of music as to the expansion of science and how music goes even beyond anything that seems related to music. It is our understanding that only university-level readers with an interest and sound knowledge in music theory, mathematics, and science who are questioning themselves on where today’s music theories come from will really appreciate the depth of this book as its style is highly academic and scientific. We would even say that if any of these fields were to be missing in the reader’s knowledge background, the one field that is absolutely compulsory to relish the underlying explanations of the correlation between music and science is a strong mastery of music theory.

The greatest attraction of this book for a music reader is to follow the author as he chronologically covers the discovery process, including struggles, hesitations, questions, assumptions, unsolved problems, experiments, opposition, and exaltation of varied scientists over the centuries, brilliantly demonstrating how music has been the inspiration, the source, the intrigue, and the motivation to test and experience mathematical hypotheses, to challenge ancient theories and to go beyond what was considered as absolutes. Starting from the Greek quadrivium educational system which included music, Pesic constantly provides solid contextual and interesting historical facts that help one
grasps the omnipresence of music in the social and educational realms of these scientists, thus greatly influencing their mindset for discovery.

As just said, music was a bench work for researchers to test and experience mathematical and physics hypotheses, to challenge ancient views, and to advance innovative ideas. We can read in chapter 2 how music brought Oresme to subtly challenge the traditional geocentric view and how it guided him to think that the heliocentric view might be the right one, all this over the debate of an audible or inaudible earth. Chapter 3 describes how Josquin also challenged the long-time, almost sacred belief that either the sun or the earth was the central immovable object of the celestial system while he proved through music that the central mode of a song could be moved and still be harmonious. Descartes experienced how “the ear can test the exact influence of the resistive air” in plucking a string and listening to its pitch at the end compared to its pitch at the beginning, thus hearing if the air has slowed the movement or accelerated it. Pesic describes how Descartes returned to “the realm of musical experience” to prove his point (Pesic, 2014, p. 95).

It is with much interest that a music reader will learn how “Mersenne verified his observations ‘very exactly more than a hundred times, on the viol and on a theorbo, as well as on two monochords’” (Pesic, 2014, p. 116); how Newton used the degree of the scale to match the colors of the rainbow and how music theory in general supported him in his quest to understand ratios being thus a “helpful touchstone for his mathematical natural philosophy” (Pesic, 2014, p. 131); how a musical device was used by Wheatstone and then by Fizeau and Foucault to test the stability of the speed of rotation and to measure the speed of light in water and air, experiments that led to the theory of relativity; how Stifel believed that since irrational proportions “were already in current musical use” they “should be mathematically acceptable” (Pesic, 2014, p. 56); how the studies of sound and vibrations led Faraday and Wheatstone to go further in their understanding and discovery of sonic transmission which led to the discovery of electromagnetism, telephony and telegraphy; and how Young started and concluded with his musical assumptions to test and prove his hypotheses on light. The author points out that “[Young’s] papers showed the importance of music and sound as he discovered his new insights; his lectures showed how he continued to rely on sound and music not only in the context of the discovery of his ideas but also in the context of their public justification and popularization” (Pesic, 2014, p. 177).

As the music student and teacher will read the details of the above summarized facts demonstrating the central role of music for the science world, they will also find the audio and video excerpts provided by the author throughout the book tremendously interesting and helpful towards grasping theoretical concepts as they give a lively practical touch to the assumptions brought forth by scientists. Such a reader will try to figure out the intervals and their ratios even before listening to the examples. Thus, the sound examples create a sense of deeper understanding by confirming or demonstrating the conclusions made mentally by the reader. The mathematics explanation of semitones, dieses, chromatics and enharmonics is truly fascinating for musicians who seek the origins of all what are today’s absolutes in music theory.

In conclusion, the author has successfully proved the undeniable place of music in science discoveries, mainly that music served as a ruler, a solid ground with tangible realities that decided the acceptance or rejection of scientific theories as well as brought scientists to see and hear, through music, evidences that were yet unknown in the mathematics and physics domains. Simply said this book richly articulates how the palpable world of music allowed scientists to leap into the unknown world of mathematics.

References