Educational Technology & Society
An International Journal

Aims and Scope

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

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

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 The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to Educational Technology & Society and three months thereafter.

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


Editors

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

Editors’ Advisors

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

Editorial Assistant

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

Technical Manager

- Panagiotis Zervas, University of Piraeus & CERTH, Greece

Associate editors

- Vladimir A Fomichev, K. E. Tsiolkovsky Russian State Tech Univ, Russia;
- Olga S Fomichova, Studio "Culture, Ecology, and Foreign Languages", Russia;
- Piet Kommers, University of Twente, The Netherlands;
- Chul-Hwan Lee, 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.

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

Executive peer-reviewers

http://www.ifets.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.
All submissions should be in electronic form. The editors will acknowledge the receipt of submission as soon as possible.
The preferred formats for submission are Word document and RTF, but editors will try their best for other formats too. For figures, GIF and JPEG (JPG) are the preferred formats. Authors must supply separate figures in one of these formats besides embedding in text.
Please provide following details with each submission: • Author(s) full name(s) including title(s), • Name of corresponding author, • Job title(s), • Organisation(s), • Full contact details of ALL authors including email address, postal address, telephone and fax numbers.
The submissions should be uploaded at http://www.ifets.info/ets_journal/upload.php. In case of difficulties, please contact kinshuk@ieee.org (Subject: Submission for Educational Technology & Society journal).
Full Length Articles

Error-Based Simulation for Error-Awareness in Learning Mechanics: An Evaluation
Tomoya Horiguchi, Isao Imai, Takahito Tomoto and Tsukasa Hirashima
1–13

Self-Observation Model Employing an Instinctive Interface for Classroom Active Learning
Gwo-Dong Chen, Narkhamid, Chin-Yeh Wang, Shu-Han Yang and Po-Yao Chao
14–26

To See or Not to See: Effects of Online Access to Peer-Generated Questions on Performance
Fu-Yun Yu and Yen-Ting Yang
27–39

Experimenting with Automatic Text-to-Diagram Conversion: A Novel Teaching Aid for the Blind People
Anirban Mukherjee, Utpal Garain and Arindam Biswas
40–53

A Study about Placement Support Using Semantic Similarity
Marco Kalz, Jan van Bruggen, Bas Giesbers, Wim Waterink, Jannes Exhuis and Rob Koper
54–64

The Effects of Game-Based Learning on Mathematical Confidence and Performance: High Ability vs. Low Ability
Oskar Ku, Sherry Y. Chen, Denise H. Wu, Andrew C. C. Lao and Tak-Wai Chan
65–78

Developing Spatial Orientation and Spatial Memory with a Treasure Hunting Game
Chien-Heng Lin, Chien-Min Chen and Yu-Chuang Lou
79–92

Utilizing a Low-Cost, Laser-Driven Interactive System (LaDIS) to Improve Learning in Developing Rural Regions
Wei-Kai Liou and Chuan-Yen Chang
93–107

Effects of Communication Competence and Social Network Centralities on Learner Performance
Il-Hyun Jo, Stephanie Kang and Meehyuan Yoon
108–120

Design, Customization and Implementation of Energy Simulation with 5E Model in Elementary Classroom
Sze Yee Lye, Loo Kang Wee, Yao Chie Kwek, Suriati Abas and Lee Yong Tay
121–137

Development and Evaluation of Across-Unit Diagnostic Feedback Mechanism for Online Learning
138–153

Knowledge Sharing among University Students Facilitated with a Creative Commons Licensing Mechanism: A Case Study in a Programming Course
Chen-Chung Liu, Chia-Ching Lin, Chun-Yi Chang and Po-Yao Chao
154–167

The Role of Community Trust and Altruism in Knowledge Sharing: An Investigation of a Virtual Community of Teacher Professionals
Hsin-Ling Chen, Hsueh-Liang Fan and Chin-Chung Tsai
168–179

Visualisation in Applied Learning Contexts: A review
Adrian Twissell
180–191

A Model for the Design of Puzzle-based Games Including Virtual and Physical Objects
Javier Melero, Davinia Hernandez-Leo
192–207

Strategies for Smooth and Effective Cross-cultural Online Collaborative Learning
Junfeng Yang, Kinshuk, Huiju Yu, Sue-Jen Chen and Ronghuiai Huang
208–221

Learning to become a teacher in the 21st century: ICT integration in Initial Teacher Education in Chile
Mario Brun and J. Enrique Hinostroza
222–238

The Game Embedded CALL System to Facilitate English Vocabulary Acquisition and Pronunciation
Shelley Shwu-Ching Young and Yi-Hsuan Wang
239–251

An Integration Architecture of Virtual Campuses with External e-Learning Tools
Antonio Navarro, Juan Cigarrrin, Francisco Huertas, Miguel Rodriguez-Artacho and Alberto Cogolludo
252–266
The Effects of Student Question-Generation with Online Prompts on Learning
  Fu-Yun Yu and Kuan-Jung Pan
  267-279

The Difference in the Online Medical Information Searching Behaviors of Hospital Patients and Their Relatives versus the General Public
  Hung-Yuan Wang, Jyh-Chong Liang and Chin-Chung Tsai
  280-290

Digital Game-Based Learning Supports Student Motivation, Cognitive Success, and Performance Outcomes
  Jeng-Chung Woo
  291-307

Scripting For Collaborative Search Computer–Supported Classroom Activities
  Renato Verduco, Leonardo Barros, Daniela Albornoz, Miguel Nussbaum and Angela McFarlane
  308-319

Effectiveness of an Electronic Performance Support System on Computer Ethics and Ethical Decision-Making Education
  Serhat Bahadır Kert, Çiğdem Uz and Zeynep Gecü
  320-331

Lost in Interaction in IMS Learning Design Runtime Environments
  Michael Derntl, Susanne Neumann and Petra Oberhuemer
  332-342

Practising Arithmetic Using Educational Video Games with an Interpersonal Computer
  Vagner Beserra, Miguel Nussbaum, Ricardo Zeni, Werner Rodríguez and Gabriel Wurman
  343-358

Languages for Specific Purposes in the Digital Era
  Reviewer(s): Nuria Otero de Juan and Jesús García Laborda
  359-361
Error-Based Simulation for Error-Awareness in Learning Mechanics: An Evaluation

Tomoya Horiguchi1*, Isao Imai2, Takahito Toumoto3 and Tsukasa Hirashima4

1Graduate School of Maritime Sciences, Kobe University, 5-1-1, Fukaeminami, Higashinada, Kobe, Hyogo 658-0022, Japan // 2Hanazono Junior High School, 4-1-1, Hanazono, Hanamigawa, Chiba 262-0025, Japan // 3Faculty of Engineering Division II, Tokyo University of Science, 1-3, Kagurazaka, Shinjyuku, Tokyo 162-8601, Japan //
4Department of Information Engineering, Hiroshima University, 1-4-1, Kagamiyama, Higashihiroshima, Hiroshima 739-8527, Japan // horiguti@maritime.kobe-u.ac.jp // isao-1@mx3.mesh.ne.jp // tomuto@ms.kagu.tus.ac.jp // tsukasa@lel.hiroshima-u.ac.jp

*Corresponding author

(Submitted October 23, 2012; Revised July 9, 2013; Accepted October 24, 2013)

ABSTRACT

Error-based simulation (EBS) has been developed to generate phenomena by using students’ erroneous ideas and also offers promise for promoting students’ awareness of errors. In this paper, we report the evaluation of EBS used in learning “normal reaction” in a junior high school. An EBS class, where students learned the concept with EBS, was compared with a usual class where students learned it as usual. We used a pre-test, post-test, delayed post-test, and an interview after the delayed post-test. In tests, three kinds of tasks were included. A learning task was used in the learning process. A complex task is a little different from a learning task and can’t be solved by memorization. A transfer task is very different from the learning task and can’t be solved without generalizing the learning. In the post-test and delayed post-test, the scores of the EBS class were significantly higher than those in the usual class. The differences of scores between these classes in the complex and transfer tasks were larger than those in the learning task. In the interview, students in the EBS class explained their solution in more conceptual ways. These results suggest that EBS contributed to the generalization and retention of the learned concept, “normal reaction.”

Keywords

Learning from errors, Awareness, Misconception, Error-based simulation

Introduction

Natural science explains and predicts natural phenomena. One of the most important purposes of elementary science education is to make students able to explain and predict natural phenomena with scientific concepts. It is, however, often difficult for students to connect abstract scientific concepts to concrete natural phenomena. As well, disconnection of them causes several serious misconceptions (Driver, Guesne, & Tiberghien, 1985; Osborne & Freyberg, 1985). Therefore, supporting students’ comprehension of this connection is a very important issue in elementary science education.

Scientific experiment or demonstration is a popular teaching method to connect them. First, a phenomenon is shown to the students, and then it is explained with scientific concepts that are the targets of teaching. Simulation-based learning environments (SLE) have been investigated to assist such “learning from experiments” and have been found useful for the introduction or acquisition of scientific concepts (Towne, de Jong, & Spada, 1993). When a student makes an error in prediction, showing the correct simulation could be useful for correcting that error. The difference between the correct simulation generated by the SLE and the phenomenon predicted by the student makes the student aware of the error. However, students sometimes have wrong concepts for explaining correct phenomena. In such cases, phenomena generated by SLE aren’t useful because the phenomena are the same as those they connected to the wrong concepts. For example, in elementary mechanics, students often answer that gravity is the only force acting on a block on a table while they predict that the block will stay at rest on the table.

Error-based simulation (EBS) addresses this problem by generating a phenomenon using students’ erroneous ideas to help them become aware of errors when they know some correct phenomena connected to their wrong concepts (Hirashima, Horiguchi, Kashihara, & Toyoda, 1998). In the above example, EBS generates an unnatural phenomenon where the block sinks into the table because gravity, which isn’t cancelled by any other force, causes a downward motion of the block. EBS is a generally useful method to generate counterexamples to students’
misconceptions or erroneous answers, and we have developed several prototype systems not only for mechanics but also for drawing (Matsuda et al., 2003) and English composition (Kunichika, Hirashima, & Takeuchi, 2006).

EBS has an important feature that other systems don’t have that can bring out the pedagogical merit in simulation-based learning. Even when a model is incalculable (e.g., because the constraints specified by the student contradict each other), our approach involves creating an EBS by relaxing some of the student’s model constraint(s). In the above example, the constraint “rigid objects (i.e., the block and table) never overlap” is violated in the simulation. Thus, EBS works as a counterexample, demonstrating that “if students’ ideas were correct, the fundamental constraint in the real world would be violated.” Because of this feature, shown the next section in detail, one can design a learning environment with EBS that has the advantages of discovery learning (Crews, Biswas, Goldman, & Bransford, 1997; Loh, Reiser, Radinsky, Edelson, Gomez, & Marshall, 2000) and directed learning (Klahr, 2009).

So far, preliminary experiments in which small number of university students learned with the above systems and teachers evaluated the functions of these systems suggested the possibility of EBS (Kunichika et al., 2006; Matsuda et al., 2003). Additionally, a preliminary test in which small number of junior high school students learned elementary mechanics with EBS suggested that EBS promoted their conceptual understanding (Horiguchi, Imai, Toumoto, & Hirashima, 2007). Though the test was conducted in a practical situation, the result wasn’t statistically validated because no control group was made for comparison. In this paper, therefore, we describe a practical use and evaluation of EBS in a controlled experiment is described. We used EBS for learning “normal reaction” in a junior high school. The results suggest that EBS contributed to generalization and retention of the learned concept “normal reaction.”

In this paper, we first introduce the outline of EBS, then describe the features of EBS compared with the related work. The purpose of this practice and the procedure of the experiments are described next, followed by its results and the discussion about them. We conclude this paper with some remarks.

**Error-based simulation as a method to promote awareness of errors**

**Error-based simulation**

In order to make students correct their erroneous idea, cognitive conflict often plays an important role (Glynn, Yeany, & Britton, 1991; Osborne & Freyberg, 1985). When the difference between an erroneous answer and the correct one is significant for them, it could cause cognitive conflict. Unexpected correct phenomena shown in experiments or simulation, for example, often have such an impact. An explanation that connects the correct concepts to the phenomena could prompt students to correct their error.

However, when the difference isn’t significant for students, it hardly causes cognitive conflict (Chinn & Brewer, 1993). For example, when a student knows the correct phenomenon and explains it with erroneous concepts, the explanation with the correct concepts may fail to prompt him/her to correct the error because the difference between the erroneous and correct concepts has less impact. In such a case, therefore, the difference should be made more “visible.” That is, what the difference implicates should be clear for students to understand its significance. We call this “error visualization” (Hirashima et al., 1998; Horiguchi et al., 1999).

Figure 1 shows the framework of error visualization with EBS. EBS is generated by mapping errors in symbolic expression to erroneous behavior. In this framework, a student’s answer should be regarded as a model of the target in question from which the behavior is generated. The difference in behavior is better to make students aware of the errors and motivate them to correct the errors. In order to use EBS effectively, we investigated the following three factors: visibility, reliability, and suggestiveness. Visibility concerns whether the difference between normal behavior and EBS is enough to make students aware of the error. Reliability pertains to whether the mapping from symbolic expression to behavior is reliable for students. Suggestiveness pertains to whether the difference between normal behavior and EBS suggests the way to correct the error. These factors are very interesting and important, especially to extend the target domain of EBS (Horiguchi et al., 1999).
We show some examples of EBS with the problem in Figure 2(a). Blocks 2 and 3 are connected with a string through a pulley. The mass of Block 1 is \( M \), Block 2 \( 2M \), and Block 3 \( 2M \). The mass of the string and the pulley is negligible. The acceleration of gravity is \( g \). All blocks move without friction. \( T \) is the tension of the string between Block 2 and Block 3. Block 2 is restricted to movement along the right side of Block 1. Therefore, normal force, \( N \), works between the two blocks, and they move as one towards the left. Because Block 1 moves, relative acceleration should be used in the equation of motion of Block 3. The horizontal component of acceleration of Block 1 is \( a_1 \), Block 2 \( a_2 \), and Block 3 \( a_3 \). The vertical component of acceleration of Block 2 is \( b_2 \). Note that \( a_2 \) is equal to \( a_1 \), and \( a_3 \) is equal to \( b_2 \), respectively. The correct set of equations of this system is shown in Figure 2(b).

Suppose a student set up an erroneous equation (3') \( -Mb_2 = Mg - T \) instead of the correct equation, (3). The set of equations (1), (2), (3'), (4), (5) and (6) is incalculable because of over-constraints. In such a case, a constraint (i.e., equation) is deleted to generate simulation. Since EBS should imply the student’s error, equation (3’) is out of the candidates for deletion. From the viewpoint of visibility, a more fundamental constraint in the real world is preferable for deletion. For example, when equation (5) is deleted, such EBS as shown in Figure 2(c) is generated, in which Block 1 and Block 2 overlap (because this equation maintains the relative acceleration between these blocks).
The student would easily recognize something is wrong with her/his equations. For another example, when equation (6) is deleted, such EBS as shown in Figure 2(d) is generated, in which the string between Block 2 and Block 3 shrinks (because this equation maintains the relative acceleration between these blocks). This EBS might have more suggestiveness than the former because it directly shows the error in the vertical acceleration of Block 2 (i.e., b2), which is included in the very error of this student (i.e., equation (3')). Thus, the effectiveness of EBS can be controlled and estimated following the above factors.

A simulator with the facility for such constraint handling is called “Robust Simulator (RSIM).” We have developed some RSIMs that check the consistency of a set of constraints and relax some of them by using heuristics, if necessary (Horiguchi & Hirashima, 2006; Horiguchi, Hirashima, & Forbus, 2012). In this practice, we embedded the facility for generating EBS into the learning environment, which is specialized in normal force in mechanics and always relaxes the rigid-objects-never-overlap constraint (described later).

In this practice, “normal reaction” is a learning target. Students are required to answer existing forces on a mechanical situation by drawing arrows of force. If the students make mistakes, EBS is generated by the students’ erroneous arrows. Most of them have enough visibility, reliability, and suggestiveness.

We introduce an example of EBS used in this practice, as shown in Figure 3. Students are shown a mechanical situation and are required to draw all the forces acting on the objects. They may make an erroneous drawing because of some misconceptions, which are regarded as the externalization of their erroneous idea. Based on the drawing, the acceleration of both objects is calculated with Newton’s second law, and their motion is simulated. In Problem 1 of Figure 3, for example, students often draw only the gravity acting on the block without the corresponding normal force as shown in Figure 4(a). In this case, the block sinks into the floor, as shown in Figure 4(b). We expect that such an unnatural phenomenon becomes a useful counterexample to students’ erroneous ideas and contributes to the correction of the errors with high and intrinsic motivation.

**Feature of EBS: From the viewpoint of constraint handling**

Simulating a model based on students’ erroneous ideas isn’t itself a new method. Many learning environments have been developed in which students construct a model and test it by simulation (Bravo, Jooldingen, & deJong, 2006;
Bredeweg, Linnebank, Bouwer, & Liem, 2009; Forbus, Carney, Sherin, & Ureel, 2004; Leelawong & Biswas, 2008). However, EBS is different in that it can bring out the pedagogical merit in the method that other systems can’t. That is, even when a model isn’t calculable because of serious conflict between constraints, our framework generates simulation by relaxing some of those constraints. Models by students often have conflicts between constraints whether the constraints are explicitly represented or not. By choosing the basic constraint(s) to be relaxed, EBS works as a counterexample (e.g., “if the model were correct, the block would sink into the floor”). Thus, students could connect the unnatural phenomena in EBS to their erroneous abstract concept.

Other systems, on the other hand, can’t generate simulation when a model isn’t calculable, so they usually give students corrective feedback on the representation of the model, which points out the erroneous parts of the representation (e.g., erroneous nodes or missing links in a concept map) and instructs students how to fix them. Even when a model is calculable, feedback is often given on its representation. However, such feedback doesn’t take advantage of simulation-based learning, that is, to connect abstract concepts to concrete phenomena. In fact, the limitation of such feedback is reported in several empirical studies (Leelawong & Biswas, 2008; Or-Bach & Bredeweg, 2013), and researchers usually try to combine it with other kinds of feedback, such as metacognitive feedback. In summary, EBS not only simulates the erroneous behavior of a model but also makes it as understandable as possible to students. Other systems rely on non-behavioral feedback when the implication of the behavior is hard to interpret.

Another possible way to deal with an incalculable model without losing the pedagogical merit of simulation is to give students another situation and make them model it. The situation should be modeled with the same target concept as the original situation. If the new model is calculable, the simulation can be generated. Otherwise, further situations are tried. However, as we discuss in the next section, using multiple situations for learning a concept might impose a heavy cognitive load on students. More seriously, since the original situation probably remains not understood, students might fail to understand the concept abstractly to form schematized knowledge.

Feature of EBS: From the viewpoint of analogy

It is well-known that most students have great difficulty in understanding “normal force.” For example, even after a teacher explains the concept in class, students often answer “only gravity” when they are asked what forces are applied on a book on a table. In this case, as we indicated in the previous section, usual experiment or demonstration can’t be used as a method to connect the concept to the phenomenon.

Using “bridging analogies” (Clement, 1993) is a very effective method to solve this problem. In this method, the gap between students’ correct belief and their misconception is bridged by a chain of intermediate analogous situations. For example, suppose students misunderstand the situation of a book on a table, which is called the “target.” First, a situation is introduced to them in which a hand is pushing down a spring on a table. Most students understand the spring pushes back up against the hand. This is called the “anchor.” Then, another situation is introduced in which a book is on a flexible board on a table. Students can understand the board pushes up the book because this situation is similar to that of the anchor. In addition, this situation is similar to that of the target. Therefore, students can connect the anchor to the target, to understand that a “normal force” is applied to the book from the table. Such an intermediate situation that shares features with both the anchor and the target is called a “bridging analogy.” It was reported that using bridging analogies in class effectively activated students’ discussion and scientific thinking, through which they understand the concept normal force (Clement, 1993).

The point of using bridging analogies is that, even in cases when usual experiment or demonstration doesn’t work, it explains the target concept (e.g., normal force) in connection with the phenomenon in some situation(s), instead of only explaining the concept itself. However, this method has the following difficulties: One must find an anchor situation that students can understand correctly. One must also find a chain of situations that bridge the gap between the anchor and the target situation. The gap between every pair of adjacent situations should be sufficiently small to be recognized that they are the same from the viewpoint of the target concept (e.g., in both the “hand-pushing-down-on-spring” and the “book-on-flexible-board” situation, students think upward [normal] force works). On the other hand, the gap should be sufficiently large to be a part of the bridge between the anchor and the target situation. This is a trade-off. Therefore, there is always the possibility that a student fails to recognize adjacent situations as the same (e.g., she/he might not accept upward [normal] force works in the “book-on-flexible-board” situation). In fact,
in the lesson designed by Clement (1993), complementary methods are used to compensate for this weakness, such as the guided discussion about the similarities and differences between situations, and the explanation with a microscopic model of normal force. Therefore, the outcome of the lesson should be regarded as the total effect of combined teaching methods including using bridging analogies.

In learning with EBS, on the other hand, the only situation a student must consider is the target situation. Instead of comparing different situations, she/he tries to explain the target by expressing her/his idea about the physical process(es) working in the situation. Any analogous situations aren’t necessary. More importantly, it is easy for a student to compare her/his different trials because the phenomena occur in the same situation. That is, they can be “well-aligned” (Gentner & Markman, 1997; Markman & Gentner, 1993). It is expected that students can easily see their differences that matter (“alignable differences”), which help them regulate the exploration by themselves to find the solution and form a conceptual understanding of the situation. Additionally, because a problem corresponds to a situation, it is relatively easy to design a sequence of problems of which situations are highly aligned (e.g., “a book on a table” and “a book on top of another on a table”). It is also expected that student can easily see their differences and apply their solution of a problem to other problems with appropriate modification, through which they might abstract the solution to form conceptual understanding and schematized knowledge.

Hypothesis

From the discussion in the previous two sections, we claim that using EBS enables one to design a learning environment that has the advantages of both discovery learning (Crews et al., 1997; Loh et al., 2000) and directed learning (Klahr, 2009). That is, since behavioral feedback is instantly given on students’ every trial (even when the model is inconsistent), using EBS promotes discovery learning. Making trials in the same situation makes it easier for students to self-regulate their exploration. At the same time, since each problem can be solved with EBS in a situation (not multiple situations), one can design a sequence of problems of which situations are highly similar and increasingly complex. Such a sequence of problems works as a guide that helps students generalize previous learning to solve a new problem of a (little) more complicated situation. This is directed learning. Therefore, we hypothesize the following: A carefully designed learning environment with EBS, in which behavioral feedback is instantly given on every trial in a problem and a sequence of well-aligned problems is provided, helps a student not only find the solution of a problem but also understand abstract concepts and form schematized knowledge by itself (without complementary methods).

In the following sections, we present the design of our system and the practice and its result with the system to verify this hypothesis.

Procedure of the practice

Purpose and method of the evaluation

The purpose of this practice is to evaluate the effect of EBS by comparing it to the effect of usual teaching from the viewpoints of transfer and retention of learning. To evaluate the transfer, we prepared three kinds of tasks. The first is a learning task that is composed of three problems, shown in Figure 3. They are used not only in the learning phase but also in all tests. Because these problems are used in the class, it is possible to gain a good score only by memorizing the correct answers. The second is a complex task composed of two problems shown in Figure 5, which consist of the same components as the problems in the learning task, but with a different number of components. Therefore, the problems are similar to the learning task, but it is impossible to gain a good score just by memorizing the answers. Generalization of number of components is required to solve the complex task. These problems were used in the post-test and delayed post-test but were not used in the pre-test and learning phase (because the teacher with 26 years of experience who conducted this practice judged the complex task was too difficult for the students as pre-test). The third is a transfer task composed of seven problems. Two problems are exampled in Figure 6. They consist of different components from the problems in the learning task. Therefore, in order to gain a good score, it is necessary to abstractly understand the relation between force and motion, not depending on the components. These problems were also used in the post-test and delayed post-test but not used in the pre-test and learning phase (because of the same reason as the complex task).
As for the examination of retention of the learning effect, we carried out the delayed post-test three months later, in addition to the post-test. We also had an interview with the students to understand how they solved the problems within one day after the day of the delayed post-test.

**Figure 5. Mechanical problems in complex task**

**Figure 6. Mechanical problems in transfer task**

**Learning environment with EBS**

For this practice, we used a learning environment that generates EBS based on students’ erroneous solutions in mechanics problems. In learning with the system, a student is provided with three problems of the learning task one by one and required to draw all forces acting on objects. After completing a drawing using a mouse, the student clicks the “done” button to see the behavior of the objects. This is called the learning phase.

In drawings, the points on which forces are acting are specified only in the neighborhood of objects’ centers or edges. The directions of forces can be specified only vertically or horizontally. The magnitudes of forces, that is, the length of arrows, can be selected from large, medium, and small. When the points, directions, and magnitudes of all forces are drawn correctly, natural motion is generated. When there are any mistakes, EBS is generated. A student can modify her/his drawing and see EBSs any number of times, until she/he completes the correct drawing for the current problem.

Occasionally, the motion of EBS is similar to natural motion. For example, when no forces are drawn in Problem 1, the block stays at rest on the floor correctly. This is the issue of visibility. In all problems used in this practice, the natural behavior is motionless. When an EBS doesn’t have enough visibility, the system directly indicates errors in the drawing. The methods to judge the visibility of EBS for moving objects with qualitative reasoning techniques were reported in Hirashima et al. (1998).

**Lessons**

This experiment was carried out with students who were in the first year of junior high school (grade 7). They were originally divided into three classes, with a total of 84 students. We assigned two of them to EBS class (54 students) and one to usual class (30 students). First, all classes worked on the pre-test. Then, all classes were provided with a lecture as usual in one class time (45 minutes) and only EBS class had additional learning time of another 45 minutes class, during which they solve three problems of learning task with the system. Therefore, the difference between
EBS class and usual class is learning with EBS. In the learning phase, each student used one system with one computer. EBS class worked on the post-test after the learning phase, while usual class after the lecture. Finally, three months later, all classes worked on the delayed post-test and were interviewed by the teacher. The pre-test included only learning task, while the post-test and delayed post-test included all three tasks. All tests were written tests.

All classes were taught by the same teacher who was in charge of science for junior high school students. In the learning phase, one assistant teacher was provided in addition to the class teacher. The teachers helped the students to use the system, while they didn’t give any hints about the solution of the problems. They also carefully observed the students’ activity. Especially, if they observed any student being seriously confused by unnatural phenomena, they were ready to stop her/him.

Results

Students’ learning activities

In the use of EBS system, all students were actively working on the problems. No student was confused by unnatural phenomena. When students saw unnatural phenomena, we observed that they were motivated to think about the cause of the error in their solutions. No students had any serious difficulties in using the system. All students completed the three problems correctly in the learning phase.

Results of scores

The results of the average scores are shown in Figures 7, 8, and 9. They show the effect of the conditions, the effect of the tests and the effect of the tasks, respectively. The statistical analysis is also been summarized in Tables 1, 2, and 3. The marking system for the tests was one point for one correct answer of an acting force; therefore the total mark for the learning task was 14 points, while for the complex task and the transfer task the marks were 19 and 30, respectively.

Figure 7 shows that in post- and delayed post-tests, and in all tasks, the scores of EBS class are higher than those of usual class (while, as shown below, there wasn’t significant difference between EBS and usual classes in pre-test). Especially, greater EBS’s effect was observed in delayed post-tests than in post-tests, and greater EBS’s effect was also observed in transfer tasks than in learning and complex tasks. Figure 8 shows that the score decreases in delayed post-tests were smaller in EBS class than in usual class. Figure 9 shows that the difficulty of tasks increases in the order of learning, complex and transfer task.

![Figure 7. Effect of conditions (Usual vs. EBS)](image1)

![Figure 8. Effect of tests (Post vs. Delayed post)](image2)
In summary, in post-test and delayed post-test, the scores of EBS class were higher than those of usual class as to all tasks. Especially, more EBS’s effect was observed in delayed post-tests than in post-tests, and so was in transfer tasks than in learning and complex tasks. Note that generalization of the learning result is necessary for answering the problems of complex and transfer task correctly. We can, therefore, conclude that the generalization and retention of the learning results were well done in the EBS class than in the usual class.

**Table 1.** Simple-simple main effects of class

<table>
<thead>
<tr>
<th>Learning task (full marks = 14)</th>
<th>Complex task (19)</th>
<th>Transfer task (30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>Post</td>
<td>Delayed</td>
</tr>
<tr>
<td>Usual class (n = 30)</td>
<td>3.6</td>
<td>12.7</td>
</tr>
<tr>
<td>EBS class (n = 54)</td>
<td>(3.4)</td>
<td>(2.2)</td>
</tr>
<tr>
<td>simple main effect</td>
<td>F = 2.058</td>
<td>F = 6.762</td>
</tr>
<tr>
<td>effect size of class</td>
<td>p &gt; .10</td>
<td>p &lt; .01</td>
</tr>
</tbody>
</table>

Average scores and SD
**Italic:** small effect (≥ .01 and < .06), underlined: medium effect (≥ .06 and < .14), bold: large effect (≥ .14) (using Cohen’s criteria [Cohen, 1998])

**Table 2.** Simple-Simple Main Effects of Task

<table>
<thead>
<tr>
<th>Learning</th>
<th>Complex</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>Delayed</td>
<td>Post</td>
</tr>
<tr>
<td>Usual</td>
<td>F = 133.212</td>
<td>F = 57.229</td>
</tr>
<tr>
<td></td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>EBS</td>
<td>F = 274.796</td>
<td>F = 520.901</td>
</tr>
<tr>
<td></td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
</tr>
</tbody>
</table>

**Table 3.** Simple-Simple Main Effects of Test

<table>
<thead>
<tr>
<th>Learning</th>
<th>Complex</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>Delayed</td>
<td>Post</td>
</tr>
<tr>
<td>Usual</td>
<td>F = 274.796</td>
<td>F = 520.901</td>
</tr>
<tr>
<td></td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>EBS</td>
<td>F = 750.349</td>
<td>F = 416.602</td>
</tr>
<tr>
<td></td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
</tr>
</tbody>
</table>

Three-factor ANOVA of 2 (class: EBS/usual) x 3 (task: learning/complex/transfer) x 3 (test: pre, post, delayed post) revealed the above observations have statistical significance. Because the interaction of the three factors was significant, and was so for all combinations of two factors, we tested the simple-simple main effect of each factor. Table 1 shows the simple-simple main effects of class factor and their effect sizes. There were significant differences in post- and delayed post-tests ($F = 6.762, p < .01$ and $F = 70.912, p < .001$ in learning task; $F = 8.628, p < .005$ and
$F = 89.047, p < .001$ in complex task; $F = 27.681, p < .001$ and $F = 48.678, p < .001$ in transfer task) while there wasn’t in pre-tests ($F = 2.058, p > .10$). Additionally, the effect sizes were greater in delayed post-tests, and in more difficult tasks ($\eta^2 = .0273$ for post-test and $\eta^2 = .281$ for delayed post-test in learning task; $\eta^2 = .0352$ for post-test and $\eta^2 = .354$ for delayed post-test in complex task; $\eta^2 = .109$ for post-test and $\eta^2 = .193$ for delayed post-test in transfer task). The effect size is calculated by squaring the correlation ratio, and indicated using Cohen’s criteria (small effect: $\geq .01$ and $< .06$, medium effect: $\geq .06$ and $< .14$, large effect: $\geq .14$) (Cohen, 1988). Tables 2 and 3 show the simple-simple main effects of task and test, respectively. According to the former, the differences of difficulty between tasks were significant ($F = 133.212, p < .001$ for post-test and $F = 57.229, p < .001$ for delayed post-test in usual class; $F = 73.173, p < .001$ for post-test and $F = 95.016, p < .001$ for delayed post-test in EBS class). According to the latter, the differences of scores between tests were also significant ($F = 274.796, p < .001$ for learning task, $F = 520.901, p < .001$ for complex task and $F = 226.940, p < .001$ for transfer task in usual class; $F = 464.900, p < .001$ for learning task, $F = 750.349, p < .001$ for complex task and $F = 416.602, p < .001$ for transfer task in EBS class).

**Results of interviews**

Within one day after the delayed post-test, the same teacher who conducted the practical use of EBS interviewed the students. During the interview, teacher only asked them to explain how they solved all seven problems in the transfer task. A total of 42 students from the EBS class and 27 students from the usual class were interviewed. The rest of the students were absent from the interview because it was conducted out of class time. The average scores of the transfer task in the delayed post-test were 18.3 in EBS class and 11.9 in usual class.

The most frequent explanation in usual class was “balance of forces.” If the students explained they tried to balance forces to solve a problem, we categorized their explanations as “force balance.” In usual class, 4.6 problems in 9 problems in average were categorized as “force balance.” Some of the students answered, “they had no idea” or “they found only gravity in a problem” (3.2 problems in average). We categorize these answers as “no idea/only gravity.” No students referred to the motion of objects (Table 4).

| Table 4. Numbers of problems explained with categories |
|-----------------|---------------|-------------------|
|                  | Usual class   | EBS class         |
|                  | (n = 27, ave = 11.9) | (n = 42, ave = 18.3) |
| force balance    | 4.6           | 1.3               |
| force and motion | 0             | 4.8               |
| no idea/only gravity | 3.2       | 1.0               |

| Table 5. Correlations between delayed post-test scores and numbers of problems explained with categories (Spearman rank correlation test) |
|-----------------|-----------------|-----------------|
|                  | Usual class (n = 27) | EBS class (n = 42) |
|                  | delayed post-test score (ave = 11.9) | delayed post-test score (ave = 18.3) |
| force balance    | $r = .84$ ***   | $r = -.22$ ***  |
|                  | $p = 2.0E-05$   | $p = .15$ ***   |
| force and motion | $r = .84$ ***   | $r = -.44$ ***  |
|                  | $p = 1.9E-05$   | $p = .004$ ***  |

*** strong positive correlation, ** medium positive correlation,
### strong negative correlation, # medium negative correlation

The relationship between the delayed post-test scores and the number of problems explained with “force balance” was investigated using Spearman rank correlation test (Table 5). There was a strong positive correlation between the two variables, $r = .84, p = 2.0E-05$. There was a strong negative correlation between the delayed post-test scores and that with “no idea/only gravity”: $r = -.84, p = 1.9E-05$. These results suggest that the students in usual class solved the problems mainly with “force balance” but did not consider the relation between force and motion.
In EBS class, by contrast, the most frequent explanations were “connection between force and motion” (4.8 problems in average). If the students’ explanations suggested that they mentally simulated the effect of forces on objects’ motion, we categorized them as “force and motion” (for example, “Since the gravity was drawing the block downward, I added the upward force for it not to fall”). On the other hand, “force balance” and “no idea/only gravity” was 1.3 and 1.0 problems in average, respectively (Table 4). There was a medium positive correlation between the delayed post-test scores and the number of problems explained with “force and motion,” \( r_s = .48, p = .0018 \). There was a medium negative correlation between the delayed post-test scores and that with “no idea/only gravity,” \( r_s = -.44, p = .004 \). The correlation between the delayed post-test scores and “force balance” was not significant, \( r_s = -.22, p = .15 \) (Table 5). These results suggest that the students in EBS class solved the problems mainly by considering the relations between force and motion.

These results suggest that the awareness for the relation between force and motion plays a crucial role in the difference between usual class students and EBS class students. Though the usual class students who didn’t consider the relation between forces and motions also got better delayed post-test scores compared with pre-test scores, their delayed post-test scores were significantly lower than EBS class students’ in all tasks. Additionally, in all tasks, the decreases between post-test and delayed post-test were smaller in EBS class than in the usual class. As indicated previously, these results suggest that EBS class students got qualitatively better understanding. We suppose it is because they considered the relation between forces and motions. EBS is a method to visualize errors by showing unnatural behaviors of objects based on students’ errors. In this practice, motion of objects was generated by using students’ erroneous forces. Therefore, it is suggested that students not only were aware of errors by observing unnatural behaviors, but also noticed the importance of the connections between force and motion. This can be considered as a normal and also general approach to solve problems in this practice. These results suggest that EBS is useful to promote students to consider the connections between forces acting on objects and their motions, although we have to examine further experiments in mechanics or other learning domains.

Discussion and concluding remarks

We previously hypothesized that a carefully designed learning environment with EBS helps a student to not only find the solution of a problem but also understand it conceptually to form schematized knowledge. The results of our practice strongly suggest this hypothesis is true. Though the teachers didn’t give any hints, all students of EBS class solved all the problems correctly by using the system. Additionally, greater EBS’s effect in delayed post-tests and transfer tasks suggests that the students of EBS class gained a conceptual understanding of which quality was confirmed in the interview with them. (To ensure fairness, we must point out that the EBS class spent more time on learning than usual class, the influence of which should be investigated in future work.)

Note that the system used in this practice doesn’t provide any “intelligent” help. It merely simulates objects’ behavior based on students’ drawing of forces by violating the “rigid-objects-never-overlap” constraint, if necessary. Without corrective feedback on the drawings, students could complete them correctly only by observing the simulation. Without metacognitive feedback, students could generalize and schematize their knowledge for future use. This outcome was, we think, obtained from the design of the system in which (a) instant feedback was given on each trial in the same situation that promoted the alignment of the trials in a problem, and (b) because the situations of a sequence of problems are highly aligned, the alignment of the solutions of the problems is promoted. Additionally, this design became possible because of the features of EBS: (1) The behavior of models can be always simulated even if they aren’t calculable, (2) The unnaturalness of EBS is controlled taking into account its visibility, reliability, and suggestiveness, (3) Students can receive meaningful behavioral feedback on their every trial in a situation (it needn’t be changed), and (4) Such correspondence between a situation and a problem makes it easier to design a well-aligned sequence of problems. Though EBS is on the lines of “modeling and simulation” methodology, its own feature presents an alternative (or a complement) to the existing methods for designing assistance in SLEs. We verified that our method does work for promoting students’ deeper understanding.

Two issues remain to be further investigated: applicability to more complex tasks and metacognitive skills students can acquire. First, the tasks used in this practice were relatively simple because they dealt with mainly one (critical) concept, “normal force.” The interface used for constructing models was also made very simple. Such simplicity might help students concentrate on comparing trials and problems to generalize what they learned. In other learning environments, more complex modeling tasks are dealt with that include a lot of parameters and difficult concepts.
(such as “stock and flow”). It is our important future work to apply EBS to other domains and more complex tasks. Second, since the results of this practice suggest the students of EBS class generalized their knowledge well, we can expect that they acquired some metacognitive skills in reflection. For generalizing and schematizing the knowledge they learned, students might reflect their solutions and problem-solving processes and acquire skills such as how to distinguish between problem-specific and generalizable elements, how to compare problems/solutions, and how to find the essence of a problem/solution. Some of these skills would be domain-independent. Their higher performance in transfer tasks suggests they had got such skills, while they might do well without such skills because our system was designed to strongly guide their learning. The detail is unknown. More elaborate experimentation is necessary to clarify this.

Currently, this is a case study. It is imperative to try EBS in other domains and more complex tasks to investigate the above issues, which would clarify its applicability, (metacognitive) learning effect, and limitation, and how it works in combination with other teaching technologies.

Acknowledgments

The authors would like to thank Prof. Dedre Gentner and Prof. Ken Forbus for their kind advices and important suggestions, which greatly improved this paper. We also thank the members of their research group, the discussion with whom was very fruitful.

References


Forbus, K. D., Carney, K., Sherin, B., & Ureel, L. (2004, August). *Qualitative modeling for middle-school students*. Paper presented at the 18th International Workshop on Qualitative Reasoning (QR04), Evanston, IL, USA.


Self-Observation Model Employing an Instinctive Interface for Classroom Active Learning

Gwo-Dong Chen¹, Nurkhamid¹,², Chin-Yeh Wang³*, Shu-Han Yang⁴ and Po-Yao Chao⁵

¹Department of Computer Science and Information Engineering, National Central University, Jhongli City, Taiwan // ²Faculty of Engineering, Yogyakarta State University, Yogyakarta, Indonesia // ³Research Center for Science & Technology for Learning, National Central University, Jhongli City, Taiwan // ⁴Department of Hospitality Management, Chien Hsin University, Jhongli City, Taiwan // ⁵Department of Information Communication, Yuan Ze University, Jhongli City, Taiwan // chen@csie.nchu.edu.tw // nurkhamid@uny.ac.id // chinyea@db.csie.nchu.edu.tw // yoko@uch.edu.tw // poyaochao@saturn.yzu.edu.tw

*Corresponding author

(Submitted March 18, 2013; Revised August 1, 2013; October 7, 2013)

ABSTRACT

In a classroom, obtaining active, whole-focused, and engaging learning results from a design is often difficult. In this study, we propose a self-observation model that employs an instinctive interface for classroom active learning. Students can communicate with virtual avatars in the vertical screen and can react naturally according to the situation and tasks. Students can also immediately observe themselves mixed with a virtual environment and therefore reflect on the necessary improvements. With the designed system, potential advantages such as motivation, enjoyment, context for situated learning, engagement, social collaboration, and role-playing might arise. To examine the idea, we conducted a case study with 60 fourth grade elementary students to investigate learners’ behavior as performers and peer audience and their perception of body movements and speech commands in the designed learning environment. The results show that the students’ image in the vertical screen affected the students and the peer audience positively. Moreover, they positively perceived their competence and their enjoyment after they performed contextual learning activities through the body movement interface.

Keywords

Active learning, Self-observation, Instinctive interface, Body movements, Classroom learning

Introduction

Active learning involves two things for learners: the learners are actively involved in the learning process—they learn by doing, actively participating in constructing knowledge, resulting in deeper and more persistent knowledge—and are being engaged in learning (Prince, 2004). These are two core elements of active learning, i.e., activity and the promotion of engagement. John Dewey coined the phrase, “learning is doing,” and later scholars identified, “knowing is doing” (Mishra, Worthington, Girod, Packard, & Thomas, 2001). More specifically, active learning may include collaborative, cooperative, and problem-based learning (Prince, 2004). Mayer (2009, pp. 21–24) categorizes active learning into two types: cognitively and physically (behaviorally) active learning. He primarily stresses utility in cognitively active learning, and this view is generally accepted in cognitivist learning theory. Indeed majority notion of active learning with technology has been of cognitive type. One example is BeyondShare (Kao, Lin, & Sun, 2008), which is an Internet-based learning environment designed to promote active learning that involved cognitive activities such as reflection, synthesis, and meaningful learning, and employed a concept map. However, when body movement is considered as an integral part of the learning process, as it is according to the embodied cognition view, properly designed physically active learning experiences should contribute to cognition and learning. For this to occur, the learning environments need to explicitly support physical activities, and current technologies have provided diverse possibilities for achieving physically active learning.

In a typical classroom setting, students can learn by passively sitting on their chairs and using minimal body movement. However, implementing body movements, rather than simply sitting for extended periods benefits learning, facilitating effective memory, fun, and activity (Jensen, 2000). Here he focuses the discussion on physically rather than cognitively active learning. James et al. (2002) provided evidence that actively manipulating objects in a virtual reality environment improved object recognition. These indicate that a learning environment must be designed to support physically active learning that comprises body movements.
In this study, we propose a self-observation model that comprises an instinctive interface for active learning, facilitating engagement, collaboration, and physical activity.

**Related work**

**Active learning with technology**

Instead of applying passive teaching and learning practices in the classroom, contemporary educators have increasingly focused on active teaching and learning. However in traditional classrooms, students learned information, blindly followed guidelines, and learned a mass of disconnected facts, but were provided no clues or context by their teachers, who authoritatively delivered knowledge (Sawyer, 2006). This practice is called instructionism, and eventually, “scientists discovered that instructionism was deeply flawed” (Sawyer, 2006, p. 2). In the early era of computer use in classrooms, their use was based on instructionism or behaviorism. However, “the computer should take on a more facilitating role, helping learners have the kind of experiences that lead to deep learning” (Sawyer, 2006, p. 8). For example, in the electronic textbook designed as an instructivist learning tool, although it possessed certain advantages compared with traditional lectures, Herrington and Standen (2000) found it boring because students experienced difficulty to apply the knowledge in real life. According to Herrington and Oliver (2000), the tool lacks authentic contexts and activities.

Harasim (2011, p. 14) noted that epistemologically, behaviorism is in one group with cognitivism. They are of objectivism perspective assuming absoluteness of knowledge based on reality; hence, learning is acquiring knowledge. Constructivism, however, is a subjectivism paradigm, which assumes that knowledge is constructed; hence, learning is doing and creating meaning. Learning in behaviorism and cognitivism tends to be passive learning in a sense that it is more instructor-centered, while learning in constructivism promotes active learning (Harasim, 2011, pp. 58–61). Active learning facilitates knowledge acquisition, critical thinking skills, problem solving, and independent thinking. Passive learning is disadvantageous, allowing teachers to become “the sage on stage,” and the lack of interactivity between teachers and students engenders quiet students (Wang, Shen, Novak, & Pan, 2009). However, obstacles arise when implementing active learning, because teachers tend to imitate the way they were taught, and most were taught as passive learners (McManus, 2001).

Technology can be applied to active learning (Shapiro, 1998; Wang et al., 2009) and may provide interactivity between learners and content knowledge (Pahl & Kenny, 2008). Jonassen et al. (2007, p.8) asserted that technology should support meaningful learning by providing a “context to support learning by doing” or active learning. Technology should no longer be used to transmit knowledge or teach. Instead, it should be used to engage students and promote critical thinking. For example, when hypertext is applied to active learning, learners can actively choose links, allowing them to effectively integrate and deeply understand knowledge, especially when the hypertext is ill-structured. In the ill-structured system, learners require cognitive effort to process the information contained in the links, understand the material, and remain oriented (Shapiro, 1998). As another example, Wang et al. (2009) designed a mobile learning environment that used real-time interaction to increase student engagement. When a video camera is used in active learning, learners can engage in meaningful activities, view their own performance, express themselves, and become creative and collaborative (Broady & Duc, 1995). Digital learning playground has also been used in active learning (Chen et al., 2013; Chen, Chuang, Nurkhamid, & Liu, 2012; Wang et al., 2010) but previous studies have not considered body movements (e.g., gesture-based interactions) or self-observations. Ozcelik and Sengul (2012) emphasized that using gestures to understand concepts can improve learning. Although active learning that leverages performance and body movements is beneficial, few researchers have explored this by using contemporary classroom technological tools.

**Self-observation for learning**

When an individual can see himself or herself while learning, it may positively affect corrective feedback, introspection, and the sense of self-assessment resulting in learning gains. Self-observation (seeing oneself) positively affects learning (Fireman, Kose, & Solomon, 2003). Observing self-action (Gupta & Bostrom, 2012) is called enactive learning. Observing the actions of others while they learn is called vicarious learning, because learners gain knowledge by modeling the behaviors of others. When foreign language students learn English orally,
Chen (2008) noted that self-assessment promoted self-confidence, additional effort, awareness of competence, and mastery in learning, suggesting that self-assessment facilitates autonomy and lifelong learning skills.

Schunk (2011) stated, “Learning occurs either enactively through actual doing or vicariously by observing models perform (e.g., live, symbolic, portrayed electronically). Enactive learning involves learning from the consequences of one’s actions.” (p. 121). Thus, self-observation can be enactive because it displays the consequences of actions onscreen. However, self-observation can also be vicarious when students feel that the technology mirrors themselves in a learning context, and this observation feels similar to watching a television.

**Instinctive interface for performing physical activity**

The story of human-computer interaction evolved from command line interfaces, to graphical-user interfaces, and then to natural user interfaces (Villaroman, Rowe, & Swan, 2011). In natural user interfaces, users interact with computing devices in intuitive and instinctive ways. Learning to use the interface is expected to be quick and easy. Furthermore, this instinctive interface provides a greater level of control compared with a keyboard and mouse (Birchfield, Ciufo, & Minyard, 2006; Shiratuddin & Wong, 2011). The Kinect device provides this type of interface, enabling users to perform physical activity by using embodied interaction. Comfortable movement is possible because Kinect is wireless, requiring no attachment between the user and the device (Ozcelik & Sengul, 2012).

Role-playing activities that involve some degree of physical activity are prevalent in current pedagogies. This type of learning usually impresses learners because they perceive, act, interact with things and events in their surroundings, and their bodies link their minds to the world (Atkinson, 2010; Barsalou, 2010). Furthermore, learning consists not only of mental activities, but it also needs the learner's body to perform and interact with the learning environment; this implies embodied cognition (Atkinson, 2010). Body movements may add contextual meaning to activities that were traditionally dominated by printed materials (Gee, 2003). Kirsh (2013) maintained that performing physical activities can enable learners to acquire more knowledge compared with simply watching others; doing may result in more effective learning compared with watching.

A similar study on using body movements for learning could be that conducted by Tolentino et al. (2009), who used a horizontal ground (floor) projection as a mixed-reality science learning space. However, by using a traditional drama stage metaphor that is similar to a traditional classroom setting, our design includes two spaces: a horizontal and a vertical screen. This does not deviate much from traditional classroom culture, in that a classroom has a blackboard facing the classroom as the center of attention—which corresponds to the vertical screen in our system. Moreover, it uses current technologies in schools, such as projectors, touchable whiteboards, and desktop computers. Adopting an authentic learning environment of a Digital Learning Playground platform (Wang et al., 2010), we further enhance the platform to include self-observation feature and embodied interaction. The following section presents additional details regarding the proposed learning environment.

**The learning environment**

The environment implements features described in the previous section, such as active learning, self-observation, and instinctive interface. Figure 1 shows the interaction setting consisting of various parts, including a vertical display, a horizontal touch-enabled board, and an action zone for the performer. Both screens display beams from the upper and front projectors. Kinect detects commands through the performer’s body movements and voice, and renders the performer’s image on the vertical screen to place and interact with learning objects. The orchestrator of the interaction between users and the system is a desktop computer. Peer students and the teacher may stand on the left and right sides of the horizontal board (Figure 2).

The horizontal screen is a shared touch-enabled space to provide menus for a teacher to preview the lesson and to explain how to play the game. It can provide a layout of board-game items, so that a group of learners can plan when collaborating on a role-playing assignment. It can also be a virtual keyboard, for example, for one of the group members to type words with during a game session. The vertical screen provides simulative and task-situational context, for example, for learners playing a cook’s role in a virtual kitchen by using their body movements. The screen can also allow learners to see virtual events that they can authentically interact with in a situated manner.
Enabling a Kinect sensor for Xbox 360 or other motion/voice controllers can facilitate instinctive interactions. Kinect detects the learner’s body movement and facilitates voice commands, for example, when choosing a type of drink. It can also produce the learner’s image on the vertical screen where the ongoing task scenarios appear. The learner, as a live task performer, with his or her body movements and/or speech commands can role-play and interact with context objects.

Figure 1. Self-observation model with instinctive interface for classroom learning

The potential pedagogical added values of the design may result from its built-in features and size. The vertical and horizontal screen sizes enable teachers and students to stand around the horizontal screen to conduct social collaborations and corrective feedback sessions. Including gaming and role-playing applications in the system should increase the motivation, enjoyment, and engagement of students. The group game function and animated multimedia support, which incorporates body movements, should engender “ownership and voice in the learning process,” and “learning in social experience” by using “multiple modes of representation” “in realistic and relevant contexts” (Honebein, 1996, pp. 11–12). When the learning context is provided onscreen during role-playing, it combines real-world learning and the traditional classroom, engaging students as if in real locations without requiring that they travel to experience the learning context. Additional pedagogical value may result from the support of vicarious (by observing others) and enactive (by doing) learning. This should facilitate a meaningful classroom learning experience for students.

Three assumptions motivate the creation of the learning environment. First, direct embodied experiences can facilitate learning as a way of constructing knowledge. In Freiler (2008), embodied learning relied on the role of the body in knowledge construction to produce deep, experiential learning. Recent focus has been paid to incorporating body movements into interface designs for learning purposes (Johnson-Glenberg, Birchfield, Savvides, & Megowan-Romanowicz, 2010; Lee, Huang, Wu, Huang, & Chen, 2012; Ozcelik & Sengul, 2012; Tolentino et al., 2009). Second, multimodal learning optimizes knowledge construction and thus improves learning, whether for languages (Macedonia & Knösche, 2011) or science and engineering (Johnson-Glenberg et al., 2010; Tolentino et al., 2009). Multimodal learning relates to knowledge that is represented using various sensory modes such as sight, sound, movement, and touch (Birchfield et al., 2006). Multimodal learning works because cognitive processing is multimodal (Atkinson, 2010). The external representations that learners experience through various perception and knowledge processing channels improves the potential to learn (Tolentino et al., 2009). Moreover, acting out knowledge facilitates verbal learning because it involves encoding and deep processing (Macedonia & Knösche, 2011). Third, because extrapolating the findings of offline self-observation from video recordings benefits learning
observing the live actions of a person while learning should also be beneficial, facilitating instant, rather than delayed feedback from teachers and peer learners.

Following these assumptions, we designed a new learning interface. The proposed interface should be engaging and enjoyable because, rather than simply using a mouse and a keyboard, we designed the system to respond to body movements and voice commands. Consequently, the body movement and voice command interfaces should provide multimodal learning channels. Based on these three assumptions, this design is feasible. However, certain caveats exist; for example, certain learners may resist or feel uncomfortable with embodied learning (Freiler, 2008), or exhibit an initial shyness to act (Birchfield et al., 2006).

This paper may exemplify the potential educational applications of Kinect (Hsu, 2011). Kinect represents a current embodied sensor technology that is ready to catch learners’ body movements, actions, and voice. To the best of our knowledge, a new feature of the proposed design is that it provides learners with the possibility to see themselves in action while they are involved in a learning scenario. The learner’s animated images appear on the vertical screen, which they and their peer learners can see (Figure 1). We call this the self-observation model.

With the proposed model we want to know about the learners’ behavior. Therefore, this study raised the following research questions: In the proposed design, what is the effect of the students’ live performance and of their speech commands and body movements on learning behaviors? In this case, we limit our observations of the students’ learning behaviors to their intrinsic learning motivation measured with the Intrinsic Motivation Inventory (IMI) survey (Schunk, Pintrich, & Meece, 2007). We chose the IMI as an instrument to understand students’ learning behavior, and specifically, whether they enjoy learning by using the system we designed. Thus, we can use the results to infer whether the design is acceptable. We hope the proposed design produces a positive effect on learning behavior.

Evaluation methods

To examine the proposed model, we implemented the system as a case study for children learning English as a second language in Taiwan. Data collection includes activity videotaping, the IMI survey (Schunk et al., 2007), and interviews.

The aim of this study was to understand the effect that the students’ live performance and their speech commands and body movements have on learning behavior, which their learning motivation represents. In this case, we believe that embedding students’ live performance in ongoing-task scenarios not only relates the “actor” and “audience” to apprehended knowledge in a real context, but also engenders productive learning performance and a positive atmosphere. Thus, we hypothesize the following:

- Students’ live performance in the ongoing-task scenarios has positive learning effects on the performer and other peers (Hypothesis 1).
- One of our assumptions was that using speech commands and body movements to act out in the task enhances learning and creates an unusual fantasy, due to near-authenticity of the role being played, not only among performers, but also among the peer audience. Therefore, we postulate that using speech commands and body movements to interact with the objects in the scenario positively affects the learning of the performer and the peers (Hypothesis 2).

Participants

The participants of this study were 60 fourth grade elementary students from two classes in Taiwan, where the experiment took place. They learn English as a second language. The English teacher of the school, who was in charge of these two classes, confirmed that the subjects’ English background knowledge was similar.
Learning material

The learning objective of the class session was about knowing food names and how to order the foods; the learning material was taken from the students’ English textbook. At the start, teachers explained to the participants what learning tasks and goals they were to accomplish. Each group of students, which comprised five people, needed to collaborate in planning their actions and managing their resources on the shared game table when their turn to play came. In this case, students ran a fast-food restaurant, and needed to decide their daily menu and staff arrangements (role assignment). Based on the assigned roles, each participant played serving a virtual customer who appeared on the vertical screen. The student playing the role of a counter assistant was in charge of taking orders from the virtual customers at a counter scenario. The starring student’s live image appeared on the screen. Meanwhile, the student acting as a trainee would listen, watch the conversations, and then input guest orders using a virtual keyboard on the game table. Afterward, the student starring as a cook used speech commands and body movements to make food based on the guest’s order in the kitchen scenario. For example, a cook had to physically flip the patty to make burgers in the virtual kitchen and respond to the beverage machine to refill a drink.

Procedure

Sixty students were divided into two 30-student groups: the experimental and control groups. The selection is based on which class they belonged to. Each group was divided into six 5-student gaming groups. Two classes took place for 80 min each for the learning activity, and were devoted to 20 min for evaluations, including interviews and questionnaires. The same teacher facilitated the process during the teaching phase. We randomly picked one class as the experimental group and another as the control group. The students of the experimental group could see their classmates’ image in the scenario; conversely, the students in the control group would not appear on the screen while acting out the conversation with the characters.

To collect data, we used videotaped activities, the IMI survey, and interviews. The IMI questionnaire was used to examine subjects’ intrinsic learning motivation with respect to the following aspects: interest/enjoyment, effort/importance, value/usefulness, perceived competence, and pressure/tension. The scores ranged from 1 (lowest) to 7 (highest). The interviews were also conducted to understand their perceptions of the learning experience.

Results

We observed that the experimental group students executed their tasks smiling (75%) and made funny faces and hand gestures (Table 1). This body language indicates their curiosity toward their images in the scenario and the learning content. Regarding the control group, they were calmer while they talked to the virtual customers. The students in the control group said the following in the interview: “I would rather not to be a counter person because I don’t like to speak out” and “I can learn how to talk to a customer, but it is not so fun doing that.” In addition, the result from the interest dimension of IMI shows that the control group demonstrates less interest in role-playing (5.81) than does the experimental group (6.38).
To ascertain active learning through observation, the performers in the control group were more likely to perform individually, and made no significant effort to seek assistance from their teammates (0%). This phenomenon highlights the interesting relationship between the performers and the teammates. Both groups reported that they exerted considerable effort and felt competent (effort and perceived competence) after performing the tasks. In addition, all the students agreed that learning by interacting within relevant contexts helped them relate to new learned knowledge (these gave a score of over 6 on the value/usefulness dimension). One of the performers in the experimental group expressed that “seeing myself in the counter dealing with the customers really engaged me in this role.”

To investigate how students perceive themselves when performing on screen and how this influences them and others, the class observation adopted two primary focuses:

- the learning behaviors of the performers, such as anxiety, active learning, enjoyment, and perception as a demonstrator, and
- the learning behaviors of others, such as attention or focus, imitative behaviors, and error correction through observation.

On average, the learners perceived that the environment was enjoyable, valuable, and essential when they used body movement and speech commands to perform contextual learning tasks. The observational data, interview, and the result of IMI were cross-analyzed in the sections that follow.

The behaviors of the performers

The interviews revealed that two students became nervous. They gave the following reasons: “I was nervous because I was afraid to give a wrong answer” and “I got nervous because I wasn’t sure my response was right.” We inferred that most students were unaware of their nervousness, and that this did not affect their performance. In addition, in contrast to their outward appearance in the scenario, having to give proper responses placed the performers under pressure. Regarding the students as task focus, observations indicated that the prompts and the assistance from the instructor and the peers also helped ease their tension of arising from the task focus.

Table 1 presents a summary of behaviors the students displayed when seeing themselves in the scenarios (Hypothesis 1). The five observational points are anxiety, active learning, enjoyment, and being perceived as demonstrators.

Table 1 provides the results obtained from our observation of anxiety levels produced from showing the students’ images in the scenario. The observational data show that both groups of performers displayed simple nervousness. The results for the pressure/tension dimension of the IMI questionnaire show that both groups’ tension is low (the highest average score is 7, whereas the lowest is 1; Table 2). Table 2 lists the results of the t-test between the CG and EG for all variables; no significant difference is indicated between the IMI variables.

<table>
<thead>
<tr>
<th>Anxiety</th>
<th>EG</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncomfortable swaying</td>
<td>8%</td>
<td>30%</td>
</tr>
<tr>
<td>Stiff body posture</td>
<td>50%</td>
<td>40%</td>
</tr>
<tr>
<td>Constant blink</td>
<td>17%</td>
<td>8%</td>
</tr>
<tr>
<td>Deep breathing</td>
<td>8%</td>
<td>40%</td>
</tr>
<tr>
<td>Frowning</td>
<td>17%</td>
<td>0%</td>
</tr>
<tr>
<td>Active learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active responding</td>
<td>58%</td>
<td>30%</td>
</tr>
<tr>
<td>Calling for support from teammates</td>
<td>17%</td>
<td>0%</td>
</tr>
<tr>
<td>Enjoyment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smile</td>
<td>75%</td>
<td>10%</td>
</tr>
<tr>
<td>Showing curiosity:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making extra movement</td>
<td>58%</td>
<td>-</td>
</tr>
<tr>
<td>Making funny faces</td>
<td>42%</td>
<td>-</td>
</tr>
<tr>
<td>Hand movements indicating touching the virtual objects</td>
<td>33%</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1. The learning behaviors of the performers (n = 60)
Perceived as a demonstrator

<table>
<thead>
<tr>
<th></th>
<th>EG</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standing up straight</td>
<td>58%</td>
<td>10%</td>
</tr>
<tr>
<td>Looking at the audiences</td>
<td>25%</td>
<td>0%</td>
</tr>
<tr>
<td>Speaking up</td>
<td>25%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Table 2. The t-test analysis of IMI of performers (n = 60)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Avg</th>
<th>Std. dev.</th>
<th>Avg</th>
<th>Std. dev</th>
<th>Sig. (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure/Tension</td>
<td>3.69</td>
<td>.48</td>
<td>3.28</td>
<td>.40</td>
<td>.570</td>
</tr>
<tr>
<td>Effort/Importance</td>
<td>6.50</td>
<td>.81</td>
<td>6.43</td>
<td>.82</td>
<td>.703</td>
</tr>
<tr>
<td>Perceived Competence</td>
<td>5.06</td>
<td>.37</td>
<td>4.97</td>
<td>.77</td>
<td>.105</td>
</tr>
<tr>
<td>Value/Usefulness</td>
<td>6.28</td>
<td>.51</td>
<td>6.47</td>
<td>.49</td>
<td>.700</td>
</tr>
<tr>
<td>Interest/Enjoyment</td>
<td>6.38</td>
<td>.23</td>
<td>5.81</td>
<td>.37</td>
<td>.059</td>
</tr>
</tbody>
</table>

Observations of the experimental group show notable behavior: some performers perceived themselves as knowledge demonstrators. Many EG performers were well aware they were performing, and intended to stand up straight (58%) and speak up (25%). Referring to the interview, six students voiced that they liked being called up to act as a model. Compared to the control group, the demonstrators’ traits were not apparent in this experiment.

The behaviors of others

In this section, we concentrate on how the non-performers participated while their teammates were performing their tasks. Table 3 presents a summary of the observations of the behaviors of the others (Hypothesis 1). Our observational phases are their attention-focus, imitative behaviors, and active learning (supporting the performer).

Table 3. The observational points of learning behaviors of the others (n=60)

<table>
<thead>
<tr>
<th>Attention-focus</th>
<th>EG</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent-minded in the task</td>
<td>3%</td>
<td>12%</td>
</tr>
<tr>
<td>Turning head toward the performer whenever performer is speaking or acting</td>
<td>97%</td>
<td>32%</td>
</tr>
<tr>
<td>When the performance is showing, the other participants are:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-participated (facing the instructor when the instruction is delivered, watching the interaction shown on the screen, observing performer’s performance)</td>
<td>70%</td>
<td>41%</td>
</tr>
<tr>
<td>Looking at the game table</td>
<td>7%</td>
<td>40%</td>
</tr>
<tr>
<td>Looking at other places</td>
<td>23%</td>
<td>19%</td>
</tr>
<tr>
<td>Immediate attention right after the image is appeared</td>
<td>97%</td>
<td>-</td>
</tr>
<tr>
<td>Imitative behaviors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imitating performers’ action</td>
<td>17%</td>
<td>12%</td>
</tr>
<tr>
<td>Simultaneously repeating the required speech</td>
<td>50%</td>
<td>18%</td>
</tr>
<tr>
<td>Active learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supporting the teammate performers (suggesting, error-correcting)</td>
<td>83%</td>
<td>35%</td>
</tr>
<tr>
<td>Supporting the performers (not aware of team competition)</td>
<td>43%</td>
<td>21%</td>
</tr>
<tr>
<td>No support</td>
<td>21%</td>
<td>70%</td>
</tr>
</tbody>
</table>

As the observational data of the peers’ focus of attention show, having the students perform in the designed scenario successfully caught the other learners’ attention (EG: 97%; CG: 32%). This is a positive learning effect because the learners were focused on the newly learned knowledge situated in an ongoing task of their classmate’s role. These live performances serve as learning inputs or examples that an instructor hopes are delivered to the other participants. Moreover, this design has motivated the peer audience to imitate the performer’s performance (EG: 50%; CG: 18%) and to show that they highly support the performers by correcting errors and making suggestions (EG: 83%, 43%; CG: 35%, 21%). Conversely, the students in the control group acted as passive knowledge receivers, and 70% of them did not show support. Approximately 43% of the peer audience was willing to help the performers, despite belonging to the other team. The interviews on the audience’s feedback of the performance revealed these statements: “I think it was very interesting when my classmate appeared on the scene” and “I wanted to know...”
whether he responded correctly or incorrectly, and I tried to learn from their mistakes and deliver a good performance during my turn.”

The behaviors of the performers in physical learning

Regarding Hypothesis 2, Tables 4 and 5 show the IMI results and the observational data regarding engagement, active learning, and fulfillment. The results for intrinsic motivation show that the motivation of both classes through the embodied interaction is high, and that it coheres with the behaviors that the observational data reveal. Regarding the performers’ behaviors, both classes displayed body expressions that demonstrated their ability to manipulate virtual objects, drawing their great curiosity to test the objects on the screen (Class A: 88%, Class B: 82%). A good number of students were highly contented and smiling (Class A: 63%, Class B: 91%) after completing the task. In relation to the non-performers, they were attracted to the task performance; some students even gave suggestions and shared the same hand movements with the performers (Figure 3). Table 4 shows a one-sample t-test that assumes a mean value of 4, the midpoint of the Likert scales from 1 (lowest) to 7 (highest), to ascertain the significance of learner responses to interacting with the system. In one sample t-test, we compared between the mean score of the sample and a known value, here the midpoint of the Likert scales, as other researchers did (for example, Jacucci et al., 2010; Mason, 2011). The results attained significance for effort/importance, value/usefulness, and interest/enjoyment, indicating that, on average, when body movement and speech commands were integrated in the system, learners found the system important, useful, and enjoyable. However, the results for the pressure/tension dimension were non-significant.

| Table 4. The t-test analysis of IMI in physical learning (n = 60, test value = 4) |
|-------------------------------|--------|--------|----------------|
| Variables                     | Avg    | Std. dev. | Sig. (p)     |
| Pressure/Tension              | 2.97   | 1.32    | .007 ns     |
| Effort/Importance             | 5.14   | 1.22    | .002 *      |
| Value/Usefulness              | 6.05   | 1.20    | .000 ***    |
| Interest/Enjoyment            | 5.58   | 1.29    | .000 ***    |

**ns** non-significant.  *significant at p < 0.05.  ***significant at p < 0.001.

| Table 5. The observational data of students’ learning behaviors in an interactive physical learning (n = 60) |
|---------------------------------------------------------------|-------|-------|
| Active learning : Performers                                  | Class A | Class B |
| Poses of readiness with both hands stretching forward          | 38%    | 36%    |
| Trailing and exploring the kitchen tools with hands            | 88%    | 82%    |
| Fulfillment : Performers                                      |        |       |
| Smile after tasks                                             | 63%    | 91%    |
| Cheering poses (Victory poses)                                | 38%    | 36%    |
| Engagement : Non-performers                                   |        |       |
| Imitate performer’s action                                    | 37%    | 33%    |
| Imitate performer’s speech                                    | 25%    | 28%    |

*Figure 3. Imitating behavior*
Discussion

We found that the involvement of body motions elicits considerable fun when performing tasks. Although some students felt stressed and shy because they saw themselves on the screen, the stress seemed not to overly affect their performance. Conversely, being in the spotlight might have encouraged students to intensify their efforts in the activities. Moreover, only students whose images appeared in the scenario could trigger the animated effects of the surrounding objects, such as touching the cash register. This “magic power” spurred students to set aside the pressure.

When learners conducted real-time observations of their self-actions while interacting with contextual objects on the vertical screen, positive learning behaviors were generated. These positive learning behaviors may result in strong learning outcomes, but those outcomes were not measured in this study. However, social cognitive theory supports that “learning interventions affect learning outcomes through reflection on observations” (Gupta & Bostrom, 2012, p. 2). Accordingly, in the proposed scheme, observing the performance of classmates might positively affect peer learners who are engaged in vicarious learning. This concurs with Gupta and Bostrom (2012), who indicated that enactive enabled and technology-mediated learning significantly and positively influence learning. However, Table 2 shows no significant difference in learner behaviors between the CG and EG.

The most interesting finding of the research was the focused attention of the peer audience while they were learning within the designed learning environment. Classmates’ images in the situated scenarios were great attention getters, and the live performance kept the students focused on the learning matter. Holding the attention of a whole class on a subject matter is a difficult task, especially under game-based learning settings. We were glad to have found a way to retain the full attention and energy of learners engaged in a learning activity. The animated image on the vertical screen may provide an additional vicarious learning channel for both the performer and the peer audience.

Conclusions

We proposed a self-observation model that employs an instinctive interface, enhancing engagement and active learning by incorporating activities, role-play, and body movements. The design in self-observation model should be suitable for classroom learning because the activity does not only engage the learner who is performing the task but also the peer learners. Peer learners ought to see the performer for some reasons. The most obvious reason is that they also will have their turn to perform so they need to prepare themselves by observing their classmate performing the task.

The proposed learning environment should contribute to the study of technology-supported active learning. Thus far, active learning has been perceived primarily in terms of cognitively active learning, and this may result from the availability of previous forms of technology. However, contemporary technological tools have the ability and potential to support additional functions. These tools allow the application of physically (behaviorally) active learning under a theoretical support from, for example, embodied cognition view.

Reflecting on the experience of designing the learning space, we learned at least two lessons. First, the scenario must be designed to engage learners in the context of the roles that they must play. Relevant virtual objects appearing on the screen must be provided, and learners should be able to manipulate these objects with their body motions. The objects should respond to the movement accordingly. Second, the learning target must be realized in the objects with which learners interact. In this way, learners have the opportunity to reflect and directly experience knowledge based on the outcome of their actions. We expect the knowledge obtained to be relevant to real-world situations.

The animated image of a performer onscreen for self-observation should assist in creating deep impressions and clear experiences that may improve learning. Therefore, we believe that enabling actions through body movement, providing context that mimics a real-world situation, and seeing an individualized animated image should facilitate deep experiences and engagement in learning. We observed that this learning mechanism truly engaged students with authentic tasks. Because of the Kinect and image-processing technologies, we were able to create a near-authentic experience that embedded students’ images in animated contexts and enabled learners to use language and body language to act out and present learned knowledge.
Finally, manipulating events and objects by using body movements and speech commands allows learners to engage with their assigned roles. We can conclude that students are impressed when they have the opportunity to use learned knowledge actively in a relevant context. This learning design with planned settings has great potential to generate a focused and engaging learning atmosphere that involves the whole class. Because our concern was to explore the learners’ behavior in relation to our design, we did not quantitatively show the learning gains regarding language learning achievement in our system, although Macedonia and Knösche (2011) suggested the beneficial effects of body movements, for example, for word memorization.

Future research could include further exploring active and engaging learning environments by including more themed activities. Involving teachers in planning their teaching and learning activities with this system, which they can readily implement with an authoring tool, is also possible. However, current limitation is that the authoring tool for the teachers has not been available yet in our system. Besides, we chose a specific subject matter, i.e., second language learning, and created the setting for this to understand the behaviors of learners when they interacted with the system. Future study may also create the settings for other subject matters, or measure learning outcomes and learning gains from the settings. It is possible that the system can become a general purpose environment to flexibly accommodate numerous subjects for a classroom.

References


To See or Not to See: Effects of Online Access to Peer-Generated Questions on Performance

Fu-Yun Yu* and Yen-Ting Yang
Institute of Education, National Cheng Kung University, No. 1, University Road, Tainan City, Taiwan // No. 1, Shürén Rd, Gangshan District, Kaohsiung City, Taiwan // fuyun.ncku@gmail.com // swallow19820307@hotmail.com

*Corresponding author

(Submitted April 23, 2013; Revised August 3, 2013; Accepted August 29, 2013)

ABSTRACT
This study examined the effects on performance of online access to peer-generated questions during question-generation activities. Two eighth grade classes (N = 63) participated in six weekly question-generation sessions to support English learning. An online student question-generation learning system was adopted. In contrast to expectations based on the literature on observational learning and scaffolding, no significant differences in English academic performance and weekly question-generation performance were found between the groups with and without online access to peer-generated questions. Nonetheless, the question-viewing function of the system used in this work induced a significant and immediate increase in QG performance for the group with access. Possible reasons for the unexpected non-significant results are provided with reference to limited attention theory and cognitive load theory, along with the significance of this study, as well as suggestions for instructional implementations and future research.

Keywords
Online learning space, Observational learning, Peer-assisted learning, Scaffolding, Student question-generation

Introduction
Questioning in all forms is considered to be an essential educational tool in different disciplines (Chin, Brown & Bruce, 2002). However, current classroom practice is dominated by teachers posing questions in order to assess the comprehension and cognitive abilities of their students. This generally gives primacy to instructor perspectives (Silver, 1994), and accentuates the summative rather than formative aspect of evaluation.

From the perspectives of constructivism, information-processing theory, and metacognition, student question-generation (SQG) creates situations that direct students to actively process information they see as important and relevant, and to activate, apply, organize, integrate, and construct/reconstruct personal knowledge while learning (Yu, Liu & Chan, 2005). These deep, constructive processes induce many psychological ones on the part of the learners (including rehearsal, organization, elaboration, reflection, planning, monitoring, evaluation, and revision), which lead to cognitive and metacognitive development (Andre & Anderson, 1978-79; Gillespie, 1990).

SQG has been found to promote student learning and growth in a number of studies (Brown & Walter, 2004; Luxton-Reilly, 2012; Rosenshine, Meister & Chapman, 1996), and there is thus clear empirical support for the teaching and use of SQG to enhance comprehension of learned content (Brown & Walter, 2004; Drake & Barlow, 2007; Gillespie, 1990), and to promote the growth of student cognitive and metacognitive strategies (Andre & Anderson, 1978-79), active learning (Barak & Rafaeli, 2004), diverse and flexible thinking (Andre & Anderson, 1978-79; Brown & Walter, 2004; English, 1997), problem-solving abilities (Dori & Herscovitz, 1999), intra-group communication (Yu & Liu, 2005), self-confidence (Whitin, 2004), and motivation (Chin et al., 2002). However, studies have also found that students have serious concerns over their lack of prior experience in SQG during formal schooling (Moses, Bjork & Goldenberg, 1993; Vreman-de Olde & de Jong, 2006), and thus doubts about their capabilities and performance at such tasks (Yu & Liu, 2005). Additionally, one study reported that the majority of students it examined viewed SQG as difficult or very difficult (Yu, 2009). To address these concerns, one of the current authors has been working on creating various theoretically sound designs to support SQG while building an empirical basis for their associated effects under a three-year project, starting in 2009.

Due to the wide support that peer-assisted learning has among academics and practitioners (Topping & Ehly, 1998), as well as the popularization of web 2.0 related notions such as “user as contributor,” “the wisdom of crowds,” “the
right to remix,” and “innovation in assembly” (Abramovich & Brouwer, 2008; O’Reilly, 2005), various ways to support learning and improve outcomes by the use of peer-generated work have been proposed. In view of the fact that online access to peers’ work allows observation and imitation, which can provide scaffolding for learning, and that technology can play a central role in facilitating this process (i.e., the collection and instant display of student-contributed items), this study examined the effects of online access to peer-generated questions. Two theoretical perspectives which suggest that online access to peers’ work may be conducive to learning are explained briefly below, before the methods used in this study are presented.

**Observational learning theory**

Bandura’s observational learning theory, which highlights the beneficial effects on cognitive development and one’s competence at a task that are derived by watching other people, is one of the most influential theories of learning and development (Lefrancois, 1999). According to Bandura, “most human behavior is learned observationally through modeling: from observing others one forms an idea of how new behaviors are performed, and on later occasions this coded information serves as a guide for action” (Bandura, 1977, p. 22). Furthermore, the process involved in observational and modeling is “one of the fundamental means by which new modes of behavior are acquired and existing patterns are modified.” (Bandura, 1986, p. 118).

Allowing students to view questions their peers have generated provides opportunities for observation and modeling, and these are expected to enhance the observers’ existing skills with regard to SQG, as well as their knowledge of the focal domain. However, as the issue of whether students actually benefit from viewing peer-generated questions while engaged in SQG has not yet been empirically examined in the literature, its effects on both academic and SQG performance are investigated in this study.

**Scaffolding**

Scaffolding, a term coined by Wood, Bruner, and Ross (1976), is a form of support that can help learners to gradually move beyond their current level of abilities (actual development) to the intended one (potential development). A number of techniques have been suggested as effective scaffolding arrangements, such as providing models or partial solutions to problems, marking critical features of artifacts or concepts, and simplifying tasks (Wu, 2010). Additionally, controlling for the elements of the task that “are initially beyond the learner’s capacity, thus permitting him to concentrate upon and complete only those elements that are within his range of competence” (Wood et al., 1976, p. 90) by providing scaffolding congruent with the learner’s current knowledge level is seen as essential for the success of this approach (Puntambekar & Hubscher, 2005).

Supporters of scaffolding emphasize its ability to help students avoid floundering and frustration (Rummel & Kramer, 2010), reduce the extraneous cognitive load induced by random trial-and-error (Sweller, 2006), decrease student working memory load while dealing with novel information (Kirschner, Sweller & Clark, 2006), and clarify procedural or conceptual misunderstandings and incomplete knowledge (Sharma & Hannafin, 2007). Generally speaking, scaffolding positively impacts student learning and has been suggested as one of the most effective instructional interventions for producing positive learning outcomes (Swanson, 1999). Recently, the positive effects of scaffolding on student cognitive outcomes were supported by a meta-analysis carried out by Belland, Walker, Olsen and Leary (2012) that examined science, technology, engineering, and mathematics education at the K-12, college, graduate, and adult levels.

Since peers are generally within each other’s zone of proximal development, peer utterances (in this case, peer-generated questions) should be more comprehensible to learners, as compared to those constructed by teachers or provided in textbooks (Ammer, 1998; Fallows & Chandramohan, 2001). Giving students access to peer-generated questions, while the task is still new to them, also allows them to see various possible forms, types, and qualities of questions. This arrangement may not only provide students with some ideas and directions for their own question-generation (QG) endeavors (thus avoiding floundering and frustration), but may also help to highlight the critical features of good questions that should be learned and imitated.
Purpose of the study and research questions

In sum, based on the concepts of observational learning and scaffolding, it is proposed that the extra space and learning opportunities provided by accessing peer-generated questions during QG may increase the learning performance associated with SQG, although the exact effects of this approach remain to be empirically substantiated. As such, this study examined the following two research questions:

- Do students with online access to peer-generated questions during SQG perform better in academic achievement as compared to those without access?
- Do students with online access to peer-generated questions during SQG have better SQG performance as compared to those without access?

While answering these questions can help in the design of better SQG activities and environments, it can also add to the broader research on constructive activities that stress the idea of ‘learning by doing,’ initially promoted by John Dewey (1963) (e.g., project-based learning, student-designed games, and so on). In response to contemporary approaches to teaching and learning that highlight the importance of students being engaged in meaningful and constructive learning activities, the supportive effects of online access to peer-generated work on the quality of the work produced and comprehension of the focal knowledge are important issues, and thus are considered in this study.

Methods

Experimental design and independent variables

This study adopted a quasi-experimental research method. Two groups were devised—the experimental group (with access to peer-generated questions) and the contrast group (without access) during SQG. Instructions and arrangements during the study were essentially the same for both groups (including weekly schedule, instructional content covered, supplemental materials, and in- and out-of-class assignments and activities), except that the access group had access to both the question-viewing and QG functions during QG activity, while the without access group only had access to the latter.

Students in the access group were able to observe questions produced by their peers the moment they were submitted, starting from the second QG activity. The question-viewing function was intentionally deactivated for the access group during their first activity, so that a baseline for the SQG abilities of both groups could be established. Finally, to account for both of the major processes of scaffolding—support building and support withdrawal (Punambeke & Hubscher 2005) – the question-viewing function was deactivated for the access group during their last activity (i.e., the sixth session).

Experimental procedures

Prior to the actual study, to ensure experimental fidelity, an instructor manual delineating the implementation procedures for the respective groups was developed and closely followed. Supplemental materials containing essential information about SQG (e.g., performance assessment, operational procedures for the adopted online system) were also developed. A pilot study was conducted to ensure that the planned procedures and time allocated for each component of the SQG activities was appropriate, and that the instruction used for the experimental implementation was clear for the targeted audience (i.e., eighth graders).

Prior to the implementation of this study, a consent form containing the following information was prepared: the purpose of the study, the information to be collected from the participants and how it would be used, an assurance that confidentiality and anonymity would be maintained, and the researcher’s contact information to enable correspondence, if required. To eliminate the possibility of experimenter bias brought about by differences in teaching experience, personality, communication capabilities, and so on, which would reduce the internal validity of this study (Brewer, 2000), two eighth grade classes (N = 64) taught by the same English teacher participated in the actual study, and were randomly assigned to the two treatment conditions after informed consent had been obtained.
As the principles of SQG are in alignment with current notions and strategies of computer-assisted language learning (e.g., an emphasis on learners using the target language and being cognitively engaged to create meanings in various contexts, as stressed by task-based instruction and communicative language teaching) (Brandl, 2007; Wang, 2009), it was introduced to support student learning of English as a second language, using the students’ regular 45-minute weekly English study session. The SQG activities were held in the school’s computer lab two weeks after the first school-wide English examination, and ended one week prior to the second exam. Two question types often used in the school-wide English examinations were chosen for the SQG activities, namely fill-in-the-blank and multiple-choice ones.

One week before the beginning of the actual study, a training session was arranged. Students were introduced to basic SQG practices, rules of thumb for generating the two chosen question types, and the operational procedures for the adopted system. Moreover, in reference to Rosenshine, Meister and Chapman (1996), effective instructional elements supporting SQG were delineated and explained to both treatment groups. These included the criteria that the teacher used to assess the performance of SQG, and models of appropriate QG (see the section on measurement for details).

Following the training session, the students were directed to generate at least four questions about recently learned English content on a weekly basis using the adopted system for a period of six consecutive weeks. Of those generated questions, at least two had to be fill-in-the-blank items and two had to be multiple-choice ones. Students were directed to generate questions about the main ideas of the study material covered each week, which were identified by the teacher prior to each SQG activity. During the study, topics covered included the past progressive tense (be + V-ing); pronunciation rules for ir, er, ur, ear, er, ar and ur; the use of when, ‘not…at all,’ infinitives, gerund (V-ing), ‘it’ as a virtual subject, ‘help’ followed by either a verb or noun, and prepositions followed by a gerund (for/at/about/by) + V-ing; and a total of 90 vocabulary items. Students were encouraged to refer to textbooks, English-Chinese dictionaries, take-home worksheets, or information found on the internet during the activities.

Starting from the second week, before proceeding to the QG tasks, whole-class feedback was given, highlighting both good work and the most common mistakes found in the student-generated questions submitted the previous week from each group. For the access group, students were also briefed about the use of the question-viewing function of the adopted system, and were advised to use their time wisely when generating and viewing questions during the allocated time frame. One week after the completion of the sixth SQG activity, teacher-designed and school-wide English exams on the study content were administered.

**Online learning system**

An online learning system with a focus on SQG (the Question Authoring and Reasoning Knowledge System, QuARKS) was adopted to support the activity. Like all similar online systems, QuARKS enables multimedia files to be included as parts of the question; texts of different fonts, sizes, and styles can be used; and questions can be easily saved, retrieved, revised, and deleted by users (see Figure 2). However, to the best of the researchers’ knowledge, QuARKS is different from other systems in at least one way— it is customizable in terms of the specific function(s) available to students at any given point in time (Yu, 2009). For the purpose of this study, both QG and question-viewing functions were activated for the experimental group, while only the QG function was activated for the contrast one (Figure 1).

For QG, students first clicked on the QG button and selected the type of question to be generated (see Figure 1), and were then directed to the appropriate space. Except for the multiple-choice questions, which demand an additional task of constructing four alternatives, students would generate a question item, provide a correct answer, and detail the main purpose or reference for each question asked (Figure 2).

With QuARKS, the teacher was able to schedule the activation or deactivation of the question-viewing function available to the access group ahead of time, or in real time. Once activated, students would click on the question-viewing function (Figure 1) to be directed to a list of peer-generated questions (see Figure 3). The list of questions could be re-sequenced by clicking on the question type, questions, version, question-poser (the exact content displayed is dependent on the identity mode chosen by this individual), submission time, and number of times
viewed labels located in the first row. Students simply had to click on the ‘browse’ button (i.e., the eye icon, 🕵️‍♂️) of any specific question item to examine its full content.

Figure 1. Learning spaces for the experimental group with both QG and question-viewing functions (left) and the contrast group with only QG function (right)

Measurements

This study explored the relative effects of having versus not having online access to peer-generated questions during QG activities on the English academic and QG performances of two groups of eighth-grade students. Each of the instruments used to achieve this is described below.
A posttest developed by the teacher was used to assess student academic performance on the covered content. It was composed of 35 fill-in-the-blank and four-option multiple-choice question items. To ensure that items were of adequate quality, only items with a test difficulty between .35 and .85 and a discrimination index above .20 were included for data analysis. As a result, one item was excluded. The Kuder and Richardson reliability coefficient was .95, reflecting the high internal consistency of the test (i.e., students with a higher level of knowledge of the tested domain tended to answer most items correctly, while students with a lower level tended to answer most items incorrectly). The average item difficulty level was .54 (ranging from .37 to .74), indicating that the items were at an appropriate level relative to the abilities of the participants. The item discrimination index ranged from .39 to 1, which shows that the test discriminated between students with higher and lower levels of knowledge of the content being assessed. In addition, a school-wide exam was administered to all the eighth graders. In total, 44 question items on the covered content were included for item analysis and, as a result, two items were deleted. The Kuder and Richardson reliability coefficient was .96 for the 42 remaining items. The average difficulty level was .64 (ranging from .40 to .80), with a discrimination level ranging from .40 to .88.

To assess student QG performance, a set of criteria was developed with reference to the Torrance creativity index (1974) and Yu and Wu's (2013) criteria for QG performance assessment. All questions that students generated during the six QG activities were analyzed, scored, and summed up according to these criteria. Specifically, each question was graded on five dimensions: fluency, complexity, elaboration, originality, and importance. For inter-rater consistency, one English teacher (with more than six years of English teaching experience) was recruited and trained to use the criteria to assess one session of SQG performance. The Kendall coefficient of concordance between the two raters was .88. Definitions and model examples of each of the five dimensions of QG performance were explained to the students during the training session, and were available for reference during the entire QG activities (see Table 1).

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Definition</th>
<th>Model example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>Free from spelling and grammatical errors; Precise meaning with one correct answer; Complete questions and annotations for fill-in-the-blank questions, and question-stems, four options, answers, and annotations for multiple-choice questions.</td>
<td>Fill-in-the-blank example: Question: My father always takes a bath in the bathroom after work.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Answer: bathroom.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Annotation: The place for taking a bath, beginning with the letter b and ending with the letter m, is bathroom.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple-choice example:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Question: I took a trip to Taipei last month.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(A) takes (B) did took (C) took (D) taking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Answer: (C).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Annotation: The verb to use with the phrase ‘a trip to a place’ is ‘take.’ As it happened in the past (last month), past tense of ‘take’ should be used (i.e., took).</td>
</tr>
<tr>
<td>Complexity</td>
<td>Involves at least two grammatical rules</td>
<td>Fill-in-the-blank example: Question: Let’s mop the floor on Sunday.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Answer: mop</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Annotation: The verb, which begins with the letter m, ends with the letter p, and is associated with ‘the floor’, is ‘mop.’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[The question tests two grammatical concepts: the use of Let’s + V and the phrase, ‘mop the floor.’]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple-choice example:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Question: Lucy made the bed last night.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(A) makes (B) made (C) take (D) did</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Answer: (B) made</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Annotation: The verb to use with ‘the bed’ is ‘make.’ As the behavior happened last night, the past tense of ‘make’ should be used, which is ‘made.’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[The question involved two grammatical concepts: phrase ‘make the bed’ and the use of past tense]</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Interconnectedness between the currently fill-in-the-blank example</td>
<td>Fill-in-the-blank example: Question: A-Mei is my favorite a ___l singer.</td>
</tr>
</tbody>
</table>

Table 1. Criteria for QG performance assessment: five dimensions with associated definitions and model examples
covered topic/unit and prior topics/units;  
Link to personal life experience;  
Plausibility of alternatives;  
Answer: aboriginal  
Annotation: The word beginning with the letter a and ending with the letter l is ‘aboriginal.’  
[A-Mei is one of the most famous singers in Taiwan. Her ethnic origin (i.e., aboriginal) is a fact well-known among students. This question contains knowledge of pop culture that is a part of students’ everyday lives]  
Multiple-choice example:  
Question: Judy ___ very unhappy after she ___ her wallet.  
(A) was; lose (B) were; lost (C) is; lost (D) was; lost  
Answer: (D) was; lost  
Annotation: the verbs used should be consistent with its subject and time tense.  
[Past tense of the verb ‘lose’ is a newly learned verb; the past tense of the be+V was previously learned; all distracters are frequently made mistakes by students]

| Originality | Uniqueness of a specific question as compared to those of peers. | Fill-in-the-blank example: | Question: John only ate a piece of p a for dinner.  
Answer: pizza  
Annotation: the word beginning with the letter p and ending with the letter a is ‘pizza.’  
[‘A piece of’ is newly learned phrase; however, ‘a piece of pizza’ was not specifically taught or mentioned in class, so this student was the only one to compose this specific question]  
Multiple-choice example:  
Question: Do you know what happened Mary ?  
(A) to (B) with (C) at (D) in  
Answer: (A) to  
Annotation: ‘to’ should be used with ‘happen’ when we want to denote something happen to someone  
[No students in the class composed questions using an indirect question] |

Answer: happened  
Annotation: ‘happen to’ is the phrase to use when denoting something happening to someone.  
[The word ‘happen’ and the phrase ‘happen to’ are important concepts to be learned]  
Multiple-choice example:  
Question: Julia is singing and dancing.  
(A) good at (B) well on (C) like (D) plays  
Answer: (A) good at  
Annotation: phrase to use is ‘be good at’  
[Be good at + sth/Ving are important grammatical concepts covered] |

**Data analysis**

The multivariate analysis of covariance technique (MANCOVA) was used to test whether there were any differences in English academic performance (measured using the teacher-developed posttest and second school-wide exam) between the treatment groups, as a result of a statistically significant moderate correlation coefficient among the two variables (Hair, Black, Brain, Anderson, & Tatham, 2006). Scores from the first school-wide exam were used as the covariate. The QG performances of students from the two groups were compared on a weekly basis using t-tests.
Results

English academic performance

Table 2 summarizes the descriptive statistics for both groups’ English performance. The assumption of the homogeneity of the regression was satisfied, with Wilks $\lambda = 0.97$, $F(1,58) = 1.04$, and $p > .05$, before proceeding to MANCOVA. The results of MANCOVA revealed no significant differences between the two treatment groups, with Wilks $\lambda = 0.95$, $F(1,60) = 1.65$, and $p > .05$.

<table>
<thead>
<tr>
<th>Treatment groups (n)</th>
<th>1st school-wide exam$^1$</th>
<th>Teacher-designed test$^2$</th>
<th>2nd school-wide exam$^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access group (30)</td>
<td>64.30 (30.19)</td>
<td>17.40 (9.79)</td>
<td>17.37</td>
</tr>
<tr>
<td>Without access group (33)</td>
<td>69.24 (23.90)</td>
<td>21.18 (9.64)</td>
<td>21.20</td>
</tr>
<tr>
<td>Total (63)</td>
<td>66.89 (26.97)</td>
<td>19.38 (9.82)</td>
<td>30.38 (10.71)</td>
</tr>
</tbody>
</table>

$^1$maximum Score: 100; $^2$maximum score: 34; $^3$maximum score: 42.

Student QG performance

The student’ week-by-week QG performance for the six sessions is given in Table 3 and graphed in Figure 4. As shown in Figure 4, the two groups started off similarly in terms of QG performance. However, once the question-viewing function was activated for the access group in the second session, they then appeared to improve significantly. On the other hand, the slope for the without access group rose steadily from the first to the fifth week with a sudden drop in the last week.

The results of the independent-sample $t$-tests carried out between the two treatment groups did not reveal any significant differences in any of the six weeks. Nevertheless, the paired-sample $t$-tests indicated a significant difference between weeks 1 and 2, $t(29)=3.097$, $p < .05$, for the access group, but not for the group without access, $t(29) = .141$, $p > .05$.

<table>
<thead>
<tr>
<th>QG</th>
<th>Access group (n = 30)</th>
<th>Without access group (n = 33)</th>
<th>$t$-test</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>25.60 (6.49)</td>
<td>25.19 (6.32)</td>
<td>.248</td>
<td>.805</td>
</tr>
<tr>
<td>2nd</td>
<td>30.37 (7.50)</td>
<td>26.84 (11.03)</td>
<td>1.461</td>
<td>.149</td>
</tr>
<tr>
<td>3rd</td>
<td>28.70 (7.66)</td>
<td>28.81 (13.07)</td>
<td>-.043</td>
<td>.966</td>
</tr>
<tr>
<td>4th</td>
<td>30.93 (9.28)</td>
<td>30.06 (8.56)</td>
<td>.384</td>
<td>.702</td>
</tr>
<tr>
<td>5th</td>
<td>29.79 (8.16)</td>
<td>30.12 (10.33)</td>
<td>-.137</td>
<td>.891</td>
</tr>
<tr>
<td>6th</td>
<td>29.97 (7.88)</td>
<td>25.91 (9.05)</td>
<td>1.878</td>
<td>.065</td>
</tr>
</tbody>
</table>

Figure 4. Week-by-week SQG performance of both groups
Discussion and conclusions

Despite the fact that studies on SQG have generally supported its positive effects on learning, a number of problems, including student concerns over the quality of the work produced, lack of experience, and perceived difficulties with regard to QG, all call for better support mechanisms or designs when this approach is being utilized. The effects that access to an online space for peer-generated questions had on student learning were thus examined in this work, based on the theories of observational learning and scaffolding.

Contrary to the authors’ expectations, the results failed to reveal any significant positive effects on student performance related to having access to peer-generated questions. There may be several plausible reasons for this. First, using a controlled experimental design, the time allocated for interacting with the online system in class was fixed and limited, and kept the same for both groups in this study. However, since people have a limited capacity for attention, and only a certain amount of information can be attended to at any given time (Ruz & Lupiáñez, 2002), this implies that students in the access group operated under greater time and resource pressure, as any time and attention used to observe peers’ work could not be used to generate questions. As such, any enhanced competencies gained from the observations might not necessarily have been applied to generating questions. A failure to find the positive effects on learning has also been found in other studies using learning time controls (for instance, Große and Renkl’s (2006) study on the effects of multiple solution methods by means of worked-out examples).

Another possible reason for the non-significant results is related to the questions that could be seen by the access group. The theories of observational learning and scaffolding both call for ‘exemplars’ or ‘models’ for observation and imitation. Although studies have found that students can be capable of generating questions that reflect cognitive skills similar to those of their instructors and that cover all the major topics of interest (Luxton-Reilly 2012), the truth of the matter is that not all peer-generated questions can qualify as ‘models,’ and not all students possess the abilities and time management skills need to find good model questions out of all the available ones, never mind being able to learn from them. To take advantage of the capabilities of networked technologies, the questions submitted by the students in the access group were immediately made available for their peers to view. If creating ‘models’ is a priority, it is clear that some delays or logistical problems may surface, even with extra technological or personnel support, and offering these would likely change the observations and assessments of the related effects.

The third and final possible reason for the non-significant findings of the current work might be understood with reference to the cognitive load theory (CLT). CLT argues that learning and carrying out related activities require learners to execute cognitive efforts within the limited capacity of their working memory to process the broad range of information associated with the task at hand. Instructional strategies that exclude unnecessary information and help learners to concentrate on relevant contents could help them to effectively and efficiently manage their limited cognitive capacity (Kirschner et al., 2006; Paas, Renkl, & Sweller, 2004; Sweller, 1988, Sweller, van Merriënboer, & Paas, 1998). While peer-generated questions could serve as worked-out examples, the best known and most widely studied of the cognitive load effects, and thus decrease the external cognitive load (Pass, & van Gog, 2006; Sweller, 2006) on the part of the question-author, students are not necessarily capable of using the resultant available working memory capacity (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Gerjets, Scheiter, & Catrambone, 2004). As has already been noted, simply presenting worked-out examples does not necessarily lead to schema construction (Renkl, 1997), possibly because learners cannot differentiate relevant information from the irrelevant, and thus may fixate on less relevant features of the examples (Ross, 1989), or the learners may lack the knowledge and skills needed to learn from the provided examples (Catrambone & Holyoak, 1989). Furthermore, inappropriate ways of designing worked-out examples might increase the learners’ cognitive load (Gerjets, Scheiter, & Catrambone, 2006; Pollock, Chandler, & Sweller, 2002). In particular, what cognitive load theorists describe as the redundancy effect (Sweller, 2006) may be present in this study. Specifically, the use of a full text with detailed information does not always lead to enhanced learning, as compared to the use of a summarized text, due to the fact that the extra material in the full text may be redundant (Jin, 2012). As noted above, whole-class feedback highlighting both good questions and the most commonly made mistakes was provided starting from the second QG session. The instructor’s weekly feedback acted as the summarized text, while accessing the peer-generated questions resembled, to some extent, the full text. As a consequence, any beneficial effects from the online observation space that may arise for the access group might have been compensated for in the without access group by the teacher’s feedback, and thus the access to the online information may have had a less significant effect than otherwise. Moreover, the unabridged, uncensored materials accessible to the access group (i.e., the peer-generated questions) may unintentionally have consumed too much of
the limited cognitive capacity of the participants, and thus reduced the benefits that may otherwise have accrued from accessing peer-generated work.

Even though the two groups did not differ in terms of QG performance in any of the weeks examined in this work, it is worth noting that this performance was significantly enhanced for the access group in the first week that the question-viewing function was activated (i.e., the second week). The significantly better QG performance reflected, at some levels, what Yu, Tsai and Wu (2013) reported in their study on SQG—the immediate, positive effects of procedural scaffolding (in the form of generic question-stems with context-specific examples), although these then faded shortly afterwards.

**Significance of the study and suggestions for instructors and future research**

To the best of the authors’ knowledge, this is the first study to empirically examine the effects of instant online access to peer-generated questions as a support mechanism for SQG. Despite being theoretically sound, this study did not confirm the superiority of having online access to peer-generated questions as compared to having no access. However, it should be noted that in the first week that the question-viewing function was activated for the access group, these learners saw a significant and immediate improvement in QG performance. In light of this, instructors dealing with students with no prior experience of QG, or other innovative constructive activities (e.g., project-based learning, student-designed gaming, and so on), might consider allowing them to have online access to peer-generated work as a form of immediate support.

Several areas for further investigation are suggested, as follows. First, in this study all QG activities were carried out in class. However, if it was designed otherwise, and thus students did not operate under a time constrained situation, and thus had more time to freely observe and construct questions, the resulting effects might have been different. That being said, the proposed effects still need to be validated empirically by a future study.

Second, almost all existing online systems that focus on QG have a peer-assessment component in place (Luxton-Reilly, 2012). With such a design, students can have access only to questions that have already passed through the peer-assessment phase, rather than to all those that have been submitted. The effects of providing online access to student-generated ‘model questions’ is one topic worth examining to extend the findings of the current study.

Finally, even though this study found that the question-viewing function did not lead to better academic or QG performance, as compared to the group without access to these questions, researchers and practitioners are advised to explore the effects of other types of scaffolding, such as reflective or metacognitive scaffolding (Ge & Land, 2004; Jackson, Krajcik & Soloway, 1998), with regard to the effects on SQG and learning.

**Acknowledgements**

This paper was funded by a research grant from the National Science Council, Taiwan (Project title: Scaffolding student-generated questions online learning activities: Research-guided approaches and comparative learning effects; project number: NSC 99-2511-S-006-015-MY3).

**References**


Experimenting with Automatic Text-to-Diagram Conversion: A Novel Teaching Aid for the Blind People

Anirban Mukherjee, Utpal Garain * and Arindam Biswas

RCC Institute of Information Technology, Indian Statistical Institute, Indian Institute of Engineering, Science and Technology, India // anirbanm.rcciit@gmail.com // utpal@isical.ac.in // barindam@gmail.com

*Corresponding author

(Submitted April 30, 2013; Revised July 25, 2013, 2011; Accepted October 10, 2013)

ABSTRACT

Diagram describing texts are integral part of science and engineering subjects including geometry, physics, engineering drawing, etc. In order to understand such text, one, at first, tries to draw or perceive the underlying diagram. For perception of the blind students such diagrams need to be drawn in some non-visual accessible form like tactile graphics. Technologies for producing tactile graphics are available but they are too expensive to be afforded by the blind students or schools in developing countries like India. As a result, science education for a large population of blind students is severely compromised. This paper proposes a novel solution to this problem. A method for digital to Braille mapping of geometry diagrams on the low-cost traditional Braille text printer is reported here. Later on, this is integrated with a previously developed text-to-diagram conversion system. Using the integrated system, a blind student can input a geometry word problem and perceive the underlying diagram on a Braille printout. The major part of the study involves rigorous evaluation of the system at a Blind school. The enthusiasm and the ability shown by the subjects in using the system strongly attest its viability as an effective teaching/learning tool for the blind students.

Keywords

Artificial intelligence, Geometry diagrams, Natural language problem, Tactile graphics, Blind people

Introduction

In many branches of study, we often encounter texts or problems that are visually represented by figures or diagrams. For example, in geometry, physics (mechanics) and in several engineering branches like mechanical, electrical, and electronics such texts appear very frequently. While solving a problem, usually, a problem statement (text) is first translated into a sketch (diagram) which visually articulates the essential problem parts; mechanical models, free-body diagrams, electrical/electronic circuits, geometry diagrams are instances of such transformation. From the point of view of understanding a problem, the representative diagrams are not only a mere convenience but also an inherent component in a person’s cognitive representation of the text or problem. A blind person might rely on and need such diagrammatic representations just as much as a sighted user does and appropriate tactile representation may play that role.

For the sighted users, standard textbooks of science and engineering are widely available that contain numerous examples illustrated with diagrams. On the contrary, similar textbooks with embossed illustrations are not easily available to the blind students. This may be attributed to the fact that producing tactile versions of large number of figures for textbooks is a time consuming, labor intensive and costly process. In a classroom of sighted students, teachers can explain any topic or problem with free-hand sketch of suitable diagrams on blackboard. They can also use computer-based teaching (CBT) tools for better presentation of text and graphics. For the blind students, there are Braille image embossers and compatible graphics programs like IVEO Viewer (Viewplus, 2013), PictureBraille (PictureBraille, 2013), TGD Pro (Duxbury Systems Products, 2013), Tiger Software (ViewPlus Tiger Software Suite, 2013), TACTICS (Way & Barner, 1997), etc. using which a teacher can generate tactile diagrams from printed or digital diagrams. Also there exist sophisticated audio-tactile, audio-haptic, and multimodal interfaces, developed to produce diagrams in a form accessible to the blind students. Some of the examples are IC2D (Hesham & James, 1999), TDraw (Kurze, 1996), NOMAD (Parkes, 1991), Talking Tactile Tablet (Landau & Gourgey, 2003), Touch Tiles (Bussel, 2003), IVEO touchpad (Krufta & Barner, 2005), DESENOX (Borges & Jansen, 1999), AudioTact/Barbieri et al., 2008) Math Class (Albert, 2006), and SALOME (Gouy-Paillet et al., 2007). Some interesting applications have been developed by Minagawa & Ohnishi (1996), Jayant et al. (2007), Watanabe et al. (2006), Guha & Anand (1992), Lahav & Mioduser (2008), and Toennies et al. (2011). But all such systems including Braille embossers are very expensive (cost of embossers ranging from US$ 5000 to 26,000) and hence not available...
in most of the blind schools in India. At best the schools can have access to traditional Braille text printers that cannot print tactile image.

Thus a large population of blind students in developing countries (Casely-Hayford & Lynch, 2003) like India grows up without any exposure to modern learning aids for diagram-based subjects like geometry. Historically, teaching geometry to these students is limited to giving basic theoretical definitions only and exercises requiring frequent diagram drawing are deliberately avoided. Though nail board or wooden pieces are sometimes used to perceive simple diagrams in lower grades, the lack of frequent access to tactile diagrams of wide varieties results in students memorizing facts as verbal assertions and this seriously limits the development of their scientific skill. In many parts of India, even today, blind students are forced to leave studying science subjects after 7th or 8th grade because of inconvenience of learning diagrams (Rahman et al., 2010). Herein lies the need for affordable and easy-to-use technology that could facilitate diagram drawing in tactile form upon reading a geometry text or problem. This paper is motivated by this need.

The paper reports a novel method by which digital diagrams can be converted to tactile diagrams using low-cost traditional Braille text printer. Keeping the primary focus on representing basic digital shapes (line and circle) in Braille, the paper further establishes how such shapes (and eventually, any geometric diagram) can be generated by the blind students themselves through text mode of input. In the process it integrates the Braille mapping module with our earlier developed system of automatic conversion of geometric text to digital diagrams (Mukherjee & Garain, 2009), the system comprising a geometry knowledgebase (Mukherjee et al., 2007), a NLP module (Mukherjee et al., 2013), and a graphics module.

Figure 1 shows a schematic diagram of the whole system that integrates the text-to-diagram conversion module with the Braille mapping module. This paper focuses on the part shown within the dotted lines. The main focus of the paper is, however, the user evaluation and impact analysis of the entire system of automatically generating geometry diagram on Braille.

**Figure 1.** The connection between the system of automatic diagram drawing from text and the proposed Braille mapping system (within dotted boundary).

### Drawing in Braille

Braille is a system commonly used by the blind people to read and write. In computer graphics, the smallest addressable picture elements are pixels. In Braille, the smallest physical unit is an embossed dot, while the smallest logical unit is a Braille cell—a 3 x 2 array of 6 dot-positions. While each pixel can display a color according to the bitmap value, each dot of a Braille cell may be in embossed state depending on the NUMBRL code (Krebs, 1977) of the character to be represented in that cell. A given NUMBRL code corresponds to a character of English alphabet or a punctuation symbol or a digit.

NUMBRL is basically a numeric code that represents the dot patterns in Braille cells. Each dot in a cell has a fixed position value (Figure 2b). The NUMBRL code of a cell is just the sum of the position values of the embossed dots of the cell. For example, consider a cell with embossed dot positions 4-2-6 (Figure 2a). The NUMBRL code for this cell is $1+20+4 = 25$. 

41
As there are 6 dots in a Braille cell the number of different patterns or characters (NUMBRL codes) that can be generated in a cell by embossing a subset of 6 dots at a time is $2^6$ or 64.

In digital graphics (Rogers, 1985), we can algorithmically select and illuminate a set of pixels to generate a line or circle. Similarly, in Braille, the first problem is to identify which dots in which cells are to be embossed for the best possible representation of a line or circle. Next cell-wise patterns of identified dots or the NUMBRL codes are passed on to a Braille text printer for producing the Braille-text version of a line or circle.

**Digital to Braille mapping**

To demonstrate the actual implementation of the Braille mapping system we have made a simulation of Braille character mapping of geometric entities. An array of 6-pixel groups displayed on the computer screen emulates the grid of 6-dot Braille cells printed by a Braille text printer on Braille sheet. One Braille dot having diameter of 1.5 mm is represented by a pixel. The distance between two adjacent dots in a cell is kept uniform (2.4 mm) while the horizontal and vertical distances between two corresponding dots in adjacent cells are set unequal (6.8 mm and 10.1 mm respectively) as usually found in the traditional Braille system. In our experiment, we have used an Automatic Braille Embosser (BPRT) developed by Webel Mediatronics Limited. This is basically a Perkins Brailler and costs about US $3000.

**Braille mapping of basic entities**

Standard graphics algorithms for drawing digital lines or circles are modified using a Braille mapping function to search Braille dots that lie closest to the line or circle path. The function measures and compares the distance of a point (on line or circle) from each of the Braille dots that fall within a small square region around the point. The dot that gives the least distance is the nearest dot found in a search and its position is saved. Pixels at all such dot positions are marked with thick dots to emulate actual embossing of dots by a Braille text printer. Position values of the selected dots in a cell are then added up to find the NUMBRL code of that cell. If no dot is selected in a cell, the NUMBRL value of that cell is 0 which implies that the cell doesn’t participate in Braille mapping of a line or circle. The NUMBRL codes of all the cells in the screen array are saved in a text file. As the printer prints the coded characters of the output file row-wise, the line or circle takes shape in Braille. This process is illustrated in Figure 3.

---

**Figure 2.**
(a) Dot positions
(b) Position values

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>o</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>o</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure 3.**
(a) Corresponding to each digital point (pm1, pm2, etc.) on a line, a Braille dot (dark dot) is selected.
(b) The NUMBRL position values shown for the selected dots in a Braille cell; The NUMBRL codes of the cells are (left to right, top to bottom): 0, 4, (20+40+1 =) 61, 40, 1, 0. The output list given by the braille function contains the following triplets: <1 1 0><1 2 4><1 3 61><2 1 40><2 2 1><2 3 0>
Braille mapping of problem diagrams

The earlier developed text to diagram conversion system (Mukherjee & Garain, 2009) generates numerical values of the parameters of all lines and circles (to be drawn) upon reading the corresponding geometric statements. For example, for a statement like ‘ABCD is a parallelogram’ the numerical values of the endpoint coordinates of the four component lines of the parallelogram ABCD are generated as:

AB: line (100,200),(50,100)
BC: line (50,100),(250,100)
CD: line (250,100),(300,200)
DA: line (300,200),(100,200)

Each of the above output is passed on as input to the Braille mapping module. The Braille line mapping function then creates four Braille lines to produce the Braille version of the parallelogram ABCD. A portion of the emulated Braille character grid and embossed dots approximating different geometric shapes are shown in Figure 4. After generating simple objects, the system is tested for diagrams comprising multiple objects. Given a geometric statement or word problem as input, the representative diagram is automatically generated in Braille. Out of a test set of 40 geometry problems, 32 were drawn correctly by the text-to-digital diagram conversion system. Again these 32 problems produced geometrically correct representations in Braille thereby yielding nearly 100% accuracy rate of text-to-Braille diagram conversion. In the present study we have only used these 32 problems corresponding to which the Braille diagrams are meaningful and useful for learning.

Still to improve the perception of a Braille figure by a blind user and also to improve drawing accuracy, some design optimization is done. Firstly, before Braille mapping, auto-generated digital diagrams are scaled up to fit to the full-screen which in-turn causes printing of enlarged diagrams on Braille paper. Secondly, selection of redundant Braille dots in representing a shape is eliminated using some heuristic functions – at the same time it is ensured that wide gap is not created between two adjacent dots. This is evident from the optimized output (from word problems) shown in Figure 5.

Figure 4. Simulation of Braille dot grid and Braille-converted geometric entities; the digital entities are shown in continuous lines while selected Braille dots are shown in thick dots. (a) line, (b) circle, (c) parallelogram, (d) triangle. Point labels are not auto-generated but put here for sake of clarity.
Figure 5. (a) Simulated diagram for the problem: ABCD is a parallelogram. X and Y are the midpoints of AD and BC. BX and DY cuts AC at M and N respectively. Prove that AM = MN = NC, (b) Simulated diagram for the problem: In triangle ABC, AB = AC. D is the midpoint of BC. From D, perpendiculurs DM, DN are drawn to AB and AC respectively. Prove that DM = DN.

Comparison with other systems

Three software tools namely BrlGraphEditor (Batusic & Urban, 2002), SparshaChitra (Lahiri et al., 2005), and the one developed by Rahman et al. (2010) can be considered close (but not similar) to our system as they can convert existing printed images into tactile form using Braille text printer. As described by Rahman et al., each block of 3x2 pixels in a digital image is considered equivalent to a Braille cell (3x2 dots). The pattern of the marked pixels in each block is mapped to yield same pattern of raised dots in the corresponding cell of the Braille grid. All the mapped cells, when printed, reproduce the digital image in tactile format. So the basic approach taken by these systems is linear mapping from an evenly spaced pixel-grid to an unevenly spaced dot-grid which naturally causes some distortion in the geometric shapes. Moreover, the process does not attempt to optimize the number or position of dots to be embossed. It causes redundant dots being selected at positions other than the best possible positions with respect to the actual entity (line or circle) path. Our system, on the contrary, logically finds the dot nearest to the entity path at any sampling position and further optimizes the number of adjacent dots getting selected to make the best possible outline-representation of an entity. To draw a quantitative comparison of the accuracy of the system by Rahman et al. vis-à-vis our system, we have measured the approximation errors of the system outputs using the MSE method. It is found that the MSE of the Braille lines, circles and diagrams drawn in the other system is 2 to 3 times greater or even more than that in our system. Thus it can be stated that unlike other systems, our system produces tactile diagrams from textual description automatically and that too with greater accuracy.

User evaluation and impact analysis

Over a period of nearly seven months pre and post experiment study was conducted at Blind Boy’s School, Narendrapur. Members of another professional organization for the blind people namely, Blind Person’s Association, Baruipur also participated in the study. Both the institutions are situated in Kolkata, India. The study and analysis reported here follows the approach similar to that adopted by Chan et al. (2006) and Huang & Shiu (2012) in education technology survey.

Study of existing practice

Given the standard of average blind schools in India, the Narendrapur school is quite rich in resource. They have couple of desktop Braille printers or TED (text embossing device) and uses DBT (Duxbury Braille Translator) software to create Braille documents and print those in raised-dot form using TED. However, not many teachers and
students have the skill of producing even simple line diagrams with DBT and TED. Still slate/stylus and nail-board system continues to be popular medium for teaching geometry in the class. The Baruipur Association has a costly Braille embosser or PED (plate embossing device) with compatible TGD (Tactile Graphics Design) software (Figure 6). But owing to the complicated process of producing tactile documents with images PED is rarely used.

Figure 6. TGD QikTac drawing interface; simple geometric shapes can be drawn on screen and then embossed using costly image embossers

Figure 7. Simple geometric shapes printed in Braille using DBT software

Need analysis

To find out the limitations and problem areas for teaching and learning geometry in a classroom of the blind students, a questionnaire survey was conducted amongst the students of grade VIII to X and teachers of the blind school separately. Group discussions were also held involving all the science teachers and a group of working professionals who are blind. Following is the summary of the facts that came up after qualitative analysis of data collected through questionnaire survey and discussion.

The school does not have enough funding to afford an image embosser which would have been an ideal solution for generating tactile diagrams to help teaching geometry. Drawing diagrams using DBT-TED is a tedious process. One has to type 6-dot characters at suitable positions such that the character chain represents a geometric shape (Figure 7). The majority’s opinion was that the nail-board or slate-stylus system is more effective than TGD or DBT though only simple diagrams can be tried with those. Overall, the inconvenience of sharing and accessing diagrams has practically limited teaching geometry to dictating facts and procedures only, deliberately omitting any reference to diagrams. The students also have option to avoid geometry questions in mathematics examinations. Some blind teachers revealed that they themselves were deprived of geometry education owing to lack of learning aids. The exercise revealed the need for a low-cost utility requiring minimal manual operation to serve as a self-learning tool for the blind students as well as a useful teaching aid for diagram-based subjects.
Selecting test subjects

To get a subjective assessment of how different categories of users react to our prototype, seven groups of test subjects as detailed in Table 3 were selected.

Table 3. User profile

<table>
<thead>
<tr>
<th>Groups</th>
<th>Composition</th>
<th>Average age</th>
<th>Visual Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I</td>
<td>4 students of grade VIII (2 male, 2 female)</td>
<td>13 years</td>
<td>Completely blind</td>
</tr>
<tr>
<td>Group II</td>
<td>4 students of grade IX (2 male, 2 female)</td>
<td>14 years</td>
<td>Completely blind</td>
</tr>
<tr>
<td>Group III</td>
<td>4 students of grade X (2 male, 2 female)</td>
<td>15 years</td>
<td>Completely blind</td>
</tr>
<tr>
<td>Group IV</td>
<td>2 senior teachers (1 male, 1 female)</td>
<td>55 years</td>
<td>Sighted</td>
</tr>
<tr>
<td>Group V</td>
<td>2 young teachers (1 male, 1 female)</td>
<td>30 years</td>
<td>Sighted</td>
</tr>
<tr>
<td>Group VI</td>
<td>2 technical professional (1 male, 1 female)</td>
<td>40 years</td>
<td>Completely blind</td>
</tr>
<tr>
<td>Group VII</td>
<td>2 non-technical persons (1 male, 1 female)</td>
<td>30 years</td>
<td>One person completely blind, the other person almost completely blind with deteriorating vision</td>
</tr>
</tbody>
</table>

The sample size in each group was kept small to impart effective training with limited number of trainers and also to ensure that each subject’s reaction or performance is observed closely. In the student groups, students from high, low and medium rank category were chosen to test whether the system is equally useful irrespective of the academic merit of the subject. In each group 50% female members were included because female literacy/education is a national issue in India; how the blind female subjects perform with a new learning tool may be significant for further study.

Training and test of recognition of simple geometric objects and relations

Firstly, the working mechanism of our integrated system, its input, and output were verbally explained to all the 7 test groups. Then practical demonstration of the system was given by presenting automatic generation of simple shapes like line, circle, rectangle, triangle (using Braille text printer) corresponding to input statements like “AB is a line”, “A circle is centered at O”, “ABCD is a rectangle”, etc. Next, figures representing relations between two geometric entities e.g., “AB is perpendicular to CD”, “Lines AB and CD intersect at E”etc. were generated and explained. Initially we found most students having difficulty in understanding the output as diagram – they rather read the diagram as Braille text, attempting to pronounce the individual characters that formed the diagram. The teachers were asked to perceive objects through touch only (without seeing) and their performance was similar. At this stage their responses with some tactile shapes printed on Braille sheet using our system was recorded and shown in Figure 8.

Each group was given 2-weeks initial training @ 1.5 hours each day until they were able to perceive Braille character chains as geometric shapes rather than text. In the beginning, concept of tracing the outline of a shape was given. We guided the students’ fingers over a figure-outline with verbal commentary and pointed out the key elements of the figure (e.g., sides and vertices of rectangle). Then we trained them to differentiate the shapes by emphasizing the attributes like orientation, length etc. Some students struggled to identify shapes as they did not carefully trace the entire outline. But most subjects could successfully retrace a shape without our help if they were prompted about the object and its starting/ending position. Gradually this support was withdrawn and after a week of practice @ 1 hour each day, the performance of each group in recognizing 10 different geometric objects were recorded (Figure 9a). The 10 objects are – point, line, square, rectangle, parallelogram, pentagon, trapezium, arc, circle and angle. Next, 10 geometric relations each involving two entities (namely, perpendicular lines, parallel lines, intersecting lines, diagonal of a rectangle, bisecting lines, bisector of an angle, tangent to a circle, diameter of a circle, two concentric
circles, two intersecting circles) were given to recognize after another 10 days of training-cum-practice @ 1 hour each day. Figure 9b shows the performance of each group.

Figure 8. Only 5% guesses were correct, the rest were incorrect; some guesses were very close like identifying a parallelogram as a rectangle

From Figure 9a we find the mean is 6.29 and standard deviation (SD) is 1.50 for the male subjects, while for the female subjects the mean and SD are 6.71 and 1.38 respectively. The t-value is 1 and $p=0.3559 (>0.1)$. In Figure 9b, mean and SD for male and female performances are 5.57, 1.72 and 5.86, 2.19 respectively. Here also the t-value is 1 and $p=0.3559 (>0.1)$. Therefore, the difference in performance of male and female in recognizing both object and object relation is not statistically significant. However, upon calculating the overall means (6.5 and 5.71) from both the charts we find the recognition accuracy diminishes by 12% when more than one entity features in the diagram.

Figure 9. (a) Recognition accuracy for objects generated in Braille, (b) Recognition accuracy for object relations generated in Braille

Students (particularly those belonging to Group II) were quite receptive and they learnt very fast. Out of 10 single objects and 10 configurations (with two objects) given for testing, 7.6 objects and 7.1 relations respectively were recognized correctly by the three student groups on an average. The average students’ performance over the two tests was therefore 7.41 or 74.1%. The average performance of the senior teachers was 45% while that of the young teachers was 62.5%. The performance of Group VI and VII were 37.5% and 60% respectively. It was noted that maximum failure occurred for circle, pentagon and parallelogram amongst objects while for tangents, bisector and concentric circles amongst the relations. Similar tests with different object-relation set were conducted at every one week interval and the performance recorded. After one month the average performance of the student groups (Group I, II & III), teacher groups (Group IV & V) and the remaining groups (Group VI & VIII) were 92.5%, 74% and 67% respectively which shows improvement with practice.
Test of classroom teaching-learning

6 experiments, spanning a duration of 3 months, were carried out in classroom environment. A teacher from Group IV or V was assigned to each of Group I, II, III and VI. Each of these four groups was given a desktop Braille text printer and the teachers were given desktop PCs with our system installed in it. Each time the teacher typed a geometry problem in his machine, the problem-diagram was drawn by the system simultaneously in the simulated form in the teachers’ machine and in the tactile form through the Braille printer accessed by the students. The test set contained 10 geometry word problems of varying hardness. It was observed that when a diagram was presented as a whole, the students at first could not follow most of the objects and their relations. It seemed difficult for the students to determine whether the elements within a bigger entity (e.g., two smaller lines within a triangle as in Figure 5b) were part of the outer outline or a detail inside the diagram. Then the teacher was asked to verbally dictate a problem one statement at a time to see whether the student could recognize those Braille objects mentioned in the statement. More help was then provided to the students who were still faltering by orientating them to a starting position (like the vertex “X” of a triangle “XYZ”) and along a tracing direction (like along a side XY from point “X”). The result was much better this time.

After this initial support, the recognition accuracy was measured in terms of number of diagrams perfectly identified and traced. If they misinterpreted any one object or relation of a diagram we considered it a failure. It was commonly found that as the number of geometric entities/relations increased the problem-diagrams became more complicated and therefore the subjects misinterpreted one or more objects/relations. With repeat tests conducted every 15 days (same problems given in random order), the performance of the subjects gradually improved (Figure 10). To minimize the chance that the students could memorize the patterns and therefore recognize correctly, they were not told about what they did was correct or not in each test. But in-between two tests usual practice (on an average 1 hour each day) was given on other simple shapes and diagrams drawn on Braille. In the last test (Test 6), Groups II and III emerged most successful (7 diagrams recognized correctly out of 10 i.e., 70% success). The score of other two Groups were 6 (60%) and 4 (40%) and the average final score of all the groups was 6 (60%). If we look at the respective minimum scores of 3 (30%), 2 (20%) and 1 (10%) in the beginning of the test series then the improvement achieved in 3 months is statistically significant (t-value 9.79, p=0.0023 < 0.01) and quite encouraging as far as trainability of the technology is concerned. This result also indicates the potential of our system as an effective classroom tool for the blind pupil.

![Figure 10. Gradual improvement of performance of each group; performance measured in terms of number of problems (along vertical axis) recognized correctly](image)

Test of self-learning

Finally the students’ self-learning ability upon using the system was tested with four groups for more than 1 month. A total of 10 problems were tested @ 2 tests every week. In between tests, the participants practiced themselves and training/help was seldom provided during this period. They were asked to type the text of a word problem as input to the system after reading the same already printed in Braille. Then they were asked to auto-generate the diagram in
Braille and explain the problem with reference to the diagram. All the subjects operated the system comfortably without any significant help and could correlate the problem texts with parts of the diagrams. The average time taken by each group from typing to tracing the auto-generated diagram for each problem is shown in Figure 11. Though initially the time consumed was quite high for all the test groups, later on it came down to 45 minutes for a 3 or 4 line problem. If we can pre-feed the problems in the system (as maximum time was taken for typing a problem) then time required by the subjects to generate and trace the diagrams would come down to around 15 minutes per problem.

![Figure 11. Time taken (in minutes) to draw and trace a problem using the new system](image)

Results and analysis

After a series of trainings followed by tests conducted for around 4.5 months, another questionnaire survey was conducted to get quantitative data on user perception about the new system. Beside the test groups (including 12 students and 8 seniors), 14 more students (i.e., a total of 26 students) participated in this survey as those students (who did not take part in the training) by this time learnt and used the system knowing from their peers of Group I, II and III. Questions were framed to bring out the average value of the evaluation indexes like satisfaction, ease of use, impact and effectiveness. The different indicators under each index and user response against those are illustrated in Tables 4 to 7. Each indicator is assessed on a five-point Likert scale that ranged from 1 (strongly disagree - SDA) to 5 (strongly agree - SA). The values in the columns SA to SDA are the percentage of number of users responded in respective category. Considering the response percentage against each category (SA to SDA) as frequency (f) of occurrence of respective scale values (x), we calculate the mean as \( \bar{x} = \frac{\sum xf}{100} \) and standard deviation as \( \sigma = \sqrt{\frac{\sum (x^2f)}{100}} - \bar{x}^2 \) against each indicator.

Table 4. Satisfaction

<table>
<thead>
<tr>
<th>Indicators</th>
<th>SA (5)</th>
<th>A (4)</th>
<th>N (3)</th>
<th>DA (2)</th>
<th>SDA (1)</th>
<th>Mean (( \bar{x} ))</th>
<th>SD (( \sigma ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfied with the tool and willing to use it in class now; could not use the existing tools so effectively</td>
<td>22</td>
<td>64</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>3.99</td>
<td>0.83</td>
</tr>
<tr>
<td>Perceived as a useful tool as any and many geometry problem can be discussed during a class</td>
<td>14</td>
<td>56</td>
<td>22</td>
<td>8</td>
<td>0</td>
<td>3.76</td>
<td>0.79</td>
</tr>
<tr>
<td>Perceived as an effective tool as there will be more time for teacher-student interaction</td>
<td>17</td>
<td>66</td>
<td>11</td>
<td>6</td>
<td>0</td>
<td>3.94</td>
<td>0.72</td>
</tr>
<tr>
<td>Perceived as a helpful tool as a student can practice geometry problems by his own thus enhancing self-efficacy and confidence level</td>
<td>33</td>
<td>56</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>4.22</td>
<td>0.63</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.98</td>
<td>0.74</td>
</tr>
</tbody>
</table>
The interface is user-friendly
The operations are simple
It requires no external help
Tracing geometric objects is easy
Flexibility of storing and reusing a diagram once produced

Average

Table 6. Impact

<table>
<thead>
<tr>
<th>Indicators</th>
<th>SA (5)</th>
<th>A (4)</th>
<th>N (3)</th>
<th>DA (2)</th>
<th>SDA (1)</th>
<th>Mean ((\bar{x}))</th>
<th>SD ((\sigma))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interesting</td>
<td>39</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.39</td>
<td>0.49</td>
</tr>
<tr>
<td>Enjoyable</td>
<td>19</td>
<td>58</td>
<td>17</td>
<td>6</td>
<td>0</td>
<td>3.90</td>
<td>0.77</td>
</tr>
<tr>
<td>Like</td>
<td>28</td>
<td>56</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>4.09</td>
<td>0.72</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.13</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 7. Effectiveness

<table>
<thead>
<tr>
<th>Indicators</th>
<th>SA (5)</th>
<th>A (4)</th>
<th>N (3)</th>
<th>DA (2)</th>
<th>SDA (1)</th>
<th>Mean ((\bar{x}))</th>
<th>SD ((\sigma))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accurately represents a problem</td>
<td>6</td>
<td>41</td>
<td>28</td>
<td>17</td>
<td>8</td>
<td>3.20</td>
<td>1.05</td>
</tr>
<tr>
<td>Minimizes user operation</td>
<td>28</td>
<td>72</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.28</td>
<td>0.45</td>
</tr>
<tr>
<td>Minimizes cognitive load of students</td>
<td>19</td>
<td>69</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>4.01</td>
<td>0.70</td>
</tr>
<tr>
<td>Minimizes effort of teacher</td>
<td>30</td>
<td>56</td>
<td>6</td>
<td>0</td>
<td>8</td>
<td>4.00</td>
<td>1.04</td>
</tr>
<tr>
<td>Increases learning outcome</td>
<td>33</td>
<td>64</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4.30</td>
<td>0.52</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.96</td>
<td>0.75</td>
</tr>
</tbody>
</table>

The average mean of the satisfaction index (Table 4) is 3.98, which implies user’s comfort and positive attitude towards using the system. Only 5% (average) users disagreed that the tool is useful, helpful and can be effectively used in the classroom. The Technology Acceptance Model theory (Davis, 1989) states that acceptance and use of a technology is determined by two factors: perceived usefulness and perceived ease of use. From the above statistics, the perceived usefulness is given by \(\bar{x} = 3.76, \sigma = 0.79\) (Table 4) and the average mean for ease of use is 3.70 (Table 5). Simple operation (\(\bar{x} = 3.96, \sigma = 0.85\)) (Table 5) and flexibility of storing the diagrams for future reuse (\(\bar{x} = 4.17, \sigma = 0.60\)) (Table 5) are the two aspects accepted by majority. However, there are several reasons behind low scoring against ‘Tracing geometric objects is easy’ (\(\bar{x} = 3.10, \sigma = 1.33\)) (Table 5). Firstly, the subjects were habituated in recognizing the 6-dot Braille characters as text only. With lot of practice and training they were gradually successful in recognizing the Braille character chains as part of diagrams. Secondly, there were some senior teachers and non-technical participants who had a mindset of not welcoming the new technology - they rather felt comfortable with the traditional methods even after accepting limitations of those methods. The young teachers are more receptive and can always be trained easily on this new technology.

Impact gives a general evaluation of the user’s reaction to the new system, especially when they used it for the first time. The average mean of 4.13 (Table 6) implies that the users liked it in-spite of possible impeding factors like typing a problem or tracing a diagram which may be time-consuming and training-dependent. Effectiveness measure as shown in Table 7 estimates whether the new system fulfills its specific educational and pedagogical objectives for the target users for whom it has been designed. The low score (\(\bar{x} = 3.20, \sigma = 1.05\)) against ‘Accurately represents a problem’ (Table 7), may be attributed to the inherent poor resolution and non-uniform spacing of the Braille dots which causes unevenness in shapes. It obviously does not mean that the Braille diagrams are geometrically incorrect, rather implies that the smoothness of the Braille lines or circles is not as good as those created using nail-board-elastic band/slate-stylus/PED which produce better perception of tactile geometric objects. Few participants had some previous exposure to figures drawn using those tools/devices. But as a teaching-learning tool our system is quite effective due to certain user-perceived advantages – firstly, diagram drawing is totally automated and it minimizes manual operations which in turn helps to reduce the cognitive load of the student and effort of the teacher.
in a class. Secondly and more importantly, the students can self-operate the system and learn without assistance. That’s why the average mean (3.96) of effectiveness (Table 7) and user’s perceived learning outcome (\(\bar{x}= 4.30, \sigma = 0.52\)) (Table 7) are quite high.

Conclusions

In the backdrop of limited facility and exposure of the blind students to tactile diagrams in many developing countries, one distinct contribution of our research is that we introduce tactile diagram drawing method based on traditional Braille printer (text-only) because such printer is the cheapest of its kind and commonly available in blind schools in India. The proposed method does not make use of any sophisticated interface rather relies on Braille character cells with uneven spacing of dots to map elements of a diagram already defined for digital display. The system evaluated in this study is capable of representing digital point, straight line, and circle, using Braille code; therefore, any simple geometric shape (like triangle, rectangle, angle etc.) and diagrams comprising many shapes and configurations can be produced using Braille text printer. The existing Braille graphics programs cannot produce shapes on Braille text printer and very limited research can be found on this mode of drawing.

Simply producing tactile diagram from digital diagram at low cost cannot resolve the actual difficulty of learning mathematics by the blind students. It could have certainly given them significant accessibility benefit, but it could not be used as a self-operated learning tool had it not been integrated with our earlier developed text-to-diagram conversion system (that produces digital diagrams from textual description in natural language (English)). As a result of this integration, a blind student can create and access diagrams without any human assistance. The system is also tested to be helpful for the teacher as it significantly eases his effort in creating and presenting accessible graphics for the blind students in a geometry class. In the review of NLP-based systems (Mukherjee & Garain, 2008) we hardly find any automated approach for producing diagrams (for blind learners) directly from word description of mathematical problems.

The core contribution of this study is the series of field tests conducted with blind students studying at different grades, teachers and blind persons in profession. To evaluate the actual utility of the low-cost Braille version of geometry diagrams we have tested the integrated system as a whole. The acceptance of our system is reflected in the interest generated in our subjects; most of them could recognize the Braille shapes with reasonable accuracy and could relate the Braille diagrams to the problem statements within a short learning time. The results of the usability tests, pre and post-test questionnaire survey underline the need and impact of such an affordable teaching-learning aid for the blind community.

The experiments with the system hint that there is enough scope for improvement in the Braille conversion algorithms. In future, we hope to make even better Braille approximation of graphic entities so that users do not confuse inner details with outlines of a diagram. Future measures like systematic training from lower classes, addition of voice guidance utility and step-by-step diagram drawing facility would definitely help in enhancing the recognition accuracy. Some other issues are to be addressed in future. One of them is addition of editing and scaling capability. Labeling of points on a diagram has been deliberately ignored in the present study (to avoid confusion of the subjects) and hence it is taken up in the 2nd phase of our research. We plan to integrate our text-to-diagram conversion module with other existing Braille graphics software tools (like QikTac, IVEO Viewer, etc.) so as to generate tactile graphics using high-end image embossers. Infrastructural constraints restricted our experiments with limited number of users. In future, experiment with larger sets of test subjects and word problems will bring out more general observations. Moreover, we can make further survey with a control group of users to compare the learning outcome in two systems.

References


A Study about Placement Support Using Semantic Similarity

Marco Kalz1*, Jan van Bruggen1, Bas Giesbers2, Wim Waterink3, Jannes Eshuis3 and Rob Koper1

1Welten Institute – Research Centre for Learning, Teaching and Technology, Open Universiteit Nederland, 6401 DL Heerlen, The Netherlands // 2Department of Technology and Operations Management, Rotterdam School of Management, Rotterdam, The Netherlands // 3School of Psychology, Open Universiteit Nederland, 6401 DL Heerlen, The Netherlands // marco.kalz@ou.nl // jan.vanbruggen@ou.nl // bgiesbers@rsm.nl // wim.waterink@ou.nl // jannes.eshuis@ou.nl // rob.koper@ou.nl

*Corresponding author

(Submitted May 5, 2013; Revised September 5, 2013; Accepted November 23, 2013)

ABSTRACT

This paper discusses Latent Semantic Analysis (LSA) as a method for the assessment of prior learning. The Accreditation of Prior Learning (APL) is a procedure to offer learners an individualized curriculum based on their prior experiences and knowledge. The placement decisions in this process are based on the analysis of student material by domain experts, making it a time-consuming and expensive process. In order to reduce the workload of these domain experts we are seeking ways in which the preprocessing and selection of student submitted material can be achieved with technological support. This approach can at the same time stimulate research about assessment in open and networked learning environments. The study was conducted in the context of a Psychology Course of the Open University of the Netherlands. The results of the study confirm our earlier findings regarding the identification of the ideal number of dimensions and the use of stopwords for small-scale corpora. Furthermore the study indicates that the application of the vector space model and dimensionality reduction produces a well performing classification model for deciding about relevant documents for APL procedures. Together we discuss methodological issues and limitations of our study whilst also providing an outlook on future research in this area.

Keywords

Methods, Assessment, Placement, Latent semantic analysis, Accreditation of prior learning, Networked learning

Introduction

In open and networked learning environments (Koper, Rusman, & Sloep, 2005) it is an important problem to support learners to find appropriate learning opportunities that fit to their competence level and prior learning. While traditional technology-enhanced learning environments are currently facing an innovation phase due to the widespread use of social and mobile media most of the time the assessment systems lag far behind the innovation processes. We see the development of assessment systems for open and networked learning environments as one of the grand challenges for technology-enhanced learning that makes it necessary to develop and evaluate alternative methods to approximate the prior knowledge of learners and to construct individual learning paths through their learning network (Janssen et al., 2010). Traditionally this problem has been addresses by intelligent tutoring systems with methods from adaptive hypermedia research like scalar models, overlay models, perturbation models or genetic models (Brusilovsky & Millan, 2007). However, it is rather a closed world in which these models operate; systems can only take account of what they represent/know about the learner in the current learning environment. All "system experiences" are lost after changing the learning environment and cannot be reused in another system. This problem has been recognized as the “open corpus” problem (Brusilovsky & Henze, 2007). While these models have their value in traditional e-learning processes in learning networks they are not applicable at all. Alternative bottom-up approaches are needed to offer personalized learning paths that do not need extensive sets of metadata to reason about prior knowledge of learners. Traditionally, higher education institutions in several European countries support the accreditation of prior learning (APL) (Merrifield et al., 2000). A typical APL procedure consists of four main phases (Van Bruggen et al., 2004):

- In a profiling phase the institution collects information about the learner’s needs and personal background.
- In the second phase learners collect and present evidence about their qualifications and experience. This evidence should support a claim for credit for the new qualification they are seeking.
In the *assessment phase* the evidence submitted by the learner is analysed and reviewed conforming to the local assessment standards. The result of this phase of the procedure is an answer to the question of whether the student should be granted recognition of their prior learning.

In the *accreditation phase* the results are verified by the department responsible for awarding the credit or recognizing the outcome of the assessment.

The procedures of APL are costly and time-consuming because they involve domain experts to assess the contents of the portfolios submitted by the students. There are two different approaches to accreditation in higher education institutions. On the one hand there is a generalized accreditation procedure based for example on certificates from vocational education which are expected to be equivalent to local certificates. On the other hand there is an individual accreditation procedure that also takes into account prior learning from non-formal and informal contexts. This second type of accreditation is seeking technological support models (Joosten-ten Brinke et al., 2008). These support models can range from a form of pre-advising the experts about which documents are relevant for the target course or study programme or it could help students to only fill their portfolios with material that is relevant to possible exemption decisions. At the same time these models have the potential to contribute to a future research agenda for technology-enhanced assessment in open and networked learning environments.

The basic assumption of our research is that prior knowledge of learners can be approximated by the content of the learner portfolio and therefore overlap between the documents in the portfolio and the courses of the plan/curriculum can act as a proxy to give exemptions and provide a personalized curriculum. The learner portfolio can consist of a variety of documents for other educational or work contexts like study assignments, thesis work or a project report. In an earlier publication (Kalz et al., 2007) we have analyzed existing technological solutions to assess prior knowledge with technological support. We have focused there on three categories of technologies, namely content similarity, metadata and ontologies. In this study we have focused on the content similarity aspect. The similarity calculation is executed in our study with the help of a dimensionality reduction technique similar to Latent Semantic Analysis (LSA).

**LSA for prior knowledge approximation**

Latent Semantic Analysis (LSA) is a theory and method for extracting and representing the contextual-usage meaning of words by statistical computations (Landauer, 2007). The latent semantic analysis process consists of several steps: In the *indexing phase* all words and documents construct a so called Term-Document-Matrix (TDM), in which all terms are listed in the columns and all documents in the rows. After the counting of the occurrence of each word in all documents several weighting and normalization options are possible. In the *dimensionality reduction* phase a mathematical function called singular-value decomposition (SVD) is applied which is similar to factor-analysis. The end result of this process is a latent semantic space, in which the input documents are represented as vectors. Documents in this space are similar if they contain words which appear in the same context and so their vectors are close together in the space providing a measurement for the similarity of text. In the *retrieval or query phase* the query text is projected into the space and the distance to the document vectors is calculated via the cosine or the Euclidian mean. Since we cannot elaborate on LSA in more detail in this paper we recommend the papers mentioned above as an introduction to LSA.

Latent Semantic Analysis has been applied to several problems in the domain of Technology-Enhanced Learning such as peer tutoring (Van Rosmalen, 2008), provision of feedback (Graesser et al., 2005), automated essay scoring (Foltz, Laham & Landauer, 1999) and the selection of educational material (Graesser et al., 2000). Zampa and Lemaire (2002) applied LSA to the user modeling problem. In their model learners learn a domain by acquiring the most important “lexemes” of a domain. The most important crucial concepts are identified beforehand and the end result is a recommendation to look at the next best item based on the theory of proximal development (Vygotsky, 1978). Wolfe et al., (1998) focused on the use of LSA for the selection of educational material.

Our application of LSA is similar to those presented here but differs in a number of important aspects. In contrast to the approach of Zampa and Lemaire (2002) we do not aim to model the student background and the learning resources beforehand. Since we used text students have written in their prior education as a proxy to their prior knowledge we have a more dynamical model which could build on learning progress documented e.g., in an electronic portfolio. Our application of LSA also differs from a “simple” essay scoring scenario where a student text
is compared to one or more clearly defined gold standard texts which provide the basis for judging about the quality of an essay. In our case the topical range of the student documents can be potentially very broad while the target documents or course units can be very similar to each other. This makes it very important to find an ideal dimensionality that results in the best discrimination between target documents.

**Study**

In this paper we present a study that has been conducted at a Psychology Course of the Open University of the Netherlands. The foundations for this study are described in (Van Bruggen et al., 2004), and research agenda for it has been presented in (Kalz et al., 2008). The study employed a vector space model and dimensionality reduction technique inspired by Latent Semantic Analysis (LSA) to calculate a similarity between documents in the learner portfolios and the content of the course units. To test the validity of the results an expert validation with domain experts is performed and the performance of using LSA as a classifier for relevant or irrelevant documents was evaluated with a Receiver Operating Characteristic (ROC curve). ROC curves are a method from signal detection theory and they are often applied to binary classification experiments and model performance evaluation.

The hypothesis tested with this study is as follows:

H1: LSA can classify documents as relevant/irrelevant in a manner comparable to human experts.

Our target for this study is to minimize the false-negative and the false-positive cases under a threshold level of 10%. Too many false-negative cases would hinder students from exemptions while too many false-positive cases would result in unnecessary work for the domain experts. In addition we checked the results we reported in a previous study about the identification of an ideal number of dimension for small corpora (Van Bruggen et al, 2006). However it is important to note that this study was not intended to evaluate the use of automated exemption rules for study programs but is intended to evaluate the applicability of the vector space model and dimensionality reduction techniques for learner placement in general. The exemption rules and problems associated with them such as trust issues, validity of thresholds etc. are not a specific problem of the method we are evaluating but of the general exemption procedure. This means that exemptions standards can differ between institutions and that they have to set their own thresholds for exemptions.

**Method**

**Study context and participants**

The study was carried out in the context of an introductory psychology course of the Open University of the Netherlands. Study participants were drawn from the 244 students that enrolled in the course around 7% submitted material to apply for the accreditation of prior learning. Overall we had a total number of 18 participants providing 28 documents to be compared to the units of the target course. Thus the total number of similarity ratings was 504 (28 documents x 18 course units).

**Study procedures**

The introductory psychology course consisted of 18 units each of them dedicated to a subtopic of psychology. The course was offered in an online environment. Before students could enter the course they had to read an introduction about the content of the units of the course. Thereafter, the students filled out a questionnaire on any prior knowledge for the course or parts of the course. Students were invited to submit materials they had produced in their prior education or working environments. Documents submitted were work reports, (bachelor / master) theses, technical reports, essays, reference lists and presentations.

We have chosen the ideal dimensionality for the study based on two performance criteria. On the one hand we employed the connection between singular values and the variance they account for in the Term-Document matrices analyzed by LSA. Our target was to reduce the variance accounted for under a threshold of 70 – 80 % of variance represented in the data. This approach is in line with other studies that have a comparable purpose. Then we used the corpus documents as queries to control the discrimination between the chapters. These queries were tested for self-
correlations and discrimination to the other chapters under different conditions. We varied the dimensions used between 5 and 1000 and we used different stopword settings (no stopword list, 30% stopword, 50% stopword list). Stopwords were removed because they either appeared very infrequently or very often in the corpus. In addition we also evaluated the use of local and global weighting options. In this process we followed a method by (Rosmalen et al., 2006) to calibrate and test several LSA parameters.

In order to evaluate the results of the study two domain experts independently rated the similarity between the student documents and the domain documents on a 5-point Likert-scale. For each student document the experts evaluated the semantic similarity and marked if they would give an exemption. In addition they wrote down how much time it took to review the material. We also interviewed one of the domain experts. We calculated a raw overall percentage of agreement between the two raters. We calculated the interrater agreement according to the consensus and the consistency of the ratings by the two judges (Stemler, 2004). The consensus of the ratings by the two domain experts was calculated using Cohen’s Kappa. The Spearman rank coefficient was used to calculate the consistency among the ratings. For the performance assessment we recoded the Likert scale into a binary scale. Here we used the most optimistic rule that a document is seen as relevant when at least one of the raters has rated the document with a value higher than 1.

The model performance assessment of our method as a classifier for relevant and irrelevant documents for APL procedures was analyzed via a confusion matrix and ROC-curve (Fawcett, 2003). ROC-Curves (Receiver Operating Characteristics) are a method from signal detection theory that has been applied to evaluate model performance assessment of classification models. In ROC curves the true positive rate (tpr) of a classifier is plotted against the false positive rate (fpr) while varying the thresholds used for the classification. One of the main advantages of the application of ROC-curves is that they are not sensitive to class skew (Hamel, 2009). With this approach we have also compared the effect of applying different weighting functions reported in the literature to contribute positively to performance in small-scale corpora (Nakov et al., 2001; Wild et al., 2005). We have compared the use of a logarithmic weighting function for a local weighting and the use of entropy and inverse document frequency for the global weighting and the combination of logarithmic and entropy weighting.

In addition we have calculated raw success scores on the document level (How many documents would have been recommended right?) and the person level (How many learners would have been exempted right if LSA would have decided about exemptions?). The corpus for the experiment consisted of the content of the 18 psychology learning units. This corpus had 28165 terms with 490431 occurrences. The corpus size was 3.1 MB. 13283 terms only appeared once in the corpus. After keeping only terms that appeared more than once the corpus was reduced to 14882 terms with 477147 occurrences.

All corpus and learner documents have been manually cut into paragraphs. The paragraph length was between 250 and 500 words. The corpus consisted in the end of 2246 paragraphs. In this study all analysis was done using the Text to Matrix Generator (TMG) - a Matlab implementation of Latent Semantic Analysis and other vector space techniques (Zeimpekis & Gallopoulos, 2006). For the experiment a script was written that calculates the mean of cosines for all paragraph to paragraph comparisons and writes down the mean correlation to all 18 chapters in a spreadsheet file.

Results

Dimensionality reduction and sensitivity

For the estimation of the ideal dimensionality we were able to reproduce results obtained in a prior explorative study (Van Bruggen et al., 2006). Figure 1 shows our research corpus with different stopping strategies applied and with different numbers of singular values.

In this figure we can see that the variance zone of 80% variance accounted for or higher starts with different numbers of singular values depending on the stopping strategy used. For no stopping this zone starts at a reduction to 194 singular values. For the 30% stopping strategy this zone starts at 530 singular values while it starts for the 50% stopping strategy at 713 singular values. The second performance criterion was tested with the discrimination between the target learning units. Figures 2 – 4 illustrate the self-correlation and the correlation to the other learning
units for the learning unit 1. These figures show the relation between singular values extracted for the analysis in relation to cosine similarity between the units. The ultimate goal for the ideal number of singular values and stopwords would be that there is a clear distinction between the units to be able to discriminate between the relevant and obsolete target learning activities. Note that we did not analyze the full range of singular values but only a selection since the figures are only used for illustrative purposes.

![Figure 1](image1.png)

*Figure 1. Correlations and variance explained for different numbers of singular values and 3 different stopping strategies*

![Figure 2](image2.png)

*Figure 2. Correlations and self-correlation for learning unit 1 with no stopping*

As we can see in figure 2 it is not possible to discriminate between unit 1 and the other learning units of the course. All units are too similar to each other to be able to discriminate between them. In figure 3 we have increased our stopword strategy to 30%.
As we can see in figure 3 the cosine values drop and discrimination between the chapters improves with a 30% stopping strategy. But still it is not easy to clearly discriminate between the chapters. In figure 4 we increased our stopping strategy to 50% stopwords.

Overall we can see that the cosine values drop the more stopwords we apply, but at the same time the discrimination between the target documents in the corpus improves sufficiently to discriminate between the chapters. This effect could be replicated for all 18 units. After having defined the ideal number of factors and stopwords for the analysis we have used these parameters (800 factors with a 50% stopword list) for querying the student documents and comparing the cosine values to the 18 learning units.
Model performance assessment

The raw agreement between the two domain experts was 95%. This high agreement between raters was mainly based on the high number of submissions not rated as relevant for the APL procedure. The interrater reliability for the raters was found to be Kappa = 0.77 (p < 0.001). According to Spearman's rho (1904) there was a high consistency between the ratings (rho = 0.75, p < 0.01). After the recoding into a binary classification the interrater agreement was Kappa = 0.74 (p < 0.001).

Only 8% of all cases (40 cases) were evaluated as relevant for APL. Because of this the data were negatively skewed. The expert data confirmed our basic assumption that content similarity is related to exemptions in APL procedures. If we take into account only cases with content similarity higher 2 on the Likert scale then 83% of the cases have been proposed for an exemption in the mean of both raters. Seven learners would have been given exemptions based on the decisions of experts which equals 38% of all participants who submitted material for the study.

In our evaluation of the impact of different weighting functions we have compared five parameters as shown in Figure 5. This figure is a plot of the true positive rate on the y-axis and the false positive rate on the x-axis.

![Figure 5. Receiver operating characteristics curve (ROC curve) for LSA with different weighting parameters](image)

There are several important aspects to be reported. In general, we can report that without taking into account the different weighting functions our method can be clearly distinguished from a random classifier that would rate documents by chance into each of the categories. Overall the use of the weighting functions reported in the literature has slightly decreased the performance of our model and the best performance could be reached without any local and global weighting. To compare this performance in more detail we provide in table 1 an overview of the size of the Area under the ROC curve (AUC) which can be used to evaluate the performance of a classification model.

<table>
<thead>
<tr>
<th>Weighting</th>
<th>AUC</th>
<th>Std. Error.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No weight</td>
<td>0.811</td>
<td>0.0262</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>0.777</td>
<td>0.0313</td>
</tr>
<tr>
<td>Entropy</td>
<td>0.708</td>
<td>0.0413</td>
</tr>
<tr>
<td>Log./Entropy</td>
<td>0.741</td>
<td>0.0386</td>
</tr>
<tr>
<td>IDF</td>
<td>0.697</td>
<td>0.0417</td>
</tr>
</tbody>
</table>
We can see that the inverse document frequency function (idf) performs worst for our model and that the logarithmic function was the second best option. If we also take into account the confidence intervals of the ROC curves we have to summarize that the weighting functions decrease the performance but not significantly because the confidence intervals of all settings overlap as we can see in figure 6.

Figure 6. Impact of local and global weighting to the model performance assessment

Thus we have decided to use the 800 dimensions, a 50% stopword list and no weighting for our model. For this settings human raters and LSA results show only a moderate agreement of Kappa = 0.33 (p < 0.001). The mean of the ratings was 0.11 (Std. error = 0.21) while the mean variance was 0.23 and the mean standard deviation 0.47. To discuss the performance of our model more into detail we provide an overview of the results with a confusion matrix (table 1). We can see that LSA could successfully classify 88 % of all cases right. The false positive cases are 8 % of all cases while 2 % of the cases fall into the false negative category.

Table 2. Classification results Human ratings vs. LSA ratings (n = 504)

<table>
<thead>
<tr>
<th>LSA rating</th>
<th>Human rating</th>
<th>Relevant</th>
<th>Irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant</td>
<td>28</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Irrelevant</td>
<td>12</td>
<td>417</td>
<td></td>
</tr>
</tbody>
</table>

The sensitivity/true positive rate of our model is 0.7 while the specificity is 0.91. The false positive rate is 0.08. Overall the negative predictive value of our model is 0.97 while the false discovery rate is 0.58. The Area Under the ROC-Curve (AUC) for our model was 0.81 (95% CI, Std Err. = 0.0262). Overall this is a sign of a good predictive model.

On the document level we have analyzed how many documents of the 36 would have been passed to the human experts if our LSA model would have been implemented to evaluate relevant documents. Here we have compared two different methods: The mean of all cosine values to each chapter for every learner document and the maximum cosine value of the comparison between learner document and all 18 chapters. The raw percentage of right classified documents of the mean method was 85 % while it was only 43 % with the maximum method. On the level of the learner our model would have recognized 7 of 12 given course unit exemptions of the human raters but it would have added 40 false positive exemptions on top for these 7 learners. Overall human experts needed in the mean 255 minutes to analyze the material provided by the students.

Discussion

Based on a method to estimate the ideal number of dimensions retained in an LSA space we could reach sufficient discrimination between the target documents and identify the ideal number of dimensions for our study. With these...
results we confirm our findings from an earlier study on LSA and small-scale corpora. This lets us conclude, that the application of a stopword strategy between 30% and 50% is needed to reach sufficient discrimination between documents for placement support on small-scale corpora. After testing different weighting options we could show that weighting does not improve the performance of our model. Overall we could reach a satisfactory classification model because of two reasons. First we could reach our self-set target of less than 10% false positive and false negative cases. Second, an AUC value of 0.81 is seen as a good performance indicator for a classification model.

In summary we have to reject the null-hypotheses showing that LSA can successfully discriminate sufficiently between relevant and irrelevant documents according to the targets we have set ourselves.

But there are several limitations to the here presented study. First of all the low number of participants and especially relevant documents is problematic in case of generalizability of the results. The data we collected for an APL procedure resulted in negatively skewed data. After discussing these findings with an APL expert we discovered that students are often not aware of what to submit for exemptions. This means that collecting more student data would likely lead to a similar skewed dataset. In fact, it is a part of the problem we are trying to solve with this technology.

Some of the false negative cases reveal that human experts decide about exemptions on more factors then just the semantic similarity between learner documents and target documents. One case was especially interesting in this regard: One of the learners submitted a very detailed description about an experiment they had conducted. This description did not contain sufficient semantic concepts which illustrated a relationship to the target documents. But human raters deducted from the document that this learner must have specific prior knowledge in psychology to be able to write such a document. For this purpose other techniques and approaches to approximate prior knowledge are needed which go beyond semantic similarity. In related experiments we have focused on semantic networks and other ways to visualize content of sets of documents for visual inspection (Kalz et al., 2009; Berlanga et al., 2011). This could address the problem of the false negative since there would be a combination of methods which would address most false negative cases from the study.

The qualitative interview showed us that the analysis model of the human experts is based on semantic similarity of documents but the cognitive process is more complicated. One domain expert described the analysis process with different steps involving keyword analysis, semantic analysis and quality ranking of the student documents. While our study confirms that the application of dimensionality reduction techniques like LSA for the support of APL procedures and the approximation of prior knowledge is a promising research and development direction the approach needs to be validated in several different contexts. In this regard we expect that a training phase is needed in all implementations to align the approach to local thresholds and local decision boundaries.

We believe that the work presented here can stimulate research about alternative bottom-up assessment methods that will play an important role in future open and networked learning environments. From a European perspective especially the development and evaluation of similar approaches in a multilingual context is a challenging research direction.

References


The Effects of Game-Based Learning on Mathematical Confidence and Performance: High Ability vs. Low Ability

Oskar Ku\textsuperscript{1}, Sherry Y. Chen\textsuperscript{2}, Denise H. Wu\textsuperscript{3}, Andrew C. C. Lao\textsuperscript{2} and Tak-Wai Chan\textsuperscript{2}

\textsuperscript{1}Department of Computer Science & Information Engineering, National Central University, Taiwan // \textsuperscript{2}Graduate Institute of Network Learning Technology, National Central University, Taiwan // \textsuperscript{3}Institute of Cognitive Neuroscience, National Central University, Taiwan // oskar@cl.ncu.edu.tw // sherry@cl.ncu.edu.tw // denisewu@cc.ncu.edu.tw // andrew@cl.ncu.edu.tw // chan@cl.ncu.edu.tw

\textsuperscript{*}Corresponding author

(Submitted May 7, 2013; Revised September 12, 2013; Accepted November 7, 2013)

ABSTRACT

Many students possess low confidence toward learning mathematics, which, in turn, may lead them to give up pursuing more mathematics knowledge. Recently, game-based learning (GBL) is regarded as a potential means in improving students’ confidence. Thus, this study tried to promote students’ confidence toward mathematics by using GBL. In addition, this study also investigated whether GBL is beneficial to all students with various abilities. The results demonstrated that this approach yielded better outcomes than the paper-based setting in both students’ confidence and students’ performance. The students with high and low levels of ability in the GBL group gained a significant improvement on the confidence toward mathematics. Additionally, low-ability students in the GBL group attained better mathematics performance than those in the paper-based setting.

Keywords

Game-based learning, Confidence, Mathematics, Performance, Ability

Introduction

Mathematics is a fundamental skill in our daily life. Humans have been applying mathematical knowledge for over 4000 years. In addition, Gauss referred mathematics as the queen of sciences (von Waltershausen, 1856), which implied the importance of mathematics. However, students often perceive mathematics as a difficult subject (Stodolsky, Salk, & Glaessner, 1991). Low confidence is one of the critical reasons that makes students feel difficult to learn mathematics. Such a negative feeling may consequently make a student give up learning mathematics (Brown, Brown, & Bibby, 2008).

In other words, self-confidence plays an important role in learning (Maclellan, 2014) because it is a predictor of a learner’s learning behavior, such as the degree of effort made and the expectation of outcomes (Schunk, 1990). Students with high self-confidence may attain better performance in tasks (Kleitman, Stankov, Allwood, Young, & Mak, 2013) and engage in target tasks more actively (Gushue, Scanlan, Panzer, & Clarke, 2006) than those who are less confident about the tasks. In addition, students with high self-confidence usually regard difficult tasks as meaningful tests (Bandura, 1994) while those with low self-confidence tend to avoid calling for help (Ryan, Patrick, & Shim, 2005). Thus, there is a need to give additional support to students with low confidence toward mathematics.

Past studies found that digital games had potentials to enhance students’ confidence (Cunningham, 1994; Radford, 2000). Furthermore, digital games can also enhance students’ learning motivation (Klawe, 1998; Nussbaum, 2007) and their learning performance (Ke & Grabowski, 2007). Therefore, embedding math learning into digital games may be a possible solution to enhance students’ self-confidence, learning motivation and learning performance.

To this end, this study investigates whether digital games can be adopted to enhance students’ confidence toward mathematics and meanwhile to improve students’ learning performance, especially for those with a low level of self-confidence toward mathematics. In addition, although past studies found the positive effects of GBL, the effects were usually reported holistically. However, it is not clear whether every student could benefit from GBL. Since the issue of individual difference is more and more important in the design of learning environments, this study further investigates how students with different levels of academic ability react to GBL. More specifically, the research questions of this study can be summarized as follows:
• Whether GBL enhances students’ confidence and learning performance toward mathematics?
• Whether students with different levels of academic ability react similarly to GBL?

Theoretical background

Confidence is a critical element in learning; past studies indicated that confidence correlates positively with performance (Al-Hebaish, 2012). This strong and positive correlation was also reported in the area of mathematics education (Stankov, Lee, Luo, & Hogan, 2012). Confidence affects an individual’s learning in various aspects. For example, there is a correlation between confidence and the elective enrollment of mathematics courses (Kleanthous & Williams, 2011; Metie, Frank, & Croft, 2007). Confidence also affects an individual’s effort; a learner with low confidence might not pay full effort to complete the task (Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998).

Recently, the results of the Trends in International Mathematics and Science Study (TIMSS) showed that Asian students tended to have low confidence toward mathematics (Mullis, Martin, & Foy, 2008). Comparing the confidence level from the eighth grade and the fourth grade, students at the eighth grade had lower confidence than students at the fourth grade. This result implies students are losing their confidence toward learning mathematics by the increase of their age. This is a warning message because low confidence may also lead students to have negative reactions to mathematics (Brown, Brown, & Bibby, 2008), which, in turn, may make them give up learning mathematics. Thus, there is a need to help students build up their confidence about learning mathematics.

Digital games may be a possible solution to address this issue. Although not all GBL studies consistently reported positive results (e.g., Chen, 2012; Ke, 2009; Kebritchi, Hirumi, & Bai, 2010; Sitzmann, 2011), the majority of the studies revealed positive effects brought by applying digital games to support learning environments. For example, GBL engages students in learning activities (Huizenga, Admiraal, Akkerman, & ten Dam, 2009; Shute, Ventura, Bauer, & Zapata-Rivera, 2009) and stimulates students’ learning motivation (Dickey, 2007; Nussbaum, 2007; Tüzün, Yılmaz-Soylu, Karakuş, Inal, & Kızilkaya, 2009). In particular, GBL was found to have the potential to enhance students’ confidence (Cunningham, 1994). For example, Radford (2000) argued that, through successfully manipulating three-dimensional objects in simulation game environments, his architecture students gained opportunities to build up confidence toward related architectural tasks. Such research implies that embedding math learning into digital games may be a possible solution to enhance students’ self-confidence toward learning mathematics.

However, previous studies mainly focused on two dependent variables: achievement and motivation (Table 1). Paucity of them investigated the effect of digital games on learners’ confidence. As mentioned earlier, confidence is also an important element, which influences the learning outcome of a learner. Thus, this study aims at investigating the effects of GBL on learners’ confidence.

<table>
<thead>
<tr>
<th>Study</th>
<th>Dependent variable(s)</th>
<th>Result(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ke &amp; Grabowski (2007)</td>
<td>Achievement</td>
<td>Students who played games had a better math performance than those who received paper-based drills.</td>
</tr>
<tr>
<td>Ke (2008)</td>
<td>Achievement</td>
<td>Students had more positive attitude on computer math gaming, but there was no significant outcome for the performance and metacognitive awareness development.</td>
</tr>
<tr>
<td>Owston, Wideman, Ronda, &amp; Brown (2009)</td>
<td>Achievement</td>
<td>Students who learned by a game development shell had a better performance than those who did not play games.</td>
</tr>
<tr>
<td>Papastergiou (2009)</td>
<td>Achievement</td>
<td>Students who played games had better performance and were more motivated than the non-gaming students.</td>
</tr>
<tr>
<td>Suh, Kim &amp; Kim (2010)</td>
<td>Achievement</td>
<td>Students who used online role-playing games for English learning had better performance than those who received face-to-face instruction.</td>
</tr>
<tr>
<td>Kebritchi, Hirumi, &amp; Bai (2010)</td>
<td>Achievement</td>
<td>Students’ achievement was improved as a result of applying GBL; however GBL did not produce significant improvement on motivation.</td>
</tr>
</tbody>
</table>
In addition, past studies reported the positive effects of GBL in a general way; however, the positive results may be contributed by a specific group of learners, instead of all learners. Because different pedagogies may be favored by different learners with different learning styles (Clewley, Chen, & Liu, 2011), GBL may not also be beneficial to all learners. For example, the multimedia elements of game-based-learning may also increase students’ cognitive load (Huang, 2011) because they have to play and learn at the same time; it might go against learners with low ability. Researchers should put more emphasis on individual differences when designing new pedagogies. On the other hand, past studies indicated that students with different academic abilities performed differently (Colquitt, LePine, & Noe, 2000). More specifically, students with a high level of academic abilities usually perform well on tasks, and vice versa. For example, MacCann, Fogarty, Zeidner, and Roberts (2011) found that students with high academic ability have higher emotional intelligence and have better problem coping performance. Thus, students’ academic ability is an important variable that may correlate with students’ task performance and should be taken into consideration when researchers analyze the effects of pedagogy. To this end, this study not only examines the effect of GBL on learners’ confidence and achievement, but also attempts to further verify the effect on students with different levels of academic ability.

### Methods

#### Experimental design

An experiment was designed to examine the effects of GBL, with a focus on confidence toward mathematics and learning performance. To this end, an experiment group (EG) learned in digital game-based environment was arranged to examine the effect of GBL while a control group (CG) learned in a paper-based condition was set to provide a comparison. Table 2 describes the distribution of participants in each group. All the participants were fourth-grade elementary school students between 10 to 11 years old. Because of the limitation of school administration and the students’ normally distribution to each class based on their academic performance, the participants were recruited via convenience sampling (i.e., two classes of students were randomly selected from all classes in the fourth grade). One of the two classes was randomly considered as the EG while the other was regarded as the CG. The experiment was conducted twice a week for five weeks; each session lasted for 20 minutes.

A computational test and a confidence scale were administered before and after the experiment. In addition, students in both groups were further divided into two subgroups: high-ability and low-ability, according to their general math performance—to investigate whether GBL is beneficial for both high and low-ability students. Thus, the experiment forms a 2 (between-subject variable: EG vs. CG) by 2 (between-subject variable: high-ability vs. low-ability) by 2 (within-subject variable: before- vs. after-learning) mixed design.

| Table 2. The numbers of the participants in the EG and CG |
|-------------|----------------|----------------|
|             | Male | Female | Total |
| EG          | 12   | 14     | 26    |
| CG          | 10   | 15     | 25    |
Instruments

Materials

This study chose mental calculation methods as the subject content of the learning material. The result of students’ previous mathematics midterm exam (Mean = 83.78, SD = 14.111, Full score = 100) demonstrated that students had prior knowledge of basic multiplication and division, irrespective of the EG or CG. In other words, all of the students had the prerequisite concepts of the learning material used in this study. To avoid the interference of standard courses, this subject matter was not taught in regular math courses. In order to prevent students from applying computation strategies mechanically, 20–30% questions that require students to refer to simple number facts were included in the learning material. The mixed patterns of arithmetic problems require students to think carefully, not just repeatedly applying the same strategy.

Mini-games

As suggested by Kafai (2001), the design of GBL can be categorized into tight coupling and loose coupling, in terms of the relationship of learning content and gameplay in GBL. The loose coupling makes the learning content interchangeable and requires few resources to implement a game, but designers may lose their attention on the specific parts of the learning content. Conversely, tight coupling puts more emphasis on the essence of the learning content, which could rich learning experience, but there is a need to use more resources to implement a game which is not reusable with different leaning content. In other words, both methods have their own strengths and weaknesses and have been applied in different contexts. For example, Anderson & Barnett (2013) applied the tight coupling approach to develop a digital game, Supercharged, to deliver basic electromagnetic concepts. The result of that study demonstrated that the game helped students obtain more learning gain and the insight of electromagnetic concepts than those learned in a traditional setting. On the other hand, Papastergiou (2009) developed a loose coupling game to teach computer memory concepts. The result indicated that students with the game-based environment gained more improvement on computer knowledge and learning motivation than those with a non-game approach. In brief, both tight coupling and loose coupling have positive effects on student learning.

However, loose coupling was selected for this study. This is due to the fact that mental calculation requires a high level of concentration for students to produce an answer, so there is a need to reduce students’ cognitive load by using the loose coupling approach. More specifically, two additional mini-games were designed in this study. One named as Battleship (Figure 1) is similar to a board game and is a turn-based strategy game in competition with a virtual competitor. A student must answer each question within a limited time; a correct answer yields an opportunity to place a bomb on the opponent’s board. However, an incorrect answer makes the student lose the chance to attack the opponent, and the virtual competitor can take its turn. The other named as Math Kicker is a soccer game, with which the participants are very good at playing (Figure 2). In this game, a student plays as a soccer player who is ready to shoot a goal. A correct answer of a question yields a successful shoot and the score is increased.

Figure 1. A screenshot of Battleship
Due to the fact that the learning material used in this experiment was mental calculation, speed is a primary factor. Therefore, the two games were designed as time-constraint games. Although students in this experiment learned individually, virtual competition was added into the mini-games to engage students in the learning activity. On the one hand, playing against opponents creates a socio-competitive situation that can promote active engagement and can provide immediate feedback to players (Vorderer, Hartmann, & Kimmt, 2003). On the other hand, the virtual opponents simulated students’ abilities. More specifically, the virtual opponents are often tied with students or a bit behind/beyond students during the process of playing the games. Thus, all students received adequate challenge meeting their abilities and could have a big chance to win unless they kept making errors. By doing so, students can immerse in the flow state (Csikszentmihályi, 1990).

Measurement

Confidence

In order to determine changes in confidence toward mathematics as a result of GBL intervention, this study employed a questionnaire modified from the confidence sub-scale of the Fennema-Sherman mathematics attitudes scales (Fennema, & Sherman, 1976), which has widely been adopted to investigate students’ opinions toward mathematics (e.g., Rattan, Good, & Dweck, 2012).

Mathematical performance

Mathematical proficiency (Kilpatrick, Swafford, & Findell, 2001) includes five components, of which the “procedural fluency” is highly related to the mental calculation. On the one hand, procedural fluency is the basis of math learning, which helps students understand that mathematics is well structured knowledge (Kilpatrick, Swafford, & Findell, 2001). On the other hand, procedural fluency emphasizes the flexibility, accuracy, and efficiency of computational skills, which is the goal of mental calculation, i.e., the subject matter of this study. Therefore, students’ computational skill was measured to examine the effect of the intervention. Two test papers for the pretest and the posttest were created to measure students’ computational skill. The test papers comprise a set of computational questions, which require students to calculate and write down their answers (e.g., 23×20 =). To avoid a ceiling effect, 250 questions were included in the test papers, so that no students could complete the test during 20 minutes. Both of the test papers for the pretest and posttest were created by two elementary school teachers; the tests had also been reviewed by the other two experienced teachers to adjust the difficulty of questions. In order to make it possible to compare students’ performance between the pretest and posttest, the level of difficulty of the pretest and posttest was maintained: the questions used in the posttest were the same as those in the pretest, except for the order of the questions. In addition, because the subjects were recruited via convenient sampling and the learning material was related to mathematics, a general mathematics test was administered before the intervention to identify whether the subjects in the CG and the EG possessed similar level of prior mathematical knowledge.
Procedures

As shown in Figure 3, the pretest and the posttest were administered a week before and after the experimental intervention. During the experiment, nine learning sessions were conducted for five consecutive weeks. Students could access the learning content via a digital document (EG) or via a paper-based handout (CG). The learning materials presented to both groups were completely the same in each session, except for the presentation medium and error feedback. The contents of the error feedback, which are the same in both conditions, include the place where an error was made and a message asking the students who committed the error to re-answer the question until the correct answer was made. Students in the EG received immediate feedback while students in the CG received the same message after their teacher marked their worksheets. The teacher also highlighted the place where an error was made on students’ worksheets and asked students to revise their wrong answers until they got the correct answer. The only difference is the timing of feedback of which immediate feedback is the innate characteristic of digital games.

![Figure 3. The procedure of this study](image)

Results and discussion

Due to the small sample size, the Mann–Whitney U test, and the Wilcoxon’s matched-pairs signed-ranks test were employed to conduct data analyses. The results from the Mann-Whitney test shows no significant difference between the EG and the CG (\(U = 312.500, p = .813\)). In other words, the EG and CG had a similar level of prior mathematical knowledge (Fig. 4).

![Figure 4. Students’ mean scores of the general mathematics test](image)
Comparison between the EG and the CG

To answer the first research question, students’ confidence and computational performance were analyzed and reported in the following subsections.

Confidence

Students’ confidence toward mathematics is shown in Figure 5. The result of the Mann-Whitney test revealed a significant difference of the gain score of confidence between the EG and the CG ($U = 175.000$, $p = .005$). In addition, the Wilcoxon’s matched-pairs signed-ranks tests revealed a significant increase of the confidence between the pretest and the posttest for EG students ($Z = 3.051$, $p = .002$), while the difference was not found on CG students ($Z = 1.107$, $p = .268$). These results indicate that the EG students’ confidence gained more improvement than the CG students’.

![Figure 5](image_url)

*Figure 5. The mean scores of students’ confidence toward mathematics*

A possible reason for such results is that games have specific goals, which provide the opportunity of attaining a winning state for students in the EG and then provide the EG students with a sense of success. Successfully completing a task enhanced students’ confidence toward the future mathematics learning tasks. This finding echoes the theory of mastery experience (Bandura, 1997), which argues that people who repeatedly gain the successful experience due to their effort can become confident when performing related tasks.

In contrast with the EG, the CG students could not realize their performance and progress immediately. In other words, the gradually increased success experience may be the key to help the EG students gain the improvement of confidence toward mathematics. However, other factors (e.g., different timing of feedback) that resulted in such a result are possible and need to be verified in our further works.

Computational performance

In this study, students’ mathematical achievement was determined by their computational abilities. More specifically, students’ mathematical achievement was defined as the amount of correct answers they produced during the assessment time. Figure 6 presents the result of students’ computational performance in the pretest and posttest. Students in both groups answered more questions correctly in the posttest than in the pretest. Both of the learning approaches seemed helpful to students. Importantly, the Mann-Whitney test indicated that the gain score in the EG was significantly greater than that in the CG ($U = 193.000$, $p = .013$). This result suggests that the game-based approach may help EG students produce more correct answers.
Comparison between EG and CG in terms of students’ mathematical ability

The previous sections demonstrated that the game-based approach may help students gain improvement on their confidence toward mathematics and mathematical performance. To answer the second research question, participants were further divided into four sub-groups in terms of their academic abilities. The results were reported in the following sections.

Students in the EG and CG were divided into high and low-ability groups according to their math performance of the general math test. The distribution of students is shown in Table 3. Students whose score is higher than the average score of their class were assigned to the high-ability group while students who performed lower than the average score were assigned to the low-ability group.

| Table 3. The numbers of students grouped by their math ability |
|------------------|------|------|
|                  | EG   | CG   |
| High ability     | 14   | 12   |
| Low ability      | 12   | 13   |

Confidence

Students’ confidence toward mathematics is illustrated in Figure 7. Low-ability students in the CG, in contrast to the other three sub-groups, demonstrated a different trend of their confidence change. More specifically, those low-ability CG students’ confidence decreased after the intervention.

Regarding the paper-based (CG) condition, the result of the Mann-Whitney test revealed a significant difference of the gain score of confidence between the high-ability and the low-ability students ($U = 40.500$, $p = .041$). The Wilcoxon’s matched-pairs signed-ranks tests revealed significant decrease of confidence between the pretest and the posttest for low-ability students ($Z = 2.398$, $p = .016$), while the difference was not found in the high-ability students ($Z = .817$, $p = .414$). This may be due to the fact that low-ability students usually have fewer chances to receive academic achievement than their high-ability peers. In addition, low-ability students have greater tendency to attribute their success to external reasons (e.g., luck), instead of internal reasons (e.g., effort), than high-ability students (Vlahović-Šterić, Vizek Vidović, & Arambašić, 1999; Weiner, 1980). Low-ability students may need more positive feedback on their performance to build their confidence than high-ability students because such feedback helps students grasp their progress and enhances students’ confidence toward the task that they are carrying out.
(Earley, Northcraft, Lee, & Lituchy, 1990; Schunk and Swartz, 1993). Students in the CG did not receive any immediate feedback about how well they had done. Thus, it may be the key issue that students did not gain significant improvement of confidence.

![Figure 7](image_url)

*Figure 7. The mean scores of confidence toward mathematics grouped by ability*

Regarding the digital game-based (EG) condition, both high and low-ability students demonstrated a similar trend of confidence change. The result of the Mann-Whitney test indicated no significant difference of the gain score of confidence between the two groups of students ($U = 81.000, p = .876$). At the same time, the Wilcoxon’s matched-pairs signed-ranks tests revealed significant differences of the confidence between the pretest and the posttest for high-ability ($Z = 2.165, p = .030$) and low-ability ($Z = 2.156, p = .031$) students. This may be due to the fact that the GBL environment provided more chances to receive positive feedback, which is useful for low-ability students. Students could rectify their errors by themselves with the feedback provided by the games. Frequently receiving the message of successfully completing learning tasks motivated low-ability EG students so that their confidence toward mathematics could be improved. This is coherent with past studies, which indicated that the feedback of successfully completing a learning task increases students’ confidence toward learning (Schunk, 1991, Pajares, 2006). On the other hand, the improvement of high-ability students’ confidence may result from completing challenging tasks since they may tend to expect challenging tasks (Li & Pan, 2009). Completing challenging tasks brings the sense of achievement to students (Dickey, 2007), which in turn raises students’ confidence (Hammond, 2004). Virtual opponents in the games simulated students’ ability to provide different levels of challenge for students with different levels of ability. Therefore, the games were still challenging for high-ability students, and this may be why the high-ability students’ confidence was improved.

In this study, both high-ability and low-ability students benefited from the positive effect of embedding learning materials into digital games, which provided students with adequate challenge, immediate feedback of their performance and opportunities so that they could gain the sense of achievement. This result indicates that GBL may help students with different levels of ability to improve confidence toward mathematics. On the one hand, concrete feedback about whether they were correct and where they made mistakes would help low-ability students consolidate their knowledge and strengthen their confidence (Kelley & McLaughlin, 2012). In this study, the immediate feedback of performance provided by the game-based environment may be the key element that supports low-ability students in learning mathematics.

On the other hand, those high-ability students in the GBL environment gained more improvement of confidence toward mathematics than their high-ability peers who gained confidence in the paper-based condition, though a significant degree is not reached. These results imply that GBL may not only be beneficial to low-ability students, but also be helpful to high-ability students, in terms of improving confidence, because digital games maintain keep providing appropriate challenge to students.
To sum up, the experience of successfully completing a task acts as a source for building low-ability students’ confidence and completing challenging tasks resulted in the improvement of high-ability students’ confidence. In addition, the findings reported in this section further indicates that the decrease of confidence toward mathematics in CG is contributed by low-ability students, whose confidence is significantly reduced in the posttest.

**Computational performance**

Students’ computational performance grouped by different levels of ability is shown in Figure 8. For the paper-based (CG) condition, the Wilcoxon’s matched-pairs signed-ranks tests indicated that both the high-ability \((Z = 3.061, p = .002)\) and the low-ability \((Z = 2.971, p = .003)\) students gained significant improvement on their computational performance. As expected, the course delivered in this study brought different levels of improvement depending on students’ abilities. The result of the Mann-Whitney test demonstrated a significant difference of the gain score of computational performance between the high-ability and the low-ability students \((U = 40.000, p = .039)\). This implied that the high-ability students in the CG gained more improvement than their low-ability peers in the same setting. This result is consistent with that of previous studies which indicated that high-ability students possessed greater learning capacity to learn new skills, such as various learning strategies (Yip, 2007) and better reasoning ability (Means & Voss, 1996). Thus, the high-ability students obtained more learning gain than low-ability students in this traditional setting.

![Figure 8. The mean scores of computational performance grouped by ability](image)

Importantly, although the students in the EG performed slightly worse than the CG students in the pretest, the high-ability students in both groups attained a comparable level of performance in the posttest. For a direct comparison of the effect between the EG and CG on low-ability students, the result of the Mann-Whitney tests demonstrated a significant difference of the gain score of math performance between the EG low-ability and the CG low-ability students \((U = 38.500, p = .032)\) while the difference of pretest between the EG low-ability and the CG low-ability students was not found \((U = 46.500, p = .085)\). This result indicated that the low-ability students in the EG gained more improvement than those in the CG.

The results reported in this section indicated that GBL may be a better approach for low-ability students, in terms of the learning gain. Although all students gained significant improvement, the improvement of the low-ability students in CG is most limited. Although the high-ability students in the EG performed a bit worse than their high-ability peers in the CG in the pretest, they achieved a comparable level of computational performance in the posttest. For low-ability students, students in the EG gained more improvement than their peers in the CG. The immediate feedback for error correction provided in the digital games may be the element that makes students reach a higher improvement. This is consistent with the results of previous studies (Brosvic, Epstein, Dihoff, & Cook, 2006;
McDaniel, Roediger, & McDermott, 2007), which indicated that immediate feedback helped learner obtain more learning gain and better retention of knowledge.

The aforementioned results suggest that the GBL approach seems to be a better approach than the paper-based approach, regardless of for low or high-ability learners. In contrast with the paper-based approach which weakened low-ability students’ confidence, GBL helps low-ability learners build their confidence toward mathematics in this study. In addition, the GBL approach also helped low-ability students gain more boost for their math skills than their low-ability peers who received paper-based intervention. As to the high-ability students, although students in both groups attained a comparable level of achievement, GBL also helped students gain significant improvement of confidence toward mathematics while those high-ability students learned with the paper-based setting did not gain significant improvement of confidence.

Conclusions

This study investigates whether GBL can enhance students’ confidence toward mathematics. The results indicate that both high-ability and low-ability students with the GBL approach gained significant improvement on their confidence toward mathematics. In contrast, students with the paper-based setting did not show a significant improvement on their confidence toward mathematics due to a significant decrease of confidence toward mathematics from low-ability students. Regarding performance, students in both conditions gained significant improvements in their performance; however, the students with the digital game-based setting gained more improvement than their peers with the paper-based setting. Furthermore, the results showed that high-ability students in both groups attained a comparable level of performance while the low-ability students in the game-based condition gained more improvement in their performance than those in the paper-based condition.

The contribution of this study can be summarized from three aspects. First, the power of mini-games: the simple mini-games used in this study helped students improve both their confidence and performance toward learning mathematics, especially for low-ability students. Second, through the bridging of the mini-games, students’ confidence toward mathematics and their calculation performance formed a bidirectional relationship and hence mutually enhanced with each other. Third, several characteristics of the mini-games which enhanced students’ confidence and performance were identified; these elements are specific goals, immediate feedback, and various levels of challenge. Specific goals provide students with a chance to obtain the sense of success which then enhance student’s confidence. Immediate feedback of students’ performance plays a supporting role, which lets students grasp their progress and directs them to move forward, especially for low-ability students. Various levels of challenge let students with diverse levels of ability enter the flow state.

This study has figured out the advantages of incorporating digital games into math learning. However, it has some limitations. First, this study was a small-scale study; a further study with a larger sample needs to be done to provide additional evidence. Second, this study used two mini-games as the learning environments; there is a need to identify the effects of the two mini-games in future study. In addition, there is also a need to conduct further research to examine how other human factors, such as cognitive styles and gender differences, affect learners’ reactions to digital games.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this study under contract numbers NSC 100-2511-S-008-014 and NSC 100-2511-S-008-013.

References


Klehanousi, I., & Williams, J. (2011). Students’ dispositions to study further mathematics in higher education: The effect of students’ mathematics self-efficacy. In M. Pylak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the seventh Congress of the European Society for Research in Mathematics Education* (pp. 1229-1238). Rzeszów, Poland: University of Rzeszów.


Personnel Psychology, 64(2), 489–528.


Developing Spatial Orientation and Spatial Memory with a Treasure Hunting Game

Chien-Heng Lin*, Chien-Min Chen2 and Yu-Chiung Lou3
1Department of Early Childhood Education, Asia University, Taiwan // 2Department of Foreign Languages and Literature, Asia University, Taiwan // 3Department of Social work, Asia University, Taiwan // chienlin@asia.edu.tw // cmchen@asia.edu.tw // yclou@asia.edu.tw

*Corresponding author

(Submitted June 3, 2013; Revised September 27, 2013; Accepted November 26, 2013)

ABSTRACT

The abilities of both spatial orientation and spatial memory play very important roles in human navigation and spatial cognition. Since such abilities are difficult to strengthen through books or classroom instruction, there are no particular curricula or methods to assist in their development. Therefore, this study develops a spatial treasure-hunting game to enhance learners’ spatial orientation and spatial memory. The game is designed on theoretical principles of spatial orientation and spatial memory. With such a design, the player can practice and learn these two spatial abilities by way of playing the game for a short period of time. This study adopts an experimental approach to test whether the game is effective in facilitating student’s development of spatial memory and spatial orientation. The findings show that this computer game is an efficient way of enhancing a learner’s spatial orientation and spatial memory over a very short period. Our research also points out that spatial orientation ability may differ between males and females, but after training through the game, the differences are reduced.

Keywords

Spatial orientation, Spatial memory, Computer game, Spatial ability

Introduction

The environment in which we live is filled with buildings, roads, and public facilities, and areas outdoors consist of a mix of various terrains, flowered landscapes, bushes, and woods. Given all of this contextual variation, for a human to navigate in such a complex environment without losing direction, a sense of personal orientation and human spatial ability play critical roles (Hung, Hwang, Lee, & Su, 2012). Spatial ability is accepted as a significant attribute in humans that is necessary to evaluate the effectiveness of learning, training, working, and playing (Rafi, Anuar, Samad, Hayati, & Mahadzir, 2005). Spatial ability is also regarded as a sort of intelligence by many researchers and identified as a cognitive process that functions when dealing with problems related to the manipulation of spatial information (Yilmaz, 2009).

Looking at our primary focus, the application of spatial ability to the daily lives of human, there are two components of spatial ability that are highly related to finding place and searching for direction: spatial orientation and spatial memory. Spatial orientation helps us to address “where is it?” and “where am I?” (Allen, 2003). Spatial memory is used to recall all spatial information related to the process of location identification and place searching, which helps the searcher orient her/his own position and the location of the searched-for object (Pentland, Anderson, Dye, & Wood, 2003). These two components of spatial competence play essential roles in location finding.

Although the two aspects of spatial ability are very important for humans in direction orientation and space location, there are no clearly defined methods of teaching available to assist students in enhancing these capabilities. Since such abilities are rather difficult to improve through books or classroom instruction, the general attitude toward this is to let them develop in their own natural way (Gittler & Gluck, 1998). Presently, there are no particular courses or methods to assist development.

A noteworthy issue, then, is how we can help students develop these spatial abilities as efficiently as possible. Due to the rapid developments in digital technology, we can adopt special features of multimedia to foster these two aspects of spatial ability by playing computer games. Computer games have proven to be very powerful tools in this regard since many games are available that involve abilities of this type (Sims & Mayer, 2002). Evidence from experimental studies has shown that spatial ability can be cultivated through proper training with digital aids (Kadam,
Sahasrabudhe, & Iyer, 2012; Subrahmanyam & Greenfield, 1994). For example, Yang and Chen (2010) designed a digital pentominoes game that tries to improve students’ spatial abilities, and the result showed that their abilities were enhanced. David (2012) revealed that people’s spatial abilities could be improved through practice with selected computer games. Do and Lee (2009) also demonstrated that a 3D computer game can be useful in improving people’s spatial abilities. Furthermore, Samsudin, Rafi, and Hanif (2011) showed that mental rotation and spatial visualization can be improved through technology-based training instruments. Finally, Wilms, Petersen, and Vangkilde (2013) claimed that intensive video action gaming can improve the encoding speed of visual information into visual memory.

**Gender differences**

Past research has claimed that there are marked differences between males and females in their spatial abilities. In tests on diverse types of spatial abilities, males perform better than females (Coluccia & Louse, 2004). However, other studies have shown that no significant differences are found between males and females in spatial visualization and spatial perception. Therefore, in tests of diverse types of spatial abilities different results might occur related to gender differences. On the capacity of spatial orientation, the majority of studies indicate that males perform better than females. Based on inductive statistics by Coluccia and Louse (2004), either in the real or simulated research environments, boys are substantially better than girls in the performance of spatial orientation. Recent studies have shown that males are also faster at locating targets in virtual environments (Lin et al., 2012).

Regarding gender differences in spatial memory, the research findings are mixed. According to Canovas, Garcia and Cimadevilla’s (2011) research findings, males outperform females in remembering the position of objects in a space. Piccardi et al. (2008) adopted the Corsi Block test to examine people's spatial memories, and the findings showed that boys perform better than girls. However, other researchers (Bull, Davidson, & Nordmann, 2010; Rahman, Bakare, & Serinsu, 2011) proposed that no gender differences exist in the ability of spatial memory. Voyer, Postma, Brake and McGinley (2007) even pointed out that females outperform males on most tests of object location memory.

Although males and females may differ in their spatial abilities, the differences can be reduced via certain learning strategies that fit males and females well (Coluccia, Louise, & Brandimonte, 2007). As Yang and Chen (2010) have shown, gender differences in spatial abilities can be reduced with the support of a digital game.

**Game design for certain instructional purpose**

Whether though brick and mortar retail markets or via the internet various types of computer games are available; however, the number of games specifically designed for certain cognitive abilities are limited (Sung, Chang, & Lee, 2008). Ideal games for such purposes should be interesting, fun, and attractive and, at the same time, meet the teaching goal prescribed by the designer.

Researchers have argued that the design of computer games has been driven by technological advances rather than theoretical principles (Hegarty, Quilici, Narayanan, Holmquist, & Moreno, 1999). In other words, the first step in creating an effective game is to analyze the substance and definition of the prescribed cognitive ability and then explore a strategy that provides an ideal process to enhance that cognitive ability.

**Research purpose**

Because there are very few computer games based on cognitive theory and because of the specific need to design domain-related multimedia learning environments, this study sets out to demonstrate how to design a computer game based on theory and domain-specific concepts. Furthermore, there is limited research on the improvement of spatial orientation and spatial memory through the design of computer games. As such, this research aims at designing a computer game based on theories related to the development of spatial ability through which students can enhance their spatial orientation and spatial memory. The second purpose of this study is to evaluate the effects of the designed game on facilitating students’ development of spatial orientation and spatial memory within the limited duration of implementation.
The theoretical basis of the design

Landmark information for spatial orientation

One relevant facet of spatial competence is spatial orientation, which is the ability of people to orientate themselves in new environments. Spatial orientation indicates an essential cognitive function through which most humans find their direction and location (Wolbers & Hegarty, 2010).

Landmark information seems to play an important role in orientation. According to Presson’s (1982) research, when identifying the precise location of a target without knowing its position beforehand, 6-8 year old children depend mainly on distance, far or near, between a landmark and the target. If the landmark is directly connected to the target, we call the landmark a “proximal cue”. On the other hand, a landmark that provides only indirect or limited information about the distance to the goal is called a “distal cue.” In other words, when searching for a position that is out of sight, people tend to use an obvious landmark or a familiar place to orient and clarify what they want to identify. Whether the target is far or near, the landmark can be used as a reference.

Spatial memory

Another relevant factor concerning spatial competence is spatial memory, which is the cognitive process that enables people to remember locations and the relationship between locations and objects in the environment. In other words, the ability to recognize and understand spatial relationships (Kozhevnikov & Hegarty, 2001). The cognitive process of spatial memory enables people to remember relationships among events and objects involved in the environment. The formation of spatial memory depends on how such sensory information is collected and applied (Johnson & Villani, 2010).

People can easily remember a location mainly because at that place, compared to other places, they made certain significant decisions (Doherty, Gale, Pellegrino, & Golledge, 1989). As to the factor, Herman & Roth (1984) discovered that if certain situations or stories can be interwoven with the location for which they have searched, people can better remember such spatial locations.

The design of the game

The strategy of the game design

The game is designed and constructed using the strategy of a treasure hunt. The framework of the game design is based on a simple idea that aims to lead the main character to seek treasures in the designed scenes and marked locations (Fig. 1). In this treasure hunt, the leading character is moved around by the player with a treasure map as the spatial guidance to find hidden treasures. Since the game of treasure hunt features the seeking of location and target, it is functional in enhancing the ability of spatial orientation.

Traditionally, the game of treasure hunt keeps a small-scale radar map on the screen. The player looks at this map to determine the direction and location of the hidden treasures. Nevertheless, the traditional game design aims at entertaining; thus, simply by watching the radar map and seeking the target, the player does not receive much direct stimulation that enhances spatial ability.

Therefore, the game design should include a function whereby the player can go through changes of location and direction and eventually find the target destination (i.e., the spatial orientation), and the player also has opportunities to practice instant memory of locations and directions of the objects or events involved (i.e., the spatial memory).
The learning of spatial memory

In the game’s design, before entering each level of the game, there will be a map providing an overall view of the scene with an arrow indicating, and speech sounds introducing, all possible treasure landmarks (Fig. 2). The player needs to quickly memorize the locations and the relative relationships and directions among all the landmarks. Upon entering the game, the view of the scene changes from an overall view to a selective partial one, which is similar to the limited viewpoint of people in a real environment. Then words and speech indicate the first treasure location, and the player uses the direction control key to move the character to seek the treasure (Fig. 3).

Given this, the player has to operate by spatial memory and quickly memorize the related landmarks and directions on the map. With advanced degrees of difficulty there are more landmark locations to be memorized.
The learning of spatial orientation

As mentioned above, spatial orientation ability means that one can identify his or her own location and direction within different and changing environments. Accordingly, the game design must contain two key elements to complete the learning of spatial orientation: Change of initial starting point and change of scenes (i.e., rotation of scenes).

In the above-mentioned order, the player memorizes the overall view of the map and enters at a random landmark as the starting point to seek the treasure. There are always six passes on the same level. The first three passes are the treasure seeker’s various initial positions (to change the starting point), and the next three passes are 90°, 180°, and 270° rotations of the scenes, as shown in Fig. 4. The player has to identify his/her location at this position, which is viewed as the base, and then go forward to a landmark and seek the treasure. The higher the level of difficulty, the more landmarks there will be.
Principles for the design of scenes

With regard to game design, in order to help the player memorize the various locations and seek the hidden treasures, there are three principles to consider: The referent landmarks, the decisive point, and the story context, which was previously discussed. These three principles assist the designer in developing a game situation that enhances the player’s spatial orientation and spatial memory.

By using these three principles as a foundation, the designer develops the treasure hunt game. Thus, this game design uses conspicuous landmarks to indicate every location of a hidden treasure, and the newly found treasure location serves as a new starting point; thus, the player goes on to seek the next landmark. At each level there is an introduction to the new theme as shown in the background of the plot.

Method

Participants

This study was conducted with 55 sixth grade students (aged 12-13 years, $M_{\text{age}}=12.4$) from an elementary school in Taiwan, who volunteered to participate. Due to differences in spatial abilities between males and females and the possibility of comparison between them, the numbers of males and females were controlled to provide a balanced sample: twenty-six male students and twenty-nine female students. The participants were randomly assigned to one of two groups: The treatment group had twenty-eight students, and the control group had twenty-seven students.

Instruments

Spatial orientation

In order to examine the participants’ spatial orientation ability, two instruments were used in this study. The first was the “Guilford-Zimmerman Spatial Orientation Test” (GZSOT) which is a traditional and the most commonly used instrument for testing spatial orientation ability (Guilford & Zimmerman, 1948). In this test, participants were shown the same landscape from two different view-points, both from the prow of the same boat so they could determine how much the position of the boat had changed in the second picture as compared with the original position in the first picture (Fig. 5). However, due to the complexity of the demonstration and application, this instrument often cannot effectively measure the ability. To improve on this weakness, Kyritsis & Gulliver (2009) created an electronic version of the GZSOT (internal reliability 0.88). Therefore, in order to avoid the possible weakness, this research adopted the newer electronic version (Fig. 5).

Figure 5. Traditional and electronic versions of Guilford-Zimmerman Spatial Orientation Test
Nevertheless, the first instrument has been criticized for one of its weak points; that is, any rotation of more than 90 degrees cannot be effectively measured (Kozhevnikov & Hegarty, 2001). In order to solve this problem, a second instrument was adopted; the “Perspective Taking Spatial Orientation Test” (PTSOT), which was developed by Hegarty & Waller (2004). This is a test of one’s ability to imagine different perspectives and orientations in space. This test can measure the perspective involving a rotation of over 90 degrees (internal reliability 0.83), as shown in Fig. 6.

![Figure 6. Perspective Taking Spatial Orientation Test](image)

Spatial memory

The Corsi Block-Tapping Task (CBTT), originally developed by Corsi (1972), was adopted to examine spatial memory. This test has frequently been used to assess short-term visuospatial memory performance in adults and children (Vandierendonck, Kemps, Fastame, & Szmalec, 2004). Fig. 7 shows the Corsi Blocks-Tapping Task.

![Figure 7. Corsi Block-Tapping Task](image)

The second instrument for examining spatial memory was a “Game-Based Spatial Memory Test” (GBSMT). This test examines how well participants improve their ability to memorize and identify their position and direction. This test was designed based on the scenes of the game and includes eight scenes and twenty questions. Each scene has
six target objects. The object of the test is to memorize the locations in the scene within a limited period of time and then indicate the position and direction of the object requested by the question, as shown in Fig. 8.

![Game-Based Spatial Memory Test](image)

**Figure 8. Game-Based Spatial Memory Test**

**Procedure**

This research adopts a pre-test and post-test design to measure the game’s effect on the students. In addition to the game’s effect on the experimental team, the game’s effect on the control team was measured. This was designed to reduce any possible inaccuracy that may have been caused by the pre-test and the post-test.

In each performance, four or five participants formed a team. Each participant operated a computer independently and maintained a certain distance from the other subjects to avoid any possible disturbance. A description of the entire process follows.

All participants were pre-tested with the GZSOT test (10min), the PTSOT (5min), the CBTT test (5min), and the GBSMT test (5min). After a five-minutes break, the game began. Once they entered the game, all participants started from the first pass and moved on to solve each pass until all were solved (about 45-55 minutes). The participants were required to wear headsets and were not allowed to have conversations with others during their participation. After the experiment was completed, they had a 10-minutes break and then the post-tests using the same instruments were conducted.

**Data analyses**

All the data collected from the pre-tests and post-tests of the GZSOT, PTSOT, CBTT, and GBSMT were coded for quantitative analyses. The GZSOT is comprised of fifty questions and for each correct answer the participant earns one point, but for a wrong answer 0.25 point is deducted, in case the participant is randomly guessing answers. The PTSOT has twelve questions and for each correct answer the participant earns one point. The CBTT is based on the participant’s ability, which is indicated by how far the participant can go in solving the passes; then the final result will be exposed, which includes the participant’s grade and total points. The GBSMT consists of twenty questions and each correct answer earns five points, with one hundred points being the maximum possible. The pre-test and post-test scores of the four exams were analyzed with descriptive statistics, including mean and standard deviation.

In order to check whether the differences between pre-test and post-test scores were significant, and to exclude certain moderating variables this research adopted the method of analysis of covariance (ANCOVA). Paired t-tests and an independent sample t-test were conducted for data analyses.
Results

Spatial orientation test

In order to examine differences in spatial orientation between the treatment and control groups, analyses were undertaken between the pre-test and post-test scores for GZSOT and PTSOT. The means and standard deviations of the participants’ pre-test and post-test scores for the PTSOT are shown in Table 1.

Analysis of covariance can exclude the interference effect from items of covariance. However, before the application of ANCOVA, we proceeded with the test of “Group regression coefficients homogeneity,” which was non-significant (data of a*x): \( F = 0.188; p = 0.666 \). This shows that the slope of the regression line in the two groups remains the same. Thus, the relationship between the covariates (pre-test scores) and the dependent variables (post-test scores) does not vary with each different handling standard of the independent variables, and this is in accordance with the homogeneity hypothesis of the group regression coefficients of covariance. Thus, the analysis of covariance can be further conducted.

The result of the ANCOVA was significant (\( F = 55.41, p = 0.00 \)), which indicated there were significant differences between the treatment group and the control group in their post-test scores under PTSOT. The adjusted post-test average scores are shown in Table 1, with treatment group significantly greater than the control group. That is, after receiving the spatial game exercises, the participants in the treatment groups made significant progress in their post-test scores under the Perspective Taking Spatial Orientation Test (PTSOT).

Table 1. Scores of pre- and post-tests in the Perspective Taking Spatial Orientation Test (PTSOT)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean (SD)</th>
<th>Adjusted Mean</th>
<th>ANCOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>28</td>
<td>7.25 (9.28)</td>
<td></td>
<td>55.415</td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>7.88 (1.31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>28</td>
<td>9.79 (1.32)</td>
<td>9.99 (a)</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>8.22 (1.05)</td>
<td>8.01 (a)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows the means and standard deviations of the student’s performance on the GZSOT in the pre-test and post-test. First, we proceeded with the test of “homogeneity of the in-group regression coefficients,” and the result was non-significant (\( F = 0.283, p = 0.597 \)). Thus, the analysis of covariance can be applied.

The subsequent ANCOVA was significant (\( F = 44.61, p = 0.00 \)), indicating that there are significant differences between the treatment group and the control group in their GZSOT post-test scores.

Table 2. Scores of pre- and post-tests in the Guilford-Zimmerman Spatial Orientation Test (GZSOT)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean (SD)</th>
<th>Adjusted Mean</th>
<th>ANCOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>28</td>
<td>18.21 (7.98)</td>
<td></td>
<td>44.61</td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>18.00 (7.40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>28</td>
<td>22.37 (8.51)</td>
<td>22.27 (a)</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>18.35 (7.18)</td>
<td>18.45 (a)</td>
<td></td>
</tr>
</tbody>
</table>

Spatial memory test

We used the Corsi Block-Tapping Task (CBTT) to examine the participant’s spatial memory. The scores from the CBTT are shown in Table 3. We conducted an ANCOVA with a single factor to test pre-test and post-test differences between the treatment and the control groups.

The test of “homogeneity of the in-group regression coefficients” was not significant (\( F = 0.139, p = 0.711 \)). This is lower than the significant level and the null hypothesis is accepted. This means that the analysis of covariance can be applied. The result of the subsequent ANCOVA was significant (\( F = 6.578, p = 0.013 \)), indicating that there are significant differences between the treatment group and the control group in their CBTT scores.

87
Table 3. Scores of pre- and post-tests in the Corsi Block-Tapping Task (CBTT)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean(SD)</th>
<th>Adjusted Mean</th>
<th>ANCOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F</td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>28</td>
<td>56.46 (19.21)</td>
<td></td>
<td>22.66</td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>53.52 (16.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>28</td>
<td>76.00 (22.24)</td>
<td>74.67 (a)</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>59.11 (16.64)</td>
<td>60.50 (a)</td>
<td></td>
</tr>
</tbody>
</table>

The scores for the Game-Based Special Memory Test (GBSMT) are shown in Table 4. The results showed that the difference between the treatment and the control groups in the test was significant ($F = 5.20, p = 0.027$). The adjusted average values of the post-test are shown in Table 4, and it can be seen that the treatment group is significantly greater than the control group in the adjusted average value of the post-test game test.

Table 4. Scores of pre- and post-tests in the Game-Based Spatial Memory Test (GBSMT)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean (SD)</th>
<th>Adjusted Mean</th>
<th>ANCOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F</td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>28</td>
<td>77.5 (14.37)</td>
<td></td>
<td>5.20</td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>78.15 (14.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>28</td>
<td>87.14 (10.49)</td>
<td>87.33 (a)</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>27</td>
<td>83.15 (11.70)</td>
<td>82.95 (a)</td>
<td></td>
</tr>
</tbody>
</table>

Gender differences

In order to examine gender differences in spatial memory and spatial orientation abilities, analyses were undertaken between boys and girls in the pre-tests and post-tests. Table 5 shows the results of the analyses of the PTSOT, which indicates there is a significant difference between boys and girls in the pre-test ($t = 4.829, p < 0.05$). Boys obviously performed better than girls on this test. Although there is a significant difference between boys and girls in the post-test of PTSOT ($t = 2.07, p = 0.048$); however, after the exercises in the spatial game, the difference between the boys' and girls' performance on the PTSOT is reduced.

Table 5. Performance of boys and girls in the pretests and posttests on the PTSOT

<table>
<thead>
<tr>
<th>Test</th>
<th>Sex</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>$t$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>M</td>
<td>13</td>
<td>7.9231</td>
<td>.75955</td>
<td>4.829</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15</td>
<td>6.6667</td>
<td>.61721</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td>M</td>
<td>13</td>
<td>10.3077</td>
<td>.85485</td>
<td>2.070</td>
<td>.048</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15</td>
<td>9.3333</td>
<td>1.49603</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of another spatial orientation test, the GZSOT, are shown in Table 6. These results indicate a significant difference between boys and girls in the pre-test and post-test performance in that boys perform better than girls ($t = 2.07, p < 0.05$).

Table 6. Performance of boys and girls in the pretests and posttests on the GZSOT

<table>
<thead>
<tr>
<th>Test</th>
<th>Sex</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>$t$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>M</td>
<td>13</td>
<td>23.2115</td>
<td>8.95698</td>
<td>3.759</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15</td>
<td>13.8833</td>
<td>3.30160</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td>M</td>
<td>13</td>
<td>27.0962</td>
<td>9.14612</td>
<td>3.160</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15</td>
<td>18.2667</td>
<td>5.41284</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The test of spatial memory in the Corsi Block-Tapping Task (CBTT) provided the results shown in Table 7, which indicate there was no significant difference between boys’ and girls’ performance on the CBTT, in either the pre-test or in the post-test.
Table 7. Performance of boys and girls in the pretests and posttests on the CBTT

<table>
<thead>
<tr>
<th>Test</th>
<th>Sex</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>M</td>
<td>13</td>
<td>59.6154</td>
<td>23.0164</td>
<td>.803</td>
<td>.429</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15</td>
<td>53.7333</td>
<td>15.4986</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td>M</td>
<td>13</td>
<td>78.9231</td>
<td>23.9007</td>
<td>.640</td>
<td>.528</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15</td>
<td>73.4667</td>
<td>21.1992</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In line with the results of the CBTT, the results for the GBSMT shown in Table 8 illustrate there was no significant difference between boys’ and girls’ performance on the game test either in the pre-test or post-test, although boys’ mean scores are slightly higher than girls’.

Table 8. Performance of boys and girls in the pretests and posttests on the GBSMT

<table>
<thead>
<tr>
<th>Test</th>
<th>Sex</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>M</td>
<td>13</td>
<td>81.1538</td>
<td>11.0215</td>
<td>1.267</td>
<td>.217</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15</td>
<td>74.3333</td>
<td>16.4606</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td>M</td>
<td>13</td>
<td>90.7692</td>
<td>5.3409</td>
<td>1.769</td>
<td>.089</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>15</td>
<td>84.0000</td>
<td>12.8452</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

Many studies have shown that multimedia computer games can improve the performance of a variety of human capabilities and intelligence (Pannese & Carlesi, 2007; Sung et al., 2008; Wilms et al., 2013). In terms of ability for spatial intelligence, many researchers have also pointed out that some of the relevant games can instantly enhance human spatial ability (David, 2012; Samsudin et al., 2011).

The authors of this study designed a treasure hunting game to train and improve students’ spatial ability, focusing particularly on spatial orientation and spatial memory. The findings show that the designed game has effectively improved the participants’ abilities in spatial orientation and spatial memory.

In two kinds of spatial orientation tests, either the small-angle test or the large-angle test, both PTSOT and GZSOT indicate that participants’ abilities in spatial orientation improved significantly after their exposure to such a designed game. This result is consistent with previous researchers’ findings that spatial capabilities can be effectively improved through the use of computer games (Yang & Chen, 2010; You, Chung, & Chen, 2008). A possible explanation for this positive result lies in the fact that the game is designed according to the theoretic principles of spatial ability, which can directly stimulate both spatial abilities. For instance, in the game, the character is constantly changing her/his starting position, and the view of the setting is constantly rotating to different angles. With the changes in location and environment, the player must constantly locate her/his position and find the relevant landmarks in order to identify the target treasures.

From the results of both spatial memory tests we found that this computer game has a positive effect on the development of spatial memory. Both spatial memory tests showed that, after about one hour practicing the game, participants greatly improved their spatial memory. The main reason for this is the involvement of spatial memory in the game. In this game’s design, every detail of the process requires the game player to use her/his spatial memory, from the simple pass of three landmarks to the complex pass of twelve landmarks, the player has to memorize within a very short period of time the spatial locations of these landmarks. Thus, this research demonstrates that a theory-based computer game can be very helpful in improving students’ spatial memory after a short period of practice, as compared with the students in the control group.

By playing the game, students practice many kinds of spatial skills, such as orientating location and direction on the map, determining the correct landmark, and memorizing related locations. This result echoes previous researchers’ findings that spatial ability can be improved through appropriate training with computer-based games (Yang & Chen, 2010).
This study also examined gender differences in spatial orientation and spatial memory. First, in the performance of spatial orientation ability, the pre-test results show that the abilities of boys are greater than those of girls. This result echoes the findings of previous studies in which males were found to perform better than females in spatial ability (Monahan, Harke, & Shelly, 2008; Yang & Chen, 2010).

After the boys and girls play the game, the boys’ post-test results were still better than the girls’; however, the difference was greatly reduced. That is to say, through the training with the game, the gap between males and females on spatial ability can be narrowed. This result is consistent with the findings of Terlecki, Newcombe and Little (2008) and Yang, Chen (2010) who suggest that, although boys comparatively have better spatial intelligence, there is almost no gap in performance between girls and boys after training with the game. The fact that there is a significant difference between boys and girls in spatial ability is likely the result of traditional social attitudes and socio-cultural factors. As observed by Voyer, Nolan, and Voyer (2000), most male-typical activities involve high spatial content, while female-typical activities do not. Therefore, less experience with spatial activities naturally affects the development of spatial ability (Saucier, McCreamy, & Saxberg, 2002). Traditionally, due to fewer opportunities for girls to practice, their spatial abilities are weaker, but whenever they have the appropriate opportunity to practice, the gap between males and females will decrease.

As to spatial memory, the boys’ pretest and posttest scores are, on average, higher than girls’ but this feature is not significant. In other words, there is no significant difference in the spatial memory between males and females. The findings of this study are consistent with the work of Maki, Yoshida and Yamaguchi (2010), who concluded that there were no significant gender differences in the performance of spatial memory. The main possible explanation for this lies in the fact that spatial memory is more like a memory capacity and, in related studies, no differences between males and females were found. Also, some studies have shown that the spatial ability construct is made of many facets, such as spatial perception, spatial visualization, mental rotation, and so on (Linn & Petersen, 1985). Nevertheless, some of those facets have demonstrated differences between males and females, but others do not. For example, the findings of Yang and Chen (2010) show that there is no difference between the male and the female performance in spatial perception and spatial visualization.

Conclusions

Previous studies have effectively increased spatial ability through the use of digital games (David, 2012; Do & Lee, 2009; Sims & Mayer, 2002). Yang and Chen’s (2010) studies further proved that the digital pentominoes game can enhance three kinds of spatial abilities: spatial perception, spatial visualization, and mental rotation. However, for spatial orientation and memory, the two direction-identification abilities, only limited studies have used digital games to improve them and even more rarely using treasure hunting games. Therefore, based on the theoretical foundation of spatial orientation and memory, this study developed a digital treasure hunting game; thus, the effects of two essential spatial factors and gender differences were examined within the game.

Our findings demonstrate that a digital spatial game is an efficient way of enhancing learners’ abilities in spatial orientation and memory within a short span of time. These findings echo related prior works in that digital games can be very effective in improving students’ spatial ability as discussed above. This study provides further evidence that, although males generally outperform females in spatial orientation, the training provided in this game can effectively reduce the differences. This result is consistent with the findings of Terlecki, Newcombe and Little (2008). In addition, the study has shown that incorporating theory into the design of digital games is functional in enhancing the learner’s acquisition of cognitive concepts, which echoes work by Sung, Chang and Lee (2008).

However, a limitation of this study is the lack of exploring the long-term influence of the game, such as the reduction of gender differences over time. Further studies can focus on discovering the long-term effects of the digital games.

This study provides more evidence to demonstrate that digital games can be designed to not only offer entertainment functions, but also contribute to educational functions. Echoing the concept of "edutainment" proposed by Druin and Solomon (1996), the focus of the issue here is the proportional use of computer games.
References


Utilizing a Low-Cost, Laser-Driven Interactive System (LaDIS) to Improve Learning in Developing Rural Regions

Wei-Kai Liou¹ and Chun-Yen Chang¹,²*

¹Science Education Center, National Taiwan Normal University, Taiwan, ROC // ²Department of Earth Sciences and Graduate Institute of Science Education, National Taiwan Normal University, Taiwan, ROC //

liouweik@ntnu.edu.tw // changcy@ntnu.edu.tw

*Corresponding author

(Submitted March 29, 2013; Revised September 20, 2013; Accepted October 19, 2013)

ABSTRACT

This study proposes an innovation Laser-Driven Interactive System (LaDIS), utilizing general IWBs (Interactive Whiteboard) didactics, to support student learning for rural and developing regions. LaDIS is a system made to support traditional classroom practices between an instructor and a group of students. This invention effectively transforms a general projection screen into an IWBs. Accompanied by a projector, laser emitter, screen, PC (Personal Computer) and web cam, LaDIS provides a viable, efficient and economical solution for e-learning in classrooms. Students in the classroom are connected to a single computer and a shared screen. LaDIS is an extensible device, for use in any IWBs and interactive teaching software. This study engages 92 first-grade students enrolled in a 5-Year Junior College; these individuals are from two classes at a college of Taoyuan County, in Taiwan. Each class consists of 46 students. Classes are twelve hours in length, for a total of six weeks in the experimental course with different didactics. One class is the control group, and it adopts ordinary didactics. The other one is an experimental group, which adopts the innovation LaDIS didactics in class. The results indicate positive impacts of the innovation LaDIS on student learning performances. Most students are highly satisfied with the system and willing to use LaDIS because of its naturally superior interactive performance in classes. LaDIS provides a low-cost classroom tool, for developing regions financially unable to afford general IWBs equipment and didactics.

Keywords

Learning, Interactive Whiteboard, Laser emitter, Web cam, Developing and rural regions, Low-Cost, Economical

Introduction

Interactive whiteboards (IWBs) are large touch-sensitive boards, which control a computer connected to a digital projector; they were originally developed for office settings. Lee & Boyle (2003) recently described the adoption of IWBs as virtually an overnight occurrence in some schools; and, that the introduction of IWBs into schools transformed the overall use of ICT (Information and Communications Technology). Smith (2005) reports, the benefits of using a graphics package to support younger pupils' handwriting skills, where gross motor movements on the IWBs helped students' handwriting on paper. Similarly, younger pupils in Goodison’s study (2002a) report a preference for using the IWBs as opposed to a computer, because they found the keyboard and mouse difficult to manipulate. As one teacher described (Austin, 2003, p. 2), ‘I can see much more evidence of learning carried from one lesson to the next because of the ability for reinforcement on the fly’. Likewise, Walker (2002b, p.2) reports on a primary teacher, who finds the ability to ‘flip back and review material’, particularly beneficial for lower ability groups and pupils with special needs. Thus, it can be seen that the versatility of IWBs extends beyond the content of lessons and activities. In traditional teaching, the teacher talks more and there is less interactive teaching (Northcote et al., 2010). There is evidence to suggest that IWBs encourage teacher-led discussions in the form of dialogic teaching (Northcote et al., 2010). Reports also show that interactive teaching, supported by the use of IWBs, can lead to positive learning outcomes (Mercer et al., 2010). Teachers report finding IWBs a flexible and versatile teaching tool across age groups and settings (e.g., Austin, 2003; Jamerson, 2002), ranging from nursery (Wood, 2001; Lee & Boyle, 2003) to higher education (Malavet, 1998; Ekhaml, 2002) and even distance education (Abrams & Haefner, 1998; Bell, 2002). Furthermore, teachers, report that IWBs extend possibilities when catering for a range of needs within a lesson. The facility to flip back and forth between pages on IWBs screen is reported as a useful technique in supporting a range of flexibly and spontaneous needs within a class (Latham, 2002; Walker, 2002b). Stallard describes the introduction of IWBs, in 29 nurseries across Birmingham, as having a profound effect on the number of pupils choosing ICT activities (Wood, 2001). She found pupils that would not normally choose to work on the computer were choosing to work on the IWBs, and observed that they could do the activities without needing the fine-motor skills required to operate a mouse. More recently, IWBs have been described as a technology hub (Miller
& Glover, 2010; Thomas & Schmid, 2010) where the teacher has access to videos, the web, animations, etc. through one device, enabling the teacher to offer a rich learning experience to students. When the IWB is set up as a technology hub, it has been proposed (Miller & Glover, 2010) that it can act as a trigger for pedagogic change (Miller & Glover, 2010). Nowadays, the traditional white board is being replaced by IWBs, which combine video/media systems. On-line IWBs use interactive vector based graphical websites to share annotations and drawings. Many schools use IWBs to connect to the Internet/district networks to develop digital video distribution systems for on-line/off-line study that can provide distance learning or e-learning courses. Although IWBs provide valuable assistance for teaching needs, its price is a major concern for many users; a common IWB, with basic functions, costs upwards of 1,500 US dollars. Therefore, how to make an economical solution for IWBs is an attractive topic for many researchers. A noteworthy solution is the Wiimote Whiteboard, by Johnny Chung Lee (Lee, 2008). However, Wiimote is geared toward the main engine in the Wii video game controller, and it uses Bluetooth wireless transmissions. Consequently, Wiimote has function limitations because it is designed specifically for use with a video game controller and not for IWBs. Rather than altering the Wiimote sensor for a secondary use, in this paper, we develop a new installation and calibration method, utilizing tools regularly available in a standard classroom, such as projector, screen and pointer. Then, we propose a more viable, efficient and economical solution, specifically designed for IWBs.

Innovation laser-driven interactive system (LaDIS) design

LaDIS is a system made to support classrooms, unable to afford traditional IWBs equipment, but in need of a fully functional technology hub. LaDIS effectively and economical, transforms a general projector and a portable or stationary screen into an IWB. Via this invention, the classroom is connected to a single computer and a shared portable or stationary screen. LaDIS is an extensible device, and may be utilized in conjunction with any IWBs interactive teaching software. The system provides alternatives of both portable and stationary screen settings to satisfy various classroom needs. Basically, the laser spot, emitted by a laser emission unit and wireless control unit, allows the laser spot to synchronize and control the cursor and software on a projection screen. The concept behind the LaDIS design is shown in Figure 1(a). The illustration carefully indicates every component in the system but in real life, LaDIS, is not as complex as the classroom illustration shown in Figure 1(a). In reality, as shown in Figure 1(b), LaDIS is a very simple system.

![Figure 1(a). The concept of LaDIS design](http://youtu.be/cbE284_Bk7k)
Figure 1(b). Set-up of a real system: (1) the webcam can be fixed on a projector (2) the projector usually hangs from the ceiling of classroom (3) the computer usually is set up inside a lectern (4) wireless control technology can connect with LaDIS and a computer without any connecting lines.

The variety of interactive functions provided by LaDIS duplicate the functions of traditional IWBs. The results from this study contribute both a novel system for synchronous distance education in an affordable manner and design insights for creators of related systems.

Image acquisition

The Laser emission unit is for emitting a laser spot on screen, and it coordinates the laser spot detected by a data processing unit through the image capture unit. The imaging is utilized to display the output image from the data processing unit. The key point of image acquisition technology, in this study, is the red visible-light filter. A light filter lens may effectively filter the Blue and Green visible light on the screen image. The wavelength of red laser spot is higher than any visible red color. This allows the optics photography-function to trace the laser emitter spot on the screen image. Utilizing this spot tracking can provide coordinates of laser spot precise.

The internal coordinates with external coordinates converter operation

The laser emission unit is for emitting a laser spot imaging on a screen. Based on the coordinates of the laser spot, (detected through the image capture unit and the coordinate converter operation) the internal coordinates may be turned into external coordinates for the data processing unit. The location of the laser spot, through image acquisition calibration, as well as an internal and external coordinates converter operation, can guide the cursor output by the data processing unit, via on screen imaging to launch synchronous movement with laser spot as shown in Figure 2.

Figure 2. Via image acquisition and calibration, the external coordinate converter operation can guide the cursor to launch synchronous movement with laser spot.
Calibration technology

The calibration utilizes an automatic four corner localization method to adjust the coordinates of the mouse cursor on the screen. This automatic calibration only takes a few seconds each time. The coordinate of the mouse cursor is followed by a defined sequence for completing the calibration. When corner is captured by the camera, the tracking process can be started; at this time, the cursor tracking is set up completely. Once the calibration stage is completed, the system starts its main loop. The main loop consists of a series of image manipulations that are performed on each frame and result in an array of coordinates pertaining to laser points on the screen. After image processing and computer software program operations, the movement of an image can be transformed by a cursor command to guide the laser point precisely as shown in Figure 3.

![Figure 3](image)

*Figure 3. (A) A red visible-light filter lens transforms the whole image (and the background) into a full red view (B) The four white corners are captured by the camera (C) The calibration comes from an automatic four corner localization method (D) Utilizing a four corner localization method, the control cursor can launch precise and synchronous movement with the laser spot* 

Wireless control technology

LaDIS also installs a wireless control electron circuit which can directly execute each kind of interactive control function completely by wireless control technology. We explored laser guided and wireless control multi-technology for their interactions and applications. The wireless control switch is used to control the function of many buttons, such as Previous, PgUp, Next, PgDn Playback, End, Esc, Black Screen, and it can, likewise, set the countdown timer, vibration alert, and other functions, as shown in Figure 4, below. This proposed system can be operated wirelessly for presentations, including multi-media interactions. LaDIS can control and interact with any software image presented on a screen.

![Figure 4](image)

*Figure 4. LaDIS: Utilizes a wireless control electron circuit to execute multiple interactive controls*
Innovation laser-driven interactive system (LaDIS) solution

Potential benefits of LaDIS

Expenses

While IWBs offer support for teaching needs, they are not an economical solution for many users; for example, a common IWBs, costs range from 1,500 to 5000 US dollars (Lee, 2008). Back lit IWBs, such as back lit LCD (Liquid Crystal Display) or even back lit projectors are more expensive and less commonly utilized in classrooms than general projectors and screens (Richard, 2006; Park, 2010). Due to financial limitations, IWBs still cannot be effectively implemented into every classroom worldwide; this is especially the case in developing country and rural regions. As proposed by Johnny Chung Lee (Lee, 2008), the Wiimote whiteboard is an example of a cost-effective solution to replicate high-cost IWBs. Wii is the official main engine of a video game system created by Nintendo in 2007. Wiimote is the Wii’s standard controller, and it uses the Bluetooth wireless transmission with main engine's segment. However, Wiimote has function limitations because it is designed specifically for use with a video game controller and not for IWBs. The main limitations of Wiimote whiteboard are its setup, and calibrations are difficult. For example, the users must connect a Bluetooth between computer and Wiimote, and then use an infrared pen to calibrate each corner on the screen each time. Using a Wiimote whiteboard, try to connect Bluetooth between computer and Wiimote and calibrate by infrared pen is inconvenient and limited as the Wii was never designed specifically to function as an IWB. Rather than altering the Wiimote sensor for a secondary use, in this paper, we develop a new installation and calibration method, by using the facilities in a classroom, such as projector, screen and pointer. LaDIS propose a more viable, efficient and economical solution, specifically for IWBs. Generally speaking, LaDIS provides automatically calibration and without any complicated connecting procedure. Classrooms can be set up with LaDIS for around $50 USD. The main parts of a LaDIS device are simply a wireless presenter and a laser diode, which cost approximately $30 USD and other hardware components of the LaDIS, such as a webcam cost $10 to $20 USD. The developed software is freely downloadable at the following website: http://webcam-whiteboard.apponic.com/download/link-1/. In comparison to the thousands of dollars that a fully operational IWBs would cost to install, LaDIS is very well suited for rural and developing regions, to cultivate interaction and information didactics, as shown in table 1. In table 1, we don't only compare the cost of the LaDIS, Interactive Whiteboard (IWBs) and Wiimote Whiteboard but also the Manipulating, Set up and function of those systems. The results are a summary from the survey of ten professional teachers in this study. They all have experience in using LaDIS, Interactive Whiteboard (IWBs) and Wiimote Whiteboard in their classes.

<table>
<thead>
<tr>
<th>Table 1. Comparison of the LaDIS, Interactive Whiteboard (IWBs) and Wiimote Whiteboard (Lee, 2008)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LaDIS (This work)</td>
</tr>
<tr>
<td>Cost</td>
</tr>
<tr>
<td>Manipulating</td>
</tr>
<tr>
<td>Set up</td>
</tr>
<tr>
<td>IWBs function</td>
</tr>
<tr>
<td>Interactive function</td>
</tr>
<tr>
<td>Presenter function</td>
</tr>
<tr>
<td>Without direct control distance limitation</td>
</tr>
<tr>
<td>Without Screen Size limit</td>
</tr>
<tr>
<td>Air mouse function</td>
</tr>
<tr>
<td>Extensive control &amp; MBWA in classroom</td>
</tr>
<tr>
<td>The user can stand a safe and comfortable distance from the projection light or associated equipment and wires</td>
</tr>
</tbody>
</table>

Note. The results are a summary from the survey of ten professional teachers in this study.

Engagement for the students

The main problem with any IWBS, including Back-lit IWBs, such as back-lit LCD (Liquid Crystal Display) or even back-lit projectors is that whoever is using it usually has their back to the class as shown in Figure 5. In general (left),
teacher’s eyes may be strained by standing close to a high luminance projector light or screen. In this study (right), a teacher can utilize the LaDIS system to stand at a comfortable distance away from the projection light or screen, and still interact fully with the screen. If there is more than one person trying to use the board, they physically block the board from the rest of the class, or they cast a shadow over it (Bell, 2002; Walker, 2003b). This often causes an instant breakdown in engagement for the students in their seats who can no longer see what is happening on the board. Once there are two or more people writing on an IWBs, other students cannot see the board. The LaDIS system can provide an economical solution, to the problems stated above, via a laser emission unit and fully wireless control functions. By utilizing a hand-held LaDIS device, a class may keep their view of the screen unhindered, even as a student or teacher interacts with the viewing screen itself. That is to say, LaDIS allows for a touch-less interactive method between teacher and student without a breakdown in engagement for any student as shown in Figure 6. In general (as seen in the left picture), when two or more students are interacting with an IWBs; the class becomes very un-interactive as the other students cannot see the board. In this study (as seen in the right picture), every student may interact with display in the classroom from their seat: in order to answer questions and remain interactive alongside each-other, without a breakdown in engagement between students, teachers and the screen. The laser emission unit is for emitting a laser spot on a screen to launch a synchronous cursor movement with the laser spot. The laser is a direct light source and has no distance limitations within a standard classroom. Employment of wireless control technology, allows every student to interact with the display from their seat. In this way, students may answer questions and stay interactive while remaining physically alongside each-other. This touch-less function allows for student, teacher and screen interaction without a breakdown in engagement between these parties.

In general, cognitive learning can be divided into cognitive structure learning theory and information-processing theory of learning. Cognitive structure learning theory considers how the individual learner understands things, in terms of developmental stages and learning styles. Information-processing theory of learning explores how the input of sensory information starts at the top, goes through sensory memory or a sensory buffer into short-term memory (STM), and hence to long-term memory (LTM) (Brown, 2007). In this study, we try to demonstrate the potential of LaDIS system on cognitive learning. Based on LaDIS system, student can engage intuitive thinking and feedback learning by discussion with teacher and classmate during class. For example, a student can point out any question/problem and operate/interactive teaching material/software, by themselves, in classroom setting, from their individual seats. A teacher can easily display cognitive representations such as enactive representation, iconic representation and symbolic representation via the LaDIS system. For example, teacher can mark, paint note, maps and zoom in any teaching material/software. Students can also participate in any learning activity in class through the LaDIS system. Teachers may, likewise, design an activity allowing students to use the LaDIS system in completing guided discovery learning. The LaDIS system can increase student’s learning motivation via diversification learning activity and more easily turns short-term memory (STM) into long-term memory (LTM). In other words, the LaDIS system is a powerful tool for engaging the cognitive learning process.

![Figure 5. In general (left). In this study (right)](image)

**Health and safety**

In general, as a teacher or student uses an IWB, their eyes may be strained by being in close proximity to a high luminance light source (Mark 2009). There is a concern often expressed regarding the health and safety implications of the multitude of wires required for IWBs and associated equipment (Bell, 2002; Smith, 2005; Cogill, 2003). Teachers also report that they need to stand to the side of the board or a shadow is cast over the screen as the light
shines on them (Bell, 2002; Walker, 2003b), a difficulty, likewise, experienced by pupils (Smith, 2005). Another virtue of the LaDIS system is that the laser emission is a directed and focused light source pointed away from the users, to the board. It provides a touch-less interactive method of use. Therefore, the user can stand a safe and comfortable distance from the projection light or associated equipment and wires while still being fully interactive with screen.

![Image](image1)

**Figure 6. In general (left), In this study (right)**

**Time wasted**

Another major disadvantage of the IWBs is the amount of time wasted as each student, one at a time, walks out to the front of the screen, selects their answer then hands the pen to the next student who walks out to the front to answer their question. This turn-taking delay and slow pace of the lesson can be very frustrating for students, especially those who do not get a go up to the front of the class due to time constraints (Smith, 2005). In this study, students can select their answer or interact with screen from their seats by LaDIS without walking up to the front of the class. In other words, teachers can save time and avoid the slow pace of students taking turns to walk up to the screen, as shown in Figure 7. Students, one at a time, walk out the front, select their answer then hand the pen to the next student who then walks out to the front to answer their question. In this study (right), students can select their answer or interactive with screen from their seats via LaDIS without walks out to the front of the classroom.

![Image](image2)

**Figure 7. In general (left), In this study (right)**

**Screen size limitation**

The height at which an IWBS is placed can be an issue, particularly where boards are permanently fixed and if pupils are to use them (Cogill, 2003). If the board is placed too low (or made small) on the wall, the screen may not be seen by pupils at the back of the class, and some functions of the board may be difficult to operate (Cogill, 2003). If the board is placed too high (or made large), then, teachers may have difficulty reaching the top. The size of the screen is a related factor to consider (Smith, 2005). The Wiimote IWBS uses coordinate tracking and guides the computer
mouse cursor by detecting the IR spot. However, Wiimote IWBS has a detection distance limitation, and the screen size is also confined within a certain short range (Lee, 2008); the designer proposed a design to utilize a web cam (optical camera) and a special light filter lens to detect the laser spot on the screen effectively. LaDIS system provides a total solution by using the facilities in a classroom, such as projector, screen, and a LaDIS device. As shown in Figure 8, the direct laser light source, allows teachers to reach the top of the screen utilizing the LaDIS system, easily. The screens, presented in the Figure 8 are the standardized size for schools, of all levels, in Taiwan. The reason why they need a large-size screen is in order to allow everyone a clear-view. In most Asian countries, including Taiwan, each class has more than 30 students. Based on this reason, LaDIS is designed to suit large-screen sizes, Asia teachers and their classes. With a projector, screen, PC equipment, already in classrooms, schools barely need to invest any extra funding for LaDIS to provide them with a viable, efficient interactive system.

Figure 8. If the screen is placed too high (or is too large), then teachers may have difficulty reaching the top of the display. In this study (right), a teacher utilizing the LaDIS system may easily point to the top of the screen

Management in classroom

Some studies indicate that Management by Walking around (MWBA) in a classroom positively affects instructional practices and student achievement (Todd, 2005). However, in order to control the interaction teaching material on the screen more effective, the teacher is restricted to stay in front of the white board. The result is that they cannot execute MWBA in a classroom by using IWBs. On contrary, owing to a touch-less and wireless interactive control device, a teacher can easily accomplish MWBA in a classroom by means of the LaDIS system. By utilizing a focused laser light source, a teacher can remain interactive with the screen and execute MWBA from any place within a classroom as shown in Figure 9(a). In general, the teacher is physically limited to staying in front of the white broad while using it. The result is that they cannot execute MWBA in a classroom while using IWBs. In this study (right), teacher can interact with the screen and execute MWBA from any part of a classroom by utilizing the LaDIS system.

Figure 9(a). In general (left), In this study (right)

LaDIS is a multi-function system. Every student may interact with display in the classroom from their seat to answer questions and remain interactive alongside each-other, without a breakdown in engagement for any student via its touch-less function; but LaDIS may also perform utilizing a touch sensitivity function as if the user were operating an IWB. This study provides teachers with alternative function options. A teacher can select different function modes depending on their teaching needs as shown in Figure 9(b). From their seat, every student may interact with the
display in the classroom, allowing them to answer questions and remain interactive alongside each other. By utilizing LaDIS’ touch-less function (left picture) this interaction occurs without a breakdown in engagement. Users may also operate LaDIS’ touch sensitivity function as an IWB (right picture).

![Figure 9(b). LaDIS’ touch-less function (left picture) and touch sensitivity function as an IWB (right picture).](image)

**Multimedia teaching method needs**

The proposed work wisely offers a viable solution to challenges faced by IWBs. The user can operate via far distance control didactics. IWBs provide an interactive, precise and multimedia teaching method in front of the screen. For a conventional IWB, the user has limited functions in presentation work by using an infrared detector and infra LED, like Wiimote. This proposed study provides a full solution for presentation needs, such as an air mouse, air presenter and air writing function as shown in Figure 10. Due to the property of far distance control, the proposed study can operate in a large classroom as well as on a large screen for more attractive didactics. Thus, we can say that it is a distance free (touch-less) detector. The trajectories of mouse movement, such as the trajectory path, icon pattern, and writing symbol, are all made possible via LaDIS, allowing for diverse computer commands and superior interaction with multimedia, multi-functional and touch-less behavior. This study also proposes a practical application on general multimedia didactics methods.

![Figure 10. The proposed design performs air writing (left), air mouse and air presenter functions (right).](image)

**Drawbacks in LaDIS**

Laser safety is the use and implementation of lasers to minimize the risk of laser accidents, especially those involving eye injuries. Since even relatively small amounts of laser light can lead to permanent eye injuries, the sale and usage of lasers are typically subject to government regulations. A Class-2 laser is safe because the blink reflex of the human eye will limit the exposure to no more than 0.25 seconds. It only applies to visible-light lasers (400–700 nm). Class-2 lasers are limited to 1 mW continuous wave, or more if the emission time is less than 0.25 seconds or if the light is not spatially coherent. Intentional suppression of the blink reflex could lead to an eye injury. Many laser pointers and measuring instruments are class 2. For class 2 laser pointers, the blink reflex of the human eye (aversion response) will prevent eye damage, unless the person deliberately stares into the beam for an extended period. Output power may be up to 1 mW. This class includes only lasers that emit visible light. Most laser pointers are in this category. In this study, our LaDIS is designed as Class-2 laser and its output power is only 1 mW. General
speaking, LaDIS is a safe device for teachers and students. However, laser beams still have potential risks if teacher or student deliberately stares into the beam for an extended period of time. Mainster et al. (2003) provides one case, an 11-year-old child who temporarily damaged her eyesight by holding an approximately 5 mW red laser pointer close to the eye and staring into the beam for 10 seconds, she experienced scotoma (a blind spot) but fully recovered after 3 months. Luttrull & Hallisey (1999) describe a similar case, a 34-year-old male who stared into the beam of a class-3, 5 mW red laser for 30 to 60 seconds, causing temporary central scotoma and visual field loss; his eyesight fully recovered within 2 days. Therefore, teacher must remind students not to stare at, or direct laser beams, at themselves or other students, as these are inherent risks and drawbacks in the LaDIS design.

Attentive student, enriched learning

Impact of LaDIS on students’ cognitive learning outcomes

Background

The samples of this study include 92 first-grade students enrolled in a 5-Year Junior College; these individuals were from two classes at a college of Taoyuan County, in Taiwan. The college, in this study, has experienced tight budgeting problems for a number of years and has no capability to afford IWBS in any classroom. Its financial situation is, likewise, similar to many schools in developing or rural regions worldwide. However, most teachers in this College have the ability, education and need to adopt IWBS didactics into their classes to improve student learning motivation, increase efficiency of teaching as well as outcomes of learning.

Participants

In this study, each class consists of 46 students, and they are assigned by normal college assignment procedures. Classes are twelve hours in length, for a total of six weeks in the experimental course with different didactics. The instructor for each group is addressed the same content and only didactics differences are: “with or without” the innovation of LaDIS in the class. One class is the control group, and it adopts ordinary didactics. The other one is an experimental group, which adopts the innovation LaDIS didactics in class. The instructor in the control group is the same as in the experimental group. It should be noted that only tradition mouse didactics are used in the control group. “Cooperative Learning” and “Enquiry Learning” activities are arranged in every class.

Data collection

Before the course, both classes took part in a pre-test of living science course achievement. After completing the course, the two groups took a post-test of living science course achievement and Student Satisfaction Index (SSI) to explore the students' learning attitudes towards the innovation LaDIS on didactics application.

Statistical analysis

As shown in Figure 11, the mean level of two classes students’ prior domain knowledge was low ($M = 42.9, SD = 9.8$) & ($M = 42.7, SD = 8.5$). After receiving the ordinary and the innovation LaDIS, two kinds of different didactics, almost all experimental group students make remarkable progress on their post-test score in the class ($M = 81.8, SD = 12.2$). The post-test score for the control group ($M = 72, SD = 14$) was modest compared to the gains of the experimental group (independent-samples $t$-test analysis on control and experimental group students’ pre-test ($t = -.08, p > 0.5$), post-test ($t = 3.567, p < .001$) scores). The results of the univariate ANCOVA analysis (covariance on the post-test scores, with students’ pre-test scores as the covariate) reveal that students taught using the innovation LaDIS didactics scored significantly higher than did students within the control group.
Main findings

According to Bloom, cognitive objectives are concerned with intellectual outcomes. The classification system ranges from lower-level knowledge outcomes to higher-level intellectual abilities and skills. In this study, we focus on understanding, including comprehension, understanding relationships, meaningful learning, remember paraphrased, defined concept and concrete concept. As shown in the Table 2, $F(1, 90) = 82.15$, $\eta^2 = 0.480$, $f = 0.96$ (very large effect size) (Cohen, 1988), this result shows the positive impacts on their achievement in living science course. The results demonstrate that “Cooperative Learning” and “Enquiry Learning” benefit from the adoption of the LaDIS interactive system. It is shown that, there is better interaction between student and student & student and teacher in the classroom due to the use of this form of interactive system, as shown in Figure 12. The aim of this study is to show how LaDIS works in a classroom, as well as to demonstrate how its performance positively impacts student’s achievement in a course. We have no intention of comparing LaDIS and IWBs, in terms of which one is better. The positive effects of IWBs application in education, is not in question. However, the cost of an IWB is 100 times that of LaDIS. The intention of this study is to provide another choice, in regard to IWBs, and solution for schools with economic concerns, especially those in rural and developing regions.

Table 2. Univariate ANCOVA analysis of covariance for student achievement

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Hypoth. SS</th>
<th>Error SS</th>
<th>Hypoth. MS</th>
<th>Error MS</th>
<th>F</th>
<th>Sig.of F</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>achievement</td>
<td>2261.49</td>
<td>2450.07</td>
<td>2261.49</td>
<td>27.52</td>
<td>82.15***</td>
<td>0.000</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Student’s SSI

As shown in Figure 12, Evaluation results for the Student Satisfaction Index (SSI) are separated into four dimensions, including: Satisfaction: meaning, whether (or not) students are satisfied with LaDIS in class; Willingness to use: indicating whether (or not) students are willingness to use LaDIS in class; Interaction: meaning, whether (or not) LaDIS increases the interactivity in class; and Learning, indicating, whether (or not) LaDIS improves learning in class. Between each two dimensions are also included two sub-items: (1) between satisfaction and learning is “motive and effect”: indicating whether (or not) LaDIS increases the motive of a student and/or effect on a student; (2) between satisfaction and willingness to use is “better than traditional and usability”: meaning, whether (or not) LaDIS better than a traditional device and LaDIS’ usability; (3) between willingness to use and interaction is “perceptive and participation”: meaning whether (or not) LaDIS increases the perceptive and participation of student; 4) between learning and interaction is “interaction with teacher and interaction with classmates”: indicating whether (or not) LaDIS increases the interaction with teachers and classmates. The results show each dimension as follows: satisfaction (90%), willingness to use (93%), interaction (95%), learning and sub-items such as motive (91%), effect
interaction with teacher (96%), interaction with classmates (93%), perceptive (93%), participation (93%), usability (91%), better than traditional devices (94%). All the above results obtain over 90% ratings on the satisfaction index. These findings demonstrate a large improvement between student’s learning and interaction in the classroom due to the use of the LaDIS interactive system. Most students are highly satisfied with, the system and willing to use the Laser-Driven Interactive System (LaDIS) because of its naturally superior interactive performance, in classes.

Figure 12. Evaluation results for the Student and Teacher Satisfaction Index (SSI)  
(URL :http://www.mysurvey.tw/hm4zYGES)

Results and discussion

Reports show that interactive teaching, supported by the use of IWBs, can lead to positive learning outcomes (Mercer et al., 2010). When the IWBs is set up as a technology hub, it has been proposed (Miller & Glover, 2010) that it can act as a trigger for pedagogic change (Miller & Glover, 2010). However, the main shortcoming of IWBs is its costs. The other disadvantages of IWBs include:

- When two or more students are interacting with an IWBs; the class becomes very un-interactive as the other students cannot see the board.
- The teacher’s eyes may be strained by standing close to a high luminance projector light or screen.
- Students usually walk out the front of the class, one at a time, selecting their answer then handing the pen to the next student who then walks out to the front to answer their question.
- If the screen is placed too high (or is too large), then teachers may have difficulty reaching the top.
- The teacher is physically limited to staying in front of the white-broad while using it.

The effect of the proposed interactive system that incorporates LaDIS into a course may be summarized as the follows: (1) With a projector, laser emitter, screen, PC and web cam, it is easy to show a viable, efficient interactive system in any classroom. (2) The proposed LaDIS system provides an economic educational system for developing and rural regions. (3) Most students are highly satisfied with, the system and willing to use the Laser-Driven Interactive System (LaDIS) because of its naturally superior interactive performance in classes. (4) This study demonstrates how LaDIS excellent performance, positively impacts students’ achievement.

LaDIS provides a new and innovative, economically effective, interactive software and presentation technique for general multimedia didactics. The proposed design utilizes a laser point which comes from a common laser pointer to emit a red spot onto a screen. It cooperates with a web cam (optic camera) and a laser point image tracing software. After some imaging and software program operations, the movement of laser point images can be transformed into a command signal to control the mouse’s cursor on the screen. In association with multimedia presentation software, the proposed design can efficiently present a superior performance with said multimedia, being multi-functional and touch-less in nature, via controlling the mouse’s locus.
The results indicate positive impacts of the innovation LaDIS on students’ learning performances. Most students are highly satisfied with the system and willing to use LaDIS because of its superior interactive nature, compared to the traditional classroom. Preliminary results suggest that the innovation of LaDIS is promising in facilitating students’ cognitive learning outcomes in didactics applications. The results demonstrate that “Cooperative Learning” and “Enquiry Learning” are fit to adopt the innovation of the LaDIS interactive system. Simply put, LaDIS improves interaction between student and student & student and teacher in the classroom. The main pedagogical advantages of LaDIS appear to be:

- Write-ability: allowing drawing, annotating and writing on a screen.
- Point-ability: allowing learners and teachers to point, physically, to interesting features on the board, rather than using a mouse pointer which can be hard to see and manipulate.
- Size: in promoting group working allowing learners and teachers to work collaboratively with software around a large, shared work area.
- Talk-ability: allowing learners and teachers to talk more directly to the audience from the front, freeing them from the PC, the keyboard and the mouse.
- Familiarity: encouraging both learners and teachers to engage with technology more comfortably through a more familiar interface.

Why LaDIS is particularly suitable for developing and rural learning settings:

- Infrastructure: 1. The key parts of a LaDIS device cost about $30USD. (The main parts of this device include a wireless presenter and a laser diode). 2. The price of a webcam ranges from $10 to $20 USD. However, a webcam is already a standard device in most laptops, so if the classroom has a laptop, then this cost can be discounted. 3. Even in developing and rural areas, a computer or Easy Personal Computer (EPC, low-cost laptop) usually part of the standard equipment available in classrooms in most countries. 4. In a standard classroom, a projector usually hangs from the ceiling of classroom. In the other words, based on the basic infrastructures existing in most schools and classrooms, such as an Easy Personal Computer and a projector, it is easy to establish an interactive classroom via LaDIS; this is the main reason why LaDIS is particularly suitable for developing and rural learning settings.

- Student needs: Student can point out, mark, maps and zoom, in or out, and operate/interactive teaching material/software by themselves, in a classroom, from their seats. Therefore, student can engage intuitive thinking and feedback learning by discussion with teacher and classmate during class. Student also can feel free to participate in any learning activity in class through the LaDIS system. Through an activity, student can complete guided discovery learning by using a LaDIS system. Without a low-cost LaDIS system, most developing country and rural area would have no chances to bring about IWBs type interactive learning into their class.

- Teacher’s learning style: Most teachers today are trained to be adept in utilizing computers. They are, likewise, trained to create various teaching materials by using computers. With information technology aids, teacher can increase student’s learning motivation and attract students’ attention (Masalela 2001). Combining computers and interactive teaching aids is a popular teaching and learning style for teachers (Abowd, 2000). However, a lot of developing country and rural area don’t have adequate budgets to support this brand-new teaching and learning style in their classes. The low-cost of the LaDIS system can be the solution, to the problem of the high costs of IWBs, in any developing country or rural area.

The aim of this study is to show how LaDIS may function in classrooms as an economical solution to IWBs. IWBs costs ($5000-$1500) almost 100 times more than LaDIS ($50-$25). Based on this study, we have the intention of providing an economical solution for schools with financial limitations, as those found in rural and developing regions. As a teacher said in our study, “If I can do the same thing with a tablet that costs $50 rather than the hefty $1500 to $5000 price tag that some IWBs cost, then that is the direction I think I would want my classroom and school to head”.

There is still more research that needs to be done. Comprehensive studies to uncover potential problems in LaDIS are being undertaken. Further studies must be conducted to explore what subjects and teaching & learning mode are suitable to adopt the innovation of the LaDIS interactive system. As we know, LaDIS is a brand-new system, for use in any IWBs and interactive teaching software. In fact, schools barely need to invest any extra funding to acquire a LaDIS. LaDIS is well suited for rural and developing regions due to its low cost. As we look to the future, we are
starting by providing LaDIS to schools, in underdeveloped areas, which can easily add this innovation, with a projector, screen and PC equipment, already found in most classrooms.

Acknowledgements

The authors acknowledge the financial support of the National Science Council of Taiwan, Republic of China (Grant No. NSC-102-2511-S-003-052-MY3.) and this research is partially supported by the “Aim for the Top University Project” of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and the “International Research-Intensive Center of Excellence Program” of NTNU and National Science Council, Taiwan, R.O.C. under Grant no. NSC 103-2911-I-003-301.

References


Effects of Communication Competence and Social Network Centralities on Learner Performance

Il-Hyun Jo*, Stephanie Kang1 and Meehyun Yoon2

1Department of Educational Technology, College of Education, Ewha Womans University, Seoul, Korea // 2Department of Learning, Design, and Technology, College of Education, University of Georgia, USA // ijj3174@gmail.com // stephaniekang83@gmail.com // meehyun@uga.edu

*Corresponding author

(Submitted August 13, 2012; Revised March 11, 2013; Accepted February 21, 2014)

ABSTRACT

Collaborative learning has become a dominant learning apparatus for higher level learning objectives. Much of the psychological and social mechanisms operating under this complex group activity, however, is not yet well understood. The purpose of this study was to investigate the effects of college students’ communication competence and degree centralities of their social networks on learning outcomes in a collaborative learning situation. The study participants were 63 students of educational technology at a women’s university in Korea. Path analyses of the data using NetMiner 3.4 and AMOS 7.0 indicated that: 1) communication competence caused trust network degree centrality ($\beta = .24$, $p = .13$), 2) communication competence caused knowledge sharing network degree centrality ($\beta = .46$, $p = .00$), 3) trust network degree centrality enhanced knowledge sharing network degree centrality ($\beta = .41$, $p = .00$), and 4) knowledge sharing network degree centrality affected individual students’ learning outcomes ($\beta = .55$, $p = .00$). The study results revealed the significant collective effects of network degree centrality measures and individual communication competence on learners’ performance. Based on these results, implications for team organization strategy and future research directions are discussed.

Keywords

Collaborative learning, Social network analysis, Communication competence

Introduction

Background and purpose of the Study

The purpose of collaborative learning is to develop students’ attitude and competence to interact and collaborate to pursue common goals, while reducing the sense of alienation and hostility from the traditional class situation that creates a high level of competitive spirit (Slavin, 1995). Unlike individual learning, collaborative learning enhances interdependency in groups based on individual learner’s competence and the sense of responsibility. Therefore, the effectiveness of collaborative learning depends on social condition in the team, as well as individual cognitive and affective factors (Johnson & Johnson, 1994).

Social Network Analysis (SNA) has received attention both as a theory and as a method to understand and analyze the social attributes of a learning group (Jo, 2008). SNA can be an appropriate measure to investigate the characteristics and pattern of interpersonal relationships and even of the communicative interaction within a certain group (Cho, Gay, Davidson & Ingraffea, 2007; Refay & Chanier, 2002). A social network approach can also be applied to measure the degree and strength of social ties within a social system and to develop models of community structures (Girvan & Newman, 2002). In this perspective, a social network acts as a conduit of information and knowledge and assists in the formation of social capital such as trust among members of the society, and is appropriate to explain collaborative learning situation (Borgatti, Everett, & Freeman, 2002).

SNA considers that the network structure and individual position in the network are the key factors that decide the group action as collaborative learning. It has its epistemological root in structuralism. However, there is a strong counter argument against this structural deterministic perspective on SNA. This counter argument is based on the critical mind that individual roles such as individual competences and effort are ignored (Jo, 2008). The argument that reflects their critical mind above is that it is still impractical to underestimate the effect of a learner’s independent role, or agency, and responsibility on the formation and development of a network, even though the attribute of the network and the individual position in that network play important roles in collaborative learning. In
this context, research examining how the individual role and responsibility influence the network is necessary. We can expect that an individual variable, especially communication competence, is related to network development in a collaborative learning situation because it directly influences communication (McCroskey, 1987), which is known as a strong predictor for dense social network.

It can therefore be inferred that communication competence enhances individual centrality and position, and that this enhanced position has a positive effect on collaborative learning outcome. However, no reported empirical research has investigated the effect of individual variables and social variables on group performance.

The purposes of this study are: 1) to investigate the effects of college students’ communication competence on the degree centralities of their social network, and 2) to analyze empirically how their communication competence is related to their learning outcome in a collaborative learning situation.

The specific study questions are:
- Does communication competence of the individual affect the degree centrality of knowledge sharing network?
- Does communication competence of the individual affect the degree centrality of trust network?
- Does trust network affect knowledge sharing network centrality?
- Does knowledge sharing network centrality affect individual learner performance?

This study is worthwhile as it examines how communication competence, an individual attribute, affects network centrality and performance in a face-to-face collaborative learning situation in university. This study is expected to contribute to real world practices by defining what instructors should consider in terms of forming and operating a collaborative learning team.

**Literature review**

**Communication skills**

*The concept of communication competence*

Communication researchers have been considered that, just as intelligence, differences in communication competence exist between individuals. When trying to communicate with any people in different relationship, people show different behavior within several types (McCroskey, 1987; Richmond & Roach, 1992). Much of this behavior is attributed to the competence of individuals (Norton, 1978). Norton’s multivariate communication style assumes that nine independent variables and one dependent variable can determine and describe communication style construct. Similarly, Rubin & Martin(1994) developed an interpersonal communication competency scale (ICCS) consisting of 10 self-reported measures of ICC skills which were reduced from the original 60 skills. In this scale, ICC refers “an impression or judgment formed about a person’s ability to manage interpersonal relationships in communication settings.” Communication competence is a natural property to some trait but is also a developed competence by interacting with the context of communication environment. For example, in public speaking, discussion in a small–group collaborative learning, and one–on–one talking with friend, different types of communication competence are expressed depending on the communication context, and a long–term exposure to a particular environment can help developing appropriate communication competence for it (McCroskey & McCorseky, 1988). In addition, we can assume that an individual who is willing to communicate both verbally and nonverbally in a situation tends to have a positive interpersonal relationship. Freeman (2004) also considered communication competence (or communication adaptability) as an enhanceable personal competence as socio-interpersonal relationship develops and at the same time, competence to be needed to develop interpersonal relationship and adapt interaction behavior.

*Communication competence and collaborative learning*

Empirical studies have been performed to identify the hypothetical relationship between communication competence and collaborative learning. The results of these studies confirm that in collaborative learning situations, an
individual’s communication competence and types of communication play an important role in learning outcome and team activities (i.e., Jeong, 2005; Allen, Long, O'Mara, & Judd, 2007).

In collaborative learning situations, learners are exposed to various types of communications. One-on-one conversation with colleagues, small group discussions with team members gathered, and presentation to target the entire class is the communication situations that are frequently performed in collaborative learning in college classes. Thus, an individual’s overall communication competences during collaborative learning activities can be expected to play an important role.

According to Richmond and McCroskey (1998), Andersen and colleagues’ (Andersen, Norton, & Nussbaum, 1981) study researched independently in face-to-face classroom environment, communication competence of teachers and students had a significant influence on cognitive learning outcomes, as well as the learner’s ongoing emotional learning intentions. In addition, an individual’s competence, which enables open, friendly, and responsive communications, creates active participation and innovation thinking in collaborative learning. Furthermore, it is reported as being related to effective behavior to deliver those innovative thinking (Allen, Long, O'Mara, & Judd, 2007).

In collaborative learning, the importance of mediating activities related to communication such as dialogue, collaboration, and social exchange has been consistently emphasized. Specifically, an individual’s communication competence positively affects various group collaborative activities and, ultimately, creates a trust network, which is a team that is able to trust each other (McCroskey, 1987). In addition, in a learner group with more excellent communication competence, more seamless sharing of knowledge activities tends to be made more frequently (Kilduff, 1992). Furthermore, it has been reported that leaders with outstanding communication competence get strong support from members, which enables the leaders to exert leadership effectively and, ultimately, contribute to organizational performance results (Richmond & McCroskey, 1998). Based on the results of these previous studies of collaborative learning situation that conducted in the context of business corporate education, it is predictable that learner’s communication competence enables more effective communication and also contributes to the learning outcome by facilitating the trust and knowledge sharing within team.

Social network analysis

Social network analysis and degree centralities

SNA is a mathematical analysis tool that displays the social structure diagrammatically in the form of nodes and links that connect each node (Jo, 2008). This analysis is based on a view of structuralism in the way that social relational phenomenon, collaboration and knowledge sharing, can be explained more appropriately by an actor’s location in the network structure than the individual property of the actor (Freeman, 2004).

In SNA, the unit of analysis is divided on the microscopic level, which is aimed at an individual actor, and the macroscopic level, which is aimed at the network. First, at the microscopic level, the act of exchanging social, cognitive, or material resource by an individual actor(a node) and the other actor(another node) pairing with a link is analyzed. This relationship between the pair may be strong or weak depending on the opponent. Degree centrality frequently used in SNA is a typical microscopic indicator that represents the degree of linkage (Wasserman & Faust, 1994).

Degree centrality and communication competence

Little previous research has examined the effect of individual variables such as communication competence on the formation and development of a social network within a group. The constructivist view that many SNA researches implicit tends to exclude the possibility that a psychological factor could have been a crucial part in deciding individual behavior (Burt, Janotta, & Mahoney, 1998; Kilduff, 1992).

Meanwhile, some social network researchers have recognized the empirical fact that an individual character plays an important role in the formulation of a relationship with others (e.g., Burt et al., 1998; Mehra, Kilduff, & Brass,
Mostly in a business corporation context, these researchers studied how the individual attribute affects the structure of the social network within organization. They found that individual communication competence not only decides the position in the network that is measured in the degree centrality within the organization (Burt et al., 1998), but that it is also a predictor that leads to social strategic distinction (Mehra et al., 2001). Specifically, people with an entrepreneurial personality take a brokerage position in a network (Burt et al., 1998), while people with an introspective personality proactively utilize the degree centrality in the network to increase their influence within the organization (Mehra et al., 2001).

Studies that attempt to determine the relationship between the type of individual communication and the network index have examined cooperation organization. Thus, from these previous studies it was hard to find direct empirical reference to connect synthetically how communication competence and type affect the social network in collaborative learning context. After examining previous studies that examined the relationship between communication and collaborative learning, collaborative learning and SNA, this study will deduce implications in collaborative learning context by connecting the findings of the two studies.

Degree centrality and learning outcome in collaborative learning

Degree centrality is calculated as the number of links that an individual actor has, and it reflects the individual actor’s network activity intuitively. In SNA, a person with a high degree centrality is a center actor who is defined as a person who has the largest number of links with other actors within a particular communication group or network (Picciano, 2002). This indicator focuses on prominent actors in the group. The person is interpreted as a positive person in the relational process. Conversely, the person whose index is low can be judged as someone who plays a marginal role in the group.

Thus, individuals with a high degree centrality are likely to have a favorable position in the network in two ways. First, relatively less effort is required when finding the information source from fellow students. Because learners with high degree centrality can find and access sources of information quickly and easily, because they are well known to fellow students in the team and class and because they know who knows what. Second, people with high degree centrality can exert a powerful influence when they acquire information from sources that they found. This is because degree centrality acts as prestige and influence within a network (Wasserman & Faust, 1994). Collaborative learning is a network phenomenon so that such general traits of degree centrality are expected to appear in collaborative situation similarly (Jo, 2008).

Based on these theoretical possibilities, empirical studies that identify the relationship between degree centrality and collaborative learning outcome have been performed. Papa and Tracy’s study showed that people with a high degree centrality within an organization also achieved high learning outcomes (Papa & Tracey, 1988). Jo conducted research on collaborative learning teams in college courses, identified how in–degree centrality of trust network affects individual outcome and how knowledge sharing out–degree centrality affects team outcome (Jo, 2008). Another study conducted in a blended learning environment integrated by classroom and e–learning which studied the correlation between relevant variables such as various network centrality and learning outcome, and discussion quality. In the results of those studies, advice network centrality and friendship network centrality had positive correlation with learning outcome, while hostile network centrality had negative correlation with most of the performance variables.

Research has been performed on the impact of social network indexes on learning outcomes for short–term performance, as well as on learning team’s learning potential and network development for the long–term impact (Jo, 2012). Palonen and Hakkarainen (2000) reported that the degree of network composition is an important precondition for the creation and sustainable development of learning community, since it helps members share cognitive and affective experience and supports informal communication and social activities (Palonen & Hakkarainen, 2000). This collective, social activity lets individual students enhance their belonging within the learning group (Haythornthwaite, 2000).
Trust network and knowledge sharing network

Individuals belong simultaneously to several networks such as high school reunion network, cooperation network in work, and weekend society network. Typical social networks that an individual student can belong to are knowledge sharing network and trust network. There is a relatively direct correlation between these two networks (Jo, 2008).

First, knowledge sharing network plays a direct role in collaborative learning. In the collaborative learning process, the team performance and knowledge expansion of individual members is expected to be improved. Especially in valuable implicit knowledge, the network with the purpose of knowledge sharing between students plays an important role (Wasko & Faraj, 2005). Thus, the process and result of collaborative learning can be influenced by the form and level of knowledge sharing network.

Meanwhile, the trust network indirectly but consistently influences the team outcome by helping the formation and development of a knowledge sharing network (Cross, Parker, & Borgatti, 2000). Kim & Mauborgne (1997) also discussed the importance of trusting atmosphere for improvement of performance by knowledge sharing. In a general knowledge creating situation that knowledge giving is not evaluated as individual level, knowledge giving to others can damage on the knowledge giver’s status of knowledge. According to previous studies, when an individual gives knowledge to others, knowledge giver loses one’s monopolistic status of knowledge (Thibaut & Kelley, 1959). Thus, if trust is not formulated in the team, people tend not to share knowledge and information with their colleagues that can improve team performance. The development of a trust network between team members is necessary to provide positive criticism on colleague’s idea (Misztal, 2001). Previous studies found that unless they trust each other, team performance cannot be improved irrespective of how outstanding and high the individual’s level of knowledge (e.g., Cross et al., 2002; Nahapit & Goshal, 1998). To sum up, trust is a factor which makes people to feel less threat when they provide or share their knowledge to other people such as colleagues. The trust network and trust itself have positive influences on knowledge sharing, and trust among team members can build a sense of fellowship so that enables smooth and active production and sharing of knowledge (Cohen & Prusak, 2001). It leads to higher performance of team and organization.

Study hypotheses

Taken together with previous studies, in collaborative learning situation, the individual communication competence of the learner improves the status in social (trust) and cognitive (knowledge sharing) network, which can have a positive effect on improving learner performance. In this, researchers grasped that mutual dependence building is required prior to knowledge sharing. Hence, the researchers formulated the following hypotheses:

Hypothesis 1. Individual communication competence will enhance knowledge sharing network degree centrality.
Hypothesis 2. Individual communication competence will enhance trust network degree centrality.
Hypothesis 3. Trust network will enhance knowledge sharing network degree centrality.
Hypothesis 4. Individual degree centrality of knowledge sharing network will enhance learner performance.

Method

Study procedure

Study Participants

The study participants were 63 students of “Theory of Corporate Education,” which was open in the department of educational technology from a women’s university located in Seoul, Korea. The objectives of this class, which was established by the principal researcher of this study, were to acquire the basic theories of corporate training and to analyze the current status of applying the approach of educational technology through business field survey.

The 63 study participants were comprised of 42 sophomores (66.7%), 9 juniors (14.3%) and 12 seniors (19.0%), of whom 52 participants (82.5%) were majoring in educational technology while 11 (17.5%) belonged to other
departments. All 63 participants were female because it was held in a women’s university. The response rate of the survey was 100% (63 out of 63).

**Process of collaborative learning class**

In order to achieve the class objectives, the students were assigned by collaborative learning task in which they should submit a report showing the differences and common elements between corporate education and educational technology after researching the current status of the training and development practices in Korea’s business corporations. In order to achieve the high–level learning objectives through conducting various activities including literature analysis, field survey and result analysis, and sharing each member’s original knowledge and experience, the collaborative learning based on team units was performed throughout the whole semester.

In the first week of the semester, the researcher organized the students into four (4)-member collaborative learning teams. According to Barkley et al., (2005), the appropriate size of team to improve social skills and mutual interdependence while minimizing the risk of free loaders by collaborative learning is 3 to 4 member group. The goal of the team composition was to create teams with diversity in terms of grade, major and matter of transferring, because a team of diverse members not only creates interdependent collaborative relations between team members (Druskat & Pescosolido, 2002), but also becomes one of the main objectives of collaborative learning based on its experience of diversity. Each team decided the team members in charge of each of the four tasks (data research, analysis & organization, writing a report and writing a presentation data) suggested by the principal researcher. These types of task are regarded as authentic, ill-structured, and requiring team collaboration. (Marks, Zaccaro, & Mathieu, 2000).

In order to expand the width of sharing knowledge and experience that may have been limited by dividing tasks for each team and individual to the level of the whole class, a jigsaw session for each task area was operated twice a month as an activity for the whole class. This provided an opportunity to share problems, data and matters discussed in each team with the whole class.

There was also an opportunity to present the interim results of each team to the whole class once every three weeks. After the other team members filled out the feedback questionnaire anonymously, it was organized as a document by the researchers in charge of the class, and the document was delivered to the relevant team. The feedback questionnaire was comprised of two qualitative and quantitative categories. In the quantitative aspect, the marks (total score: 100) for each category, including the feedback on “presentation data” (detailed items were logicality & systemicity of composition, and fulfillment of the analysis date) and “presentation activity” (detailed items were persuasive power of the presenter and the presentation, and propriety of using the presentation time), were provided to each relevant team. The qualitative evaluation involved filling out a form about the desired requirements. Therefore, this class included the collaborative learning not only for each team, but also for the whole class as in the jigsaw session and feedback from other team members.

**Some post-hoc observations**

As reported previously, this study relied on data from self-reported surveys, which may be vulnerable to respondents’ subjectivity and social desirability. To triangulate the interpretation of the main metric data, a follow-up 2 hour group interview with 15 students who volunteered. The interview data provided some additional information. First, according to the interviewees, the respondents had enough time to carefully read the survey instrument, which provided a relevant level of engagement when answering the survey. Second, the interviewees unanimously agreed that the most relevant group activity was jigsaw session, from which the students gained fresher task perspectives and opportunity to test the original teams’ approaches.

**Measuring tools**

**Measuring procedure for each time**

In the first week of the semester, individual communication competence was measured. During the semester, the team collaborative learning was continuously conducted. In this process, the researchers carried out an office hour
Measuring communication competence

In order to measure individual communication competence, this study used the Self–Perceived Communication Competence (SPCC) scale developed by McCroskey and his colleague (McCroskey & McCroskey, 1988). This measuring tool is comprised of 12 questions that are supposed to be answered with percentages. To answer the relevant questions, a respondent marks 0 points if he/she thinks that he/she cannot do it at all, and marks 100 points if he/she thinks that he/she can do it completely. In order to add up the results of responses for each question, the SPCC tool calculates the average value of the relevant question for four conditions (one–to–one, small group, meeting, public speech) and for three objects (friend, acquaintance, stranger). For example, the score about the stranger is the response average of the average value of four related questions (1, 4, 7, 10), while the score of the group condition is the response average of the average value of three related questions (4, 9, 11). The final score was inserted into the communication competence variable within the model as the average value of the scores of each of seven subordinate areas of SPCC. The reliability (α) of internal questions for each subordinate area was .74 (public), .73 (meeting), .79 (group), .84 (one–to–one), .87 (stranger), .85 (acquaintance), and .81 (friend), while the overall reliability was .93.

Measuring the degree centrality of social networking

In order to secure the source data needed for measuring individual degree centrality within the class, first the tool of “name–generator” was applied. The questions were comprised of asking to write down names of ‘the five most trusted friends’ and “the five friends who give knowledge the most frequently” in order. By providing different scores from five to one to people mentioned from the 1st to the 5th ranking, the researchers gave the value of degree centrality increased individually. By inputting the matrix data of the added degree centrality between “the person pointing out” and the “person pointed out” into NetMiner 3.4 (analytical tool for social networking), the vector of individual degree centrality was calculated.

Measuring the results of learning

The learning results were calculated based on the score added by two subordinate variables such as the individual learner’s understanding of the contents of the class and the excellence of the team assignment. The individual learner’s understanding of the contents of the class (first subordinate variable) was calculated through the total score of the mid–term and final exams that were comprised of ten short–answer questions. The team unit result (second subordinate variable) was calculated through the score of the report–type team project assignment. The title of this report was “Analysis of the current status of applying the instructional system design (ISD) within corporate training in Korea”. In order to write the report, the analytic activities and collecting field data such as interview or survey were required to be included.

For the individual exam score and the evaluation of the team assignment, the marking standard for each type was prepared in advance. After discussing the marking standard with two doctors of educational technology and the researchers, they were marked independently, and the average score of each marker’s score was used for the final score. The assignment score and the sum of midterm and final score were calculated on the scale of one hundred points.

This class condition can reveal the differences of the learners’ efforts put in to prepare the team assignment and midterm/final exams. Such differentiated efforts could work as a chaotic variable. In order to control this, the researchers set up the contribution level weight of both reference variables equally as 50:50, which was announced in the beginning of the semester.
Data analysis

In order for the descriptive statistical analysis of calculation and diagram for the trust network of each collaborative learning team and individual degree centrality within the network for sharing knowledge, the analysis of degree centrality of NetMiner 3.4 (Analyze → Centrality → Degree) was used. The path analysis of AMOS 7.0 was conducted to analyze the causal relations between the variables of the study model. As a method to estimate the parameter for hypothesis testing, the maximum likelihood method was applied while the significance level (α) was set up as .05 in accordance with custom. The study path model and the hypothesis related to each path are shown in (Figure 1) below.

![Figure 1. Path model and hypothesis](image)

Results

Basic analysis

Before the path analysis, the normality and multicollinearity of four variables of communication competence, degree centrality of trust network, degree centrality of knowledge sharing network and collaborative learning outcome of college students were tested.

In the examination results of kurtosis, degree of scattering, none of the four variables violated normality. However, there was a slight tendency to a more pronounced positive skew. This observation shows an important feature in common with Barabasi’s model of “scale free network”. According to Barabasi (2002), networks appear to have positively skewed distribution or power-law distribution in which most nodes have a relatively small number of ties, but a small number of nodes have a disproportionately large number of ties. Since path analysis is known as robust with respect to skewed distribution of predictor variables, and no significant outliers were detected, we decided to conduct significance test as planned.

Learning outcome was measured systematically as described in Method section. However, positive skew was observed in the criteria variable, the learning outcome. Since violation of normality in criterion variable is a threat to path analysis, we conducted statistical manipulation utilizing z-score conversion to alleviate the problem, which did not gain significant improvement. Thus, the p value reported should be conservatively interpreted.
Visual observation did not detect any significant outlier, which is major threat to the linear regression. Visual inspection of data plots and frequency distributions did not identify any outlier.

The descriptive statistics and the degree normality of each variable are summarized in Table 1 below.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average</th>
<th>Standard Deviation</th>
<th>Kurtosis</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>communication competence</td>
<td>76.11</td>
<td>14.20</td>
<td>-.02</td>
<td>-.29</td>
</tr>
<tr>
<td>degree centrality of trust network</td>
<td>18.52</td>
<td>11.02</td>
<td>-.28</td>
<td>.28</td>
</tr>
<tr>
<td>degree centrality of knowledge sharing network</td>
<td>6.12</td>
<td>4.20</td>
<td>.15</td>
<td>.06</td>
</tr>
<tr>
<td>learning outcome</td>
<td>91.21</td>
<td>6.81</td>
<td>-.01</td>
<td>-.25</td>
</tr>
</tbody>
</table>

Next, the tolerance analysis was conducted in order to decide the existence of multicollinearity between the variables. A tolerance of lower than .10 indicated the possibility of multicollinearity (Byrne, 2010). The tolerance result of .545–.681 indicated that the multicollinearity was not detected. Since the major statistical conditions were satisfied, each hypothesis was then tested by the path analysis as planned.

**Hypothesis Testing**

Correlation analysis was conducted before the path analysis. Because it can provide necessary information for path analysis by simple pair-wise analysis, it can provide the necessary information for the path analysis that considers the direct and indirect relationships among variables (Hoyle, 1995). The correlation matrix is presented in Table 2 below.

<table>
<thead>
<tr>
<th>No.</th>
<th>communication competence(1)</th>
<th>degree centrality of trust network (2)</th>
<th>degree centrality of knowledge sharing network (3)</th>
<th>learning outcome(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>.30&lt;.05</td>
<td>.49&lt;.01</td>
<td>.56&lt;.01</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>.55&lt;.01</td>
<td>.43&lt;.01</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.47&lt;.01</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

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

Next, the path analysis was conducted to test each of the five hypotheses. The results are presented in Table 3 below.

<table>
<thead>
<tr>
<th>No.</th>
<th>Hypothesis</th>
<th>standardized path coefficient (β)</th>
<th>Significance Level (p)</th>
<th>Acceptance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>communication competence → degree centrality of knowledge sharing network</td>
<td>.46</td>
<td>.00</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td>communication competence → degree centrality of trust network</td>
<td>.24</td>
<td>.13</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>degree centrality of trust network → degree centrality of knowledge sharing network</td>
<td>.41</td>
<td>.00</td>
<td>Y</td>
</tr>
<tr>
<td>4</td>
<td>degree centrality of knowledge sharing network → learning outcome</td>
<td>.55</td>
<td>.00</td>
<td>Y</td>
</tr>
</tbody>
</table>
Conclusions

Summary of findings

The four main study findings are summarized as follows.

First, communication competence, which is an individual attribute that enables the possessor to occupy the center of the knowledge sharing network, affected collaborative learning outcomes. What is remarkable about learning outcomes of knowledge sharing network is that they are highly explanatory ($\beta = .549, p = .000$). Therefore, if the student’s communication competence in class is high, the degree centrality of knowledge sharing network is increased and consequently collaborative learning outcomes are increased.

Second, the trust network is not determined only by individual communication competence. Trust is determined by predictability of others in certain circumstance (Misztal, 2001), and needs accumulated group or social experience. It doesn't necessarily mean that communication competence doesn’t have predictability on learning outcome (Jo, 2012), but that the effect of the variable is relatively low compared to other variables inserted into the research model. As shown in Table 2, when communication competence and degree centrality of trust network link are focused and the relations of other variables are ruled out, there is a meaningful relation between them($r=.302, p<.05$).

Third, centrality of knowledge network predicts collaborative learning outcomes with high predictability. The chance of mutual cooperation should be expanded to share knowledge with the whole class as well as the team for increasing the effectiveness of collaborative learning.

Fourth, degree centrality of trust network predicts centrality of knowledge network as result of precedent studies (e.g., Jo, 2008; Wasko & Faraj, 2005). This empirically confirmed again that mutual trust should be required to share the knowledge as shown in Social Capital Theory (Misztal, 2001; Nahapiet & Goshal, 1998; Wasko & Faraj, 2005).

Implications

Several implications and future research area related to collaborative learning can be proposed based on the study results.

First, to improve effect of collaborative learning, developing individual communication competence is required. Most of the existing studies for improving the effect of collaborative learning aim to improve the inside condition of collaborative learning by enhancing the team building, type and structure of entitlement, and extent of task distribution. The study results revealed that individual communication competence affects knowledge sharing in a collaborative learning context. Thus, when teams are built, individual communication competence should be considered in the short term, detailed characteristics of communication competence-related knowledge sharing should be established, and the development needs to be studied.

Second, developing instructional strategies of various collaborative learning styles is needed for preserving and expanding the trust of inter-learners, which is a precious network property. However, the problem is that universities are expanding individual competition for good grades because of the unemployment crisis these days, especially in Korea. The spread of this phenomenon of individualism to universities may reduce the meaning and role of collaborative learning and may harm trust network among co-learners. According to Druskat and Kayes (2000), if the team is built by learners’ autonomy, the team tends to build a familiarity of learners than their diversity. The more students feel time pressure and a team competitive sense, the more this tendency is increased (Druskat & Pescosolido, 2002). However if the collaborative learning team members know each other well, and have little regard for the learning of others, getting a grade is effective in the short term but expanding learners’ knowledge is restrained in the long term (Mehra, et al., 2001). Thus, instructors need to develop team building deeply for effective collaborative learning (Druskat & Kayes, 2000). Upon study completion, the researchers interviewed some of the participants in the group and confirmed the reasonableness of this argument. The participants stated that if they make collaborative learning team autonomically, they tend not to accept new members who are unfamiliar because accepting unfamiliar members is too risky to maximize the team assignment grade in the short term. This team
building tendency is safe in the short term but the width of the student relationships is restrained in the long term; furthermore, the improvement in the trust network may be reduced.

**Restriction and follow-up**

Due to the following study limitations, extreme care should be taken in interpreting the results.

First, all the participations were women since this research was conducted in women’s university. According to a previous study, women respond to relation (or network) sensitively (Palonen et al, 2000). The types of communication are gender specific (Jeong, 2005). Therefore, these study results cannot be generalized. Follow-up study needs to be conducted with a mixed group of participants.

Second, the communication competence was self-reported. From the study results, recognizing competence for oneself is classified with real competence notionally. Therefore, the results of measuring real communication competence may be different from those of this study according to performance assessment. Thus, in following studies, use of performance-based measures of communication competence is recommended.

Third, an attempt was made to analyze the structure of social and contents of interaction, which is a delimitation of quantitative social network analysis method, i.e., quantitative analysis. Especially, contents and meaning of shared language and narrative within the collaborative learning team should be identified and be analyzed poly-synthetically with the quantitative index of network analysis.

Fourth, we did not investigate whether a jigsaw session or the activity of the whole class contributes to the interdependence of the teams. In line with the previous study recommendations, further qualitative observation and rich analysis is needed in order to identify what collaborative learning strategy is necessary for expanding the trust network beyond the individual or team.

Fifth, the present study analyzed the prediction model. For more practical value, the model should be extended to include the effects of interventions for the enhancement of the learning outcome.

**Acknowledgements**

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2013S1A5A2A03044140).

**References**


Design, Customization and Implementation of Energy Simulation with 5E Model in Elementary Classroom

Sze Yee Lye1*, Loo Kang Wee1, Yao Chie Kwek2, Suriati Abas3 and Lee Yong Tay2
1Educational Technology Division, Ministry of Education, Singapore // 2Beacon Primary School, Singapore //
3Indiana University, United States of America // lye_sze_yee@moe.edu.sg // wee_loo_kang@moe.edu.sg //
	kwek_yao_chie@moe.edu.sg // sabas@indiana.edu // tay_lee_yong@moe.edu.sg
*Corresponding author
(Submitted April 20, 2013; Revised July 27, 2013; Accepted December 23, 2013)

ABSTRACT
Science simulations are popular among educators as such simulations afford for multiple visual representation and interactivity. Despite the popularity and abundance on the internet, our literature review suggested little research has been conducted on the use of simulation in elementary school. Thus, an exploratory pilot case study was conducted to address this research gap. In this study, an open source energy simulation was remixed for use in elementary school targeted at the Grade 4 & 5 students as an after-school enrichment program. We proposed 3 stages: design, customization and implementation, to provide useful insights with the aim to allow other educators to conduct their own remixed simulation lessons. The simulation design principles (e.g., learning outcomes and colour coding) with the corresponding TPACK construct that emerged from the design and customization stages were reported. Such simulation design principles would be useful to interested educators and researchers who wish to adapt and use simulation or teach others how to remix simulation. Data from the multiple sources (e.g., field observations, surveys, design notes and existing simulations) indicated that students enjoyed learning with the remixed energy simulation.

Keywords
Science simulation, Elementary school, Open source physics, Energy, TPACK

Introduction

Science simulations are best described as “software programs that allow students to explore complex interactions among dynamic variables that model real-life situations” (Park, Lee, & Kim, 2009, p. 649). Like Rutten, van Joolingen, and van der Veen (2012), we view interactivity to be an essential feature of science simulations and this is what makes them distinct from non-interactive based animations. Simulations are capable of accepting inputs and presenting the computational results in multiple representations like graphs or tables. Such interactivity and multiple representations afforded by simulation lends itself as a handy tool for inquiry-based learning, a common approach adopted by science educators (S. Chen, 2010). Moreover, a wealth of such simulations is freely available and accessible online (S. Chen, 2010). With their educational usefulness and availability, such simulations, inevitably, are becoming popular among Science educators.

Despite the pervasive use of simulations in schools and considerable research reporting on older students (Y. L. Chen, Hong, Sung, & Chang, 2011; Lamb & Annetta, 2013; Rutten et al., 2012; Scalise et al., 2011), there is paucity of research regarding the use of simulations in elementary school settings (Smetana & Bell, 2012). To address this gap, we report a study that explores the design, customization and implementation of energy simulation (Gallis, 2010) in an elementary classroom. We modified this open source energy simulation developed by Esquembre (2012). Taking on the dual role of a designer and teacher, we modified the simulation for use to suit the elementary classroom and also implemented the lesson the based on the BCSE 5E model (Bybee et al., 2006).

Literature review

Inquiry learning with science simulation

Inquiry learning has been an essential part of Science education throughout the world. Students are situated in an inductive learning mode, in which deductions on science principles are made based on their experience with the
students are learning by doing and are no longer just the passive receivers of
question formulations can present the data in table or graphical form
more, the simulation offers (structured inquiry), teacher guided
instructional goal
achievement of
does not contribute to the
Extraneous Processing
Reducing Extraneous
Processing.

Science simulations are probably suitable for these inquiry-based activities due to its capability of displaying
multiple representations and its interactivity (Chen, Wu, & Jen, 2013). In using the simulations to conduct
investigations, the students are learning together with technology (Jonassen, Howland, Marra, & Crismond, 2008).
Using simulation as their partner in learning, these simulations can present the data in table or graphical form
automatically. Hence, the students’ cognitive resources are freed up so that they can now focus on higher-order
thinking processes (e.g., analyzing the results and designing investigations). Furthermore, the simulation offers
multiple representations (e.g., word, pictures, diagrams, graphs and table of values) of the same or related concepts.
These multiple representations complement one another and students can thus now “integrate information from the
various representations to achieve insights that would otherwise be difficult to achieve with only a single
representation” (Wong, Sng, Ng, & Wee, 2011, p. 178).

With a deeper understanding of the concept, the students can now be in a better position to respond to questions, to
evaluate and communicate their findings. In using the science simulations, students are allowed to manipulate the
experimental variables to test their hypothesis. This allows for “genuine interactivity in terms of active learner
contribution and engagement” (Hennessy, Deanev, & Ruthven, 2006, p. 702). With this interactivity feature, students
can now make use of the simulations and design their own investigation; by varying different parameters and collect
their data accordingly to test their hypothesis.

With these affordances, students can now be actively engaged in inquiry-based activities (e.g., data collection and
presentation, changing variables and hypothesis testing). Simulations are highly dynamic and have the potential to
enhance students’ learning gains (Chen & Howard, 2010; Lindgren & Schwartz, 2009; Scalise et al., 2011). By
experimenting with the simulations, the students are learning by doing and are no longer just the passive receivers of
information in a typical teacher-centered class setting. Such student-centered learning with technology is, of course,
the much preferred approach for the contemporary researchers (Jonassen et al., 2008; S.-H. Liu, 2011).

Reducing cognitive load in science simulations

Cognitive load refers to the “strain that is put on working memory by the processing requirements of learning task”
(Driscoll, 2005, p. 136). We believe well designed simulation is a key to supporting learning as students can focus on
the learning task and not struggle with simulation. For example, too little information represented does little to aid
learning (Blake & Scanlon, 2007; Chen et al., 2011; van der Meij & de Jong, 2011) while too much information may
result in too high a cognitive load to integrate information from the various representations (Sweller, 2005). Thus,
in a well-designed simulation, the cognitive load should be minimized (Scalise, et al., 2011) as eliminating features
“that are not necessary for learning will help students to focus on the learning processes that matter” (de Jong, 2010, p.
109). Here, we present reputed educational psychologist, Mayer (2008) 10 evidence-based principles to reduce
cognitive load (See Table 1).

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reducing Extraneous Processing.</td>
<td>Coherence</td>
<td>Remove extra information which is not needed achieve instructional goal</td>
</tr>
<tr>
<td>Extraneous Processing does not contribute to the achievement of instructional goal</td>
<td>Signaling</td>
<td>Highlight essential material</td>
</tr>
<tr>
<td></td>
<td>Redundancy</td>
<td>Do not add text to narration</td>
</tr>
<tr>
<td></td>
<td>Spatial contiguity</td>
<td>Place related representation (i.e., words, diagrams) close to one another</td>
</tr>
<tr>
<td></td>
<td>Temporal contiguity</td>
<td>Present corresponding narration and animation concurrently</td>
</tr>
</tbody>
</table>
Managing Essential Processing

Essential Processing is required for the achievement of instructional goal.

Segmenting
Learner can control pace of animation

Pretraining
Provide pretraining in the name, location, and characteristics of key components

Modality
Present words as spoken text rather than printed text.

Fostering Generative Processing

Generative processing involves organization and integrating material

Multimedia
Use words and pictures rather than words alone.

Personalization
Use words in conversational style rather than formal style.

Other than the principles suggested by Mayer (2008), we can also consider other ways to optimize working memory. According to a study by Özcelik, Karakus, Kursun and Cagiltay (2009), colour coding increased both retention and transfer performance. Such enhanced learning can be attributed to the ease of locating the related information (e.g., graphic or text) and hence leading to a better integration of the multimedia material.

Research gap in elementary science simulation

There is no doubt that both researchers and educators are interested in using simulations for classroom teaching. For example, high school students used electricity and collision cart simulations in the studies, of Liu & Su (2011) and Wee (2012) respectively, while researchers were investigating the effect of instructional support on learning with simulation for eighth graders (Eckhardt, Uhrahne, Conrad, & Harms, 2013). However, there is little discussion about the use of simulations in elementary school settings. We conducted searches in the database Web of Science using descriptors "simulation" and "elementary school" and “science” for articles published since 2008 on April 19 2013. The search yielded only three results (Jaakkola & Nurmi, 2008; Jaakkola, Nurmi, & Veermans, 2011; Unlu & Dokme, 2011). Similar conclusion was also reported in the study of Smetana and Bell (2012) who examined 61 empirical studies in ERIC and Education Fulltext databases. Only two (3.3%) studies were conducted in elementary school setting. It is evident from these findings that there is indeed a paucity of research done in elementary school setting.

Energy simulation

The conservation law of energy states that in any closed system (including the Universe), the total quantity of energy remains fixed shares the same big idea about fundamental nature such as conservation of momentum in a closed system, will always been both an universal and relatively difficult (Papadouris & Constantinou, 2011; Stylianidou, 2002). This can be due to the abstract nature of the topic. We postulate that traditional pen paper teaching methods does not allow primary school students to apply this concept in other related areas like energy change in chemical reaction in their later schooling stage (Nordine, Krajcik, & Fortus, 2011). Therefore, we propose using well designed simulation with high interactivity and reduced cognitive load could afford for deeper learning as students can become “active agent in the process of knowledge acquisition” (Rutten et al., 2012, p. 137).

A search for energy-related simulations was conducted in the 5 popular simulations sites (NTNUJAVA Virtual Physics Laboratory, Java Applets on Physics, PhET Project, MyPhysicsLab and Interactive Physics and Math with Java) as suggested by Chen (2010) in her review on free simulation-based virtual laboratories. There are, of course, good energy simulations designed but they are not customized for use in elementary school (e.g., with extra information like acceleration due to free fall).

Open source physics simulation

Open source physics simulations are characterized by access to the source codes. These codes are modified using the free Easy Java Simulation (EJS) available at http://www.um.es/fem/EjsWiki/Main/Download (see).
The move towards open source simulations can be viewed as part of the increasingly popular movement towards open content like the massive open online course (DeSilets, 2013) and open textbook (Wiley, Hilton, Ellington, & Hall, 2012). Open Content is one of the emerging technology trend highlighted in the reputed 2013 Horizon Report (Johnson et al., 2013). They described open content as “materials that are freely copyable, freely remixable, and free of barriers to access, sharing, and educational use” (p. 7). This is very much similar to the spirit of open source physics in which the newly created or remixed simulations are shared with the community.

In this paper, we took an existing roller coaster simulation Open Source Roller Coaster (Gallis, 2010) and infused reducing cognitive loading principles using EJS. In the spirit of open content, we released our primary school version here http://www.phy.ntnu.edu.tw/ntnujava/index.php?topic=2383.0. Such enhancement to the existing energy simulation is, of course, much more efficient than creating a completely new simulation. Moreover, like the other open source physics simulations, the energy simulation is built by physics professors modeled using Physics equations. These simulations are more realistic thus aid the understanding of scientific concepts better Blake and Scanlon (Blake & Scanlon, 2007).

TPACK and open source energy simulation

Koehler and Mishra (2005) first mooted technological pedagogical content knowledge (TPACK) which extends Shulman’s conceptual model (1987) of Pedagogy Content Knowledge (PCK). TPACK refers to “refers to the synthesized form of knowledge for the purpose of integrating ICT/educational technology into classroom teaching and learning” (Chai, Koh, & Tsai, 2013, p. 32). As a framework, the constructs of TPACK are technological knowledge (TK), pedagogy knowledge (PK), and content knowledge (CK), Pedagogy Content Knowledge (PCK), Technology Content Knowledge (TCK), Technology Pedagogy Content (TPK) and Technology Pedagogy Content Knowledge (TPACK). The TPACK framework can be used in the design of teacher education program for designing technology-based lessons (Chai et al., 2013).

One possible strategy to develop teachers’ capability in infusing technology in classroom teaching is through the design of the educational technology to be used in authentic situations (Chai, Koh, Tsai, & Tan, 2011; Koehler & Mishra, 2005; Voogt, Fisser, Pareja Roblin, Tondeur, & van Braak, 2013). In a similar manner, designing open source simulation allows the teachers to experience the rich interplay of the relationships between technology,
pedagogy and content (Koehler, Mishra, & Yahya, 2007). In the use of open source simulations, they can even have a richer experience as they are now involved in modifying the simulation to suit their needs. In such modification, teachers can deepen their TPACK as the design can be seen as an “a process of weaving together components of technology, content, and pedagogy” (Koehler & Mishra, 2005, p. 135). They need to integrate the technology knowledge (what EJS can do) and pedagogical content knowledge (i.e., how to teach energy concept). TPACK, as knowledge, in this case, can transform teaching of science as the teacher is now able to customize the simulations to meet their specific pedagogical needs. Such customization by teachers is rarely published as the codes are not easily available or there is not relatively simple to use simulation author toolkits (like EJS or Molecular Workbench).

In remixing the simulation, the teachers are taking on the dual role of both the teacher and the designer. The dual role is advantageous, as the intention of the designer and the teacher will definitely be aligned. Teachers know the learning needs of the students in class and are clearly in the best position to design the resources to best fit their classroom practice (Hjalmarson & Diefes-Dux, 2008). Such tailored-made simulation is aligned to the students’ need and intended learning outcomes and thus are more likely to be adopted by teachers (Moizer, Lean, Towler, & Abbey, 2009). As the teachers are involved in designing simulation, they were, of course, aware of the affordances of the simulations and thus were able to better utilize them in the class. With holistic view of the relationship and the overlapping of the technology, pedagogy and content knowledge, they were now in a better position to use the simulation meaningfully in class.

**Stages in using open source energy simulation**

In this study, we proposed classifying the use of open source simulation into 3 stages: design, customization and implementation. Such classification can possibly guide the interested educators and researchers in the use of open source simulation. In the design stage, we decided on the possible features to be added. This was a paper design exercise before the simulation was actually changed using on EJS. Based on features, the simulation that was close to what was required was chosen. This would ensure that the open source simulation could be customized in the shortest possible time. In this case, the Open Source Roller Coaster (Gallis, 2010) was selected. Moreover, students were more likely to connect the simulation experience to the roller coaster real life experience.

The customized simulation was implemented in an enrichment lesson. 35 Grade 4 and Grade 5 mixed-ability students took part in this study. With no prior knowledge on energy, the lesson served as an introduction to energy. The objective of the lesson was to investigate energy conversion (in this case, it is the conversion of Gravitational Potential Energy to Kinetic Energy and vice versa). Energy simulation was used to support inquiry-based activities (i.e., like investigate the relationship between height and potential energy). The simulation can be downloaded at: https://dl.dropbox.com/u/24511248/JaveEJS/ejs_RollerCoasterV4.jar.

A guided inquiry-based lesson with the instructional guidance (i.e., reinforcing the energy concept and introducing simulation) was adopted rather than the pure discovery approach (Eckhardt et al., 2013). Learning is likely to be best supported by “instructional guidance rather than pure discovery, and curricular focus rather than unstructured exploration” (Mayer, 2004, p. 14). Similarly, in the study of simulations on middle school students, Chen and Howard (2010) also suggested that appropriate instructional guidance could enhance students’ learning experience. This study was conducted over two 2.5-hr lessons using the popular constructivist-based BCSE 5E Model (Bybee et al., 2006). See Table 4 for details.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Description</th>
<th>Activities Implemented</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engage</td>
<td>Elicit their prior understanding</td>
<td>You Tube Videos</td>
</tr>
<tr>
<td>Explore</td>
<td>Develop their understanding</td>
<td>Exploring the Energy Simulation</td>
</tr>
<tr>
<td>Explain</td>
<td>Demonstrate their understanding</td>
<td>Google Form Submission and Inquiry-based Worksheet</td>
</tr>
<tr>
<td>Elaborate</td>
<td>Deepen their understanding</td>
<td>Roller Coaster (Figure 9) to challenge the students’ perceived perception on energy</td>
</tr>
<tr>
<td>Evaluate</td>
<td>Assess their understanding</td>
<td>Change the settings of the Roller Coaster Simulation or Pendulum Simulation (Figure 2.) to record the highest kinetic energy</td>
</tr>
</tbody>
</table>
Purpose of the study

The purpose of this exploratory pilot study was to document how we designed, customized and implemented the simulation in an elementary class. It sought to provide evidences of increased interactive engagement the use of simulation both from the teachers and the students’ perspectives. This study would be of interest to both teachers and researchers who want to adapt and use simulations or teach others how to modify the simulation. The study was guided by the following questions:

- What are the generic principles and its corresponding TPACK construct that guides the design and customization of simulation?
- What are the students’ learning experiences with simulation?

Research design

A case study approach was adopted so as to seek a deeper understanding on how the open source energy simulation was adapted and used in class. In this study, the case represented the stages of design, customization and implementation of the energy simulation in an elementary school in Singapore.

To develop converging lines of inquiry, multiple sources of evidence were collected (Yin, 2008). Furthermore, findings derived and triangulated from the different data source are likely to be deemed more reliable and convincing. More details on data source and analysis and the corresponding process can be found in Table 3. Both directed content and conventional content analysis were used for qualitative data (Hsieh & Shannon, 2005). For example, predetermined categories derived from the literature review on principles to reduce cognitive load served as a guide in analyzing the simulations during the customization stage.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Data Source</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>Design Notes</td>
<td>Conventional content</td>
</tr>
<tr>
<td></td>
<td>Energy Simulation</td>
<td>analysis emerging categories)</td>
</tr>
<tr>
<td></td>
<td>Other Simulations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other research articles</td>
<td></td>
</tr>
<tr>
<td>Customization</td>
<td>Energy Simulation</td>
<td>Directed content analysis</td>
</tr>
</tbody>
</table>
Other Simulations

Implementation
- Field notes on lesson observations
- Students’ Google form submission (Open ended question)
- Focus Group Discussion with 12 students (from different abilities)
- Students Perception survey (Comments)
- Students Perception survey (5-point Likert Item)
- Students’ Worksheet

Conventional content
- Descriptive Statistics

Results and discussion

PK in design stage

The design features added in the simulation was noted (see Table 4 and Figure 3). They were further studied to surface the generic underlying principles and distill to the generic TPACK construct that could help guide the design process of other similar simulation. In this case, it is the pedagogy knowledge (PK) which is the knowledge “to teach a subject matter without reference towards content” (Chai et al., 2013, p. 33).

Table 4. Design

<table>
<thead>
<tr>
<th>Features</th>
<th>Principles</th>
<th>TPACK Construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Box about highest point and lowest point</td>
<td>Learning outcome</td>
<td>PK</td>
</tr>
<tr>
<td>Relationship between height and gravitational potential energy, speed and kinetic energy</td>
<td>Learning outcome</td>
<td>PK</td>
</tr>
<tr>
<td>Display of Graph and Values</td>
<td>Inquiry-based activities</td>
<td>PK</td>
</tr>
<tr>
<td>Controls for variables (like starting height and the mass)</td>
<td>Inquiry-based activities</td>
<td>PK</td>
</tr>
<tr>
<td>Added in Circular Loop</td>
<td>Teachers’ feedback</td>
<td>PK</td>
</tr>
<tr>
<td>Display of maximum and minimum height on the graphic</td>
<td>Teachers’ feedback</td>
<td>PK</td>
</tr>
<tr>
<td>Display of real GPE and KE</td>
<td>Fidelity towards actual phenomenon:</td>
<td>PK</td>
</tr>
<tr>
<td>Newly designed roller coaster</td>
<td>Students’ misconception</td>
<td>PK</td>
</tr>
</tbody>
</table>

From the analysis, it was evident that pedagogical knowledge guided the design as it was our intention to make the simulations educationally sound. Our other simulations, design notes were also examined for the principles emerging
from the analysis of the current simulation. This helped to triangulate the findings. The design principles were described below:

- Learning outcomes: The simulation was modified such that the learning outcomes (like information about highest point and lowest point) can be made more explicit from the simulation. It will aim to persuade and convince the students of the intended concept based on their observation. This is hardly surprising as this is the primary purpose of the customization is to suit the localized setting (a Singapore elementary school). This factor was evident in the other collision cart (Wee, 2012) in which the momentum equation was added to the existing simulation (Figure 4).

- Inquiry based activities: The simulation was further modified so that it could support inquiry-based activities (e.g., investigation, testing their hypothesis and evaluation of their findings). This was achieved by providing graph display and table of values, or given the students to control the variables like starting height. Similarly in the friction simulation, students could control the mass and the applied force (Figure 5) while values were displayed in collision cart simulation (Figure 4).

- Feedback from other teachers: Feedback was solicited from other teachers regarding our first version of the design via an online survey. We acted upon the feedback and further enhanced the simulation. Such practice was adopted in other simulations too, as reflected by the designer notes for the Ripple Tank Simulation and Cooling/Heating Simulation (Figure 6). Such feedback was based on pedagogical reasoning so that students were able to learn better with simulations.
Fidelity towards actual phenomenon: The simulations must be of correct level of representation (not too complex or simplified for the targeted audience) to bring about a meaningful and deep learning experience for the students (Blake & Scanlon, 2007). Open Source Physics simulations are modelled closely after the real-life phenomenon as they were based on Physics concepts. Similarly, in our modification of the energy simulations, we would base our codes on the Physics concepts. The gravitational potential energy (GPE) changed accordingly to height according the GPE equation (Figure 7). For the collision cart simulation, Newton’s 3rd Law was clearly evident in the force graph (Figure 8).

**Figure 6. Teacher’s feedback**

**Figure 7. Energy law in roller coaster simulation**

**Figure 8. Newton third law in collision cart**
Students’ misconceptions: Based on our collective teaching experience, we surfaced the possible misconceptions that students were likely to have. They would usually assume that the object can only reach the same height as the initial position. Another roller coaster (with the initial position lower than the highest position) was created to challenge the students. They were to make the roller coaster reach the highest point. This allowed the students to experience cognitive dissonance so as to bring about a conceptual change (Chen, Pan, Sung, & Chang, 2013; Smetana & Bell, 2012).

**TPK in customization stage**

The energy simulation was examined using the principles for reducing cognitive load (see Reducing cognitive load in science simulations). Some principles are not applicable in the design of the simulation (i.e., narration-related principles). Clearly, such principles belong to the technology pedagogy knowledge (TPK) as we used the technology (what EJS can do) to make the simulation more pedagogically sound. TPK is the “knowledge of the existence and specifications of various technologies to enable teaching approaches without reference towards subject matter” (Chai et al., 2013, p. 33) and hence the similar principles can be adopted for customizing other simulations. To triangulate the findings, other simulations we customized were also inspected using the same lens. Please refer to Table 5 for more details.

**Table 5. Customization**

<table>
<thead>
<tr>
<th>Customization Principle</th>
<th>TPACK Construct</th>
<th>Energy Simulation</th>
<th>Other Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coherence</td>
<td>TPK</td>
<td>Acceleration due to gravity (g), normal reaction force (R), energy loss (k) are removed. Information does not contribute to the learning outcome (Figure 10)</td>
<td>Forces (e.g., normal reaction force) removed in the inclined ramp simulation for elementary school (Figure 10. Roller coaster simulation)</td>
</tr>
<tr>
<td>Signaling</td>
<td>TPK</td>
<td>Indicate the highest and lowest height in the graphic display (Figure 10).</td>
<td>Indicate the height of the ramp in the inclined ramp simulation (Figure 10. Roller coaster simulation)</td>
</tr>
</tbody>
</table>
Spatial contiguity  
TPK  
Indicate the highest and lowest point in the graphic display  
Display of velocity and height information in symbolic (length of arrow) and numerical form near the object.  
Modification of Total Energy bar such it is made up of KE and PE energy bars  
Change the orientation of KE bar such that it is moving downwards (Figure 10).

Colour Coding  
TPK  
Consistent colour scheme for concepts related to potential energy and kinetic energy.

Display of velocity information in symbolic (length of arrow) and numerical form near the object in the inclined ramp simulation. The length of the arrow depends on the velocity (Figure 10. Roller coaster simulation).

Consistent colour scheme in collision cart simulation (Figure 11. Inclined plane simulation)

---

**Figure 10.** Roller coaster simulation

---

**Figure 11.** Inclined plane simulation

---

Original simulation  
Customized simulation

Original simulation  
Customized simulation
Learning experience in implementation stage

The qualitative data (i.e., field notes, focus group discussion and the Google form submission) were examined to uncover the possible emerging categories for students' learning experience. From the analysis, the students had a positive learning experience as they found the features of simulation useful and were engaged in self-directed learning. Such conclusion is line with literature review on simulation as described in the earlier section on Inquiry learning with science. The findings were further supported by the survey which was on 5 – point Likert Scale and the students' worksheet. Out of the possible 35 students, 27 students completed the survey while 35 students' worksheets were collected.

Positive learning experience

From field observations, comments in the survey and focus group discussion, students were engaged in learning with the simulations. They had no trouble navigating and exploring the simulation. Most of them found the lesson “cool” and “fun” and used the term “play” to describe their learning experience.

As the roller coaster could be customized, they were investigating the energy relationship using different way. Most students were able to articulate the energy principles during field observation. This observation also concurred with the Google Form Submission. Students also showed their understanding of the concepts in the hard copy worksheets. Majority of the students (30) were able to state the energy conversion correctly. This positive learning experience also concurred with the perception survey in which the students enjoyed and found it easy to use and learned with the simulation (see Table 6).

<table>
<thead>
<tr>
<th>Table 6. Positive Learning experience</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I enjoy learning about science using simulation.</td>
<td>4.15</td>
<td>1.06</td>
</tr>
<tr>
<td>I am able to learn about energy on my own using the simulation.</td>
<td>3.74</td>
<td>1.11</td>
</tr>
<tr>
<td>I find it easy to use the simulation.</td>
<td>3.89</td>
<td>1.19</td>
</tr>
<tr>
<td>Reliability Test: Cronbach’s Alpha</td>
<td>0.886</td>
<td></td>
</tr>
</tbody>
</table>
Features

The features (e.g., colour scheme, bar graph, numerical information) conceived during the design and the customization stages helped the student in understanding the energy concepts (see Table 7). The survey findings were further supported by focus group discussion. The representative comments were:

- *The indication of height (blue line) helps me to see the relationship between GPE and height.*
- *KE depends on the speed.*
- *The yellow arrow on the simulation shows that when the speed is low (like 7.3 m/s), the KE is 26.3J. But when the arrow is longer (faster) (like 11.3m/s) the KE is 60.5J.*
- *The total energy is made up of KE and GPE. The bar at the sides shows the bar. The TE bar shows both KE and GPE.*

<table>
<thead>
<tr>
<th>Table 7. Design features</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>The colour scheme helps me to understand the energy better.</td>
<td>4.41</td>
<td>0.97</td>
</tr>
<tr>
<td>The bar graphs helps me to understand the relationship between KE, GPE and Total energy.</td>
<td>4.33</td>
<td>1.04</td>
</tr>
<tr>
<td>The numerical information helps me to understand the relationship between KE and GPE</td>
<td>3.96</td>
<td>1.12</td>
</tr>
<tr>
<td>The indication of height (blue line) helps me to see the relationship between GPE and height.</td>
<td>4.18</td>
<td>1.08</td>
</tr>
<tr>
<td>The indication of speed (yellow line) helps me to see the relationship between KE and speed.</td>
<td>4.30</td>
<td>1.03</td>
</tr>
<tr>
<td>Reliability Test: Cronbach’s Alpha</td>
<td></td>
<td>0.932</td>
</tr>
</tbody>
</table>

However, the weaker student during focus group discussion indicated that they would need teachers’ guidance to point out the design features (e.g., consistent colour scheme) so that they could discover the energy relationship on their own.

Self-directed learning with simulation

Analysis from the various data sources indicated that students were engaged in self-directed learning with simulation. They enjoyed the interactivity and the experimentation as one student commented in the survey “I like the fact that we can experiment with the simulation by ourselves, and the simulation is very interactive.” The focus group discussion also revealed similar results with students enjoying the process of controlling the variables and evaluating their hypothesis on their own. Even if they did not fully comprehend what the teacher had explained, they could make use of the simulation to monitor their learning as one student pointed out “We don’t understand we can check ourselves”.

It was interesting to note that students were evaluating their hypothesis in a way we had not thought of. For example, some students used the starting speed and starting height to discover about the kinetic and gravitational potential energy:

- *If I adjust the height where the roller-coaster starts from, the amount of gravitational potential energy would be higher/lower as shown on the bar.*

From field observations, the students figured out how to customize the roller coaster on their own. They enjoyed the customization process as evident from the perception survey (M=4.24, SD=1.23). They even began experimenting with different variables (i.e., starting speed and starting height) so that the roller coaster could arrive at a higher point than the initial position. Most students were able to address their own alternative misconception before they were addressed formally. From Google Form submission, the students showed that they understood that for the object to have maximum kinetic energy; they had to “make its starting height very high”. Their experience was aptly put across by one student during the focus group discussion:

- *I enjoy customizing the stimulation as I change the ramps, height and speed according to my own fancy and changing the ramps helps me know how I can get from the start point to the end point quickly/slowly, and figure out why some ramps don’t allow me to arrive at the end point as quickly as the other ramps.*
Future work

Simulations can be used to provide differentiated instructions. With EJS, different versions of the same simulation can be created to cater to the different needs of the students. The expertise reversal effect can be considered to guide the design of the different versions of the simulation. In expertise reversal effect, students with different prior knowledge need different design features as features that “are effective with low-knowledge individuals can lose their effectiveness and even have negative consequences for more proficient learners” (Kalyuga, 2007, p. 510). During the focus group discussion, the weaker students found the graph difficult to understand as they “don’t understand what the graph is moving for.” They would prefer the table of values (Figure 13).

For further customization, we can consider not showing the graph in the simulation for this group of students so that they can just focus on the table of values. More interactivity can be provided for the better students. In study of Park et al., (2009), it was reported that students with higher prior knowledge performed better with high-interactivity simulations while students with low prior knowledge reacted more positively with low-interactivity simulations.

![Figure 13. Graph and table of value](image)

Conclusions

In this study, the energy simulation was adapted and implemented in an elementary school. We analyzed the findings from both the teachers’ and students’ perspectives. Due to the exploratory approach, the findings might not be generalizable but can be transferred to similar context. In this study, we had surfaced the generic principles with its corresponding TPACK construct in the design and customization stages. These would be useful to interested educators and researcher who wish to adapt and use simulation. For teacher-educators, the TPACK construct that emerged from the stages could provide some insights on designing lessons on customizing simulations. The findings also revealed that the students enjoyed learning with simulation. The positive learning experience could be due to deliberate features put in (e.g., the indication height) and the self-directed learning afforded by the simulation (interactivity and experimentation).

Acknowledgements

We wish to express our deepest gratitude to the following group of people or organizations for making this research possible:

- Francisco Esquembre, Fu-Kwun Hwang and Wolfgang Christian for their contribution to the open source physics simulation.
- Michael Gallis for creating and sharing his original simulation.
References


Development and Evaluation of Across-Unit Diagnostic Feedback Mechanism for Online Learning

Jian-Wei Lin¹, Yuan-Cheng Lai², Yu-Chin Szu³, Ching-Neng Lai⁴, Yuh-Shy Chuang¹ and Yen-Hung Chen⁵

¹Chien Hsin University, Taiwan // ²National Taiwan University of Science and Technology, Taiwan // ³St. John's University, Taiwan // ⁴Hsing Wu University, Taiwan // ⁵Institute for Information Industry, Taiwan // jwlin@uch.edu.tw // laiyc@cs.ntust.edu.tw // belinda@mail.sju.edu.tw // Sebastian.lai@gmail.com // yschuang@uch.edu.tw // pplong@gmail.com

*Corresponding author

(Submitted March 23, 2013; Revised August 13, 2013; Accepted September 8, 2013)

ABSTRACT

Solving well-structured problems often requires using considerable related concepts which are usually scattered and introduced throughout different learning units of a subject. In addition, poor learning of related concepts of preceding units may block the learning of subsequent units, and eventually leads to the inability to solve well-structured problems of a subject. Thus, this work proposes using across-unit diagnostic feedback, which can identify weak concepts not only within a unit but also in different units. Furthermore, the provided feedback can be used to recommend remedial learning paths for students, and inform the students the priority of the paths to understand which weak unit should be remedied first and which weak concepts within a unit should be remedied first. Students can refer to the instructions and use the provided corresponding remedial materials to conduct remedial learning in a systematic way. To discriminate the learning effect among various feedback types, this project will compare the proposed system, the Across-Unit Diagnostic Feedback System (AUDFS), with two other systems, the Single-Unit Diagnostic Feedback System (SUDFS) and the Traditional Feedback System (TFS). Experiment results show that the proposed system significantly enhanced learning achievement and the ability to solve well-structured problems for students. The mean student retention time of the proposed system is significantly higher than that of other systems, indicating that the proposed system enables sustained connection between students and the system. Additionally, positive correlations exist between student retention time of the proposed system and student post-test scores. Through a questionnaire and interviews, most students expressed positive attitudes to the proposed system.

Keywords

E-learning, Diagnostic feedback, Remedial learning, Online assessments

Introduction

Traditional education breaks an entire course into units and requires that learners focus on each unit independently, such that learners are often unable to perceive the big picture before all units are presented (Wang, Peng, Cheng, Zhou, and Liu, 2011). Conceptual understanding of a domain is often lacking in books and learning materials. Moreover, after the author teaching computer science courses a number of times, it was clear that most students prepared for their midterm or final exam at the last minute. That is, students rarely reviewed the material of previous units. Without prior knowledge of preceding units, students typically had difficulty in understanding subsequent units. The vicious cycle may propagate until the semester ends, resulting in reduced learning efficiency for the students and leaving students poorly prepared for midterms and final exams. Prior knowledge plays an important role in learning new concepts, such that students typically require prerequisite knowledge to learn a new concept (Ausubel, 1968). Lin, Lin, and Huang (2011) developed a test-based diagnostic system to assist instructors and students in diagnosing prior knowledge before new instruction is undertaken.

Moreover, the most common problems for students to solve in scientific subjects at universities are well-structured problems that feature logical, story- and rule-based problems with pre-defined steps and correct solutions (Jonassen, 1997). Arithmetic problems test the scientific conceptual understanding of students, requiring comprehension of a finite number of regular concepts, rules, and knowledge which may span a subject (Laxman, 2010). Without comprehension of preceding units, this can be a hurdle to the acquisition of follow-up knowledge and further reduce a student’s ability to solve well-structured problems.
Formative assessment plays an essential role in learning processes, through improving student learning efficiency by offering feedback rather than evaluation for course grades (Pachler, Daly, Mor, and Mellar, 2010). Students who have received feedback on their performance may take steps to remedy whatever weak concepts the assessment has exposed (Buchanan, 2000). Conducting one assessment for every unit throughout a course is very likely to stimulate students to practice and review preceding units, in turn benefiting learning in subsequent units.

However, learning feedback should show more information, instead of just simple results (correct/incorrect) or correct answers. That is, besides informing learners of their correctness, learners need to be told where their problem-solving process went wrong and coached from that point onward (Jonassen, 1997). An effective coaching method, “diagnostic feedback,” contributes to learning effectiveness. There are a handful of literatures which applied data mining and fuzzy theories to diagnose learning barricades or to developing adaptive learning for each individual (Chen and Bai, 2009; Chen, Hsieh, and Hsu, 2007; Lee, Lee, and Leu, 2009). Chen and Bai (2009) presented a method to diagnose learning barriers of learners based on fuzzy rules. Chen et al. (2007) presented a learning diagnosis system based on association rules for learning mathematics in an elementary school, trying to discover any learner misconceptions according to the responses of incorrect testing items during the learning processes. Lee et al. (2009) developed an intelligent concept diagnostic system based on a technique of data mining, enabling teachers to diagnose the learning barriers and misconception of learners with constructed concept maps.

However, most related studies provide diagnostic feedback for one assessment (i.e., one specific unit) instead of for an entire subject (i.e., a series of related assessments). For example, Chen, Hsieh, and Hsu (2007) designed a learning diagnostic system based on association rules for the course unit of “Fractions” in a mathematics course at an elementary school. Hwang (2003) developed a method to identify remedial learning paths for an assessment in the unit called “Number expressions and operators” in a science course at an elementary school. Chen (2011) proposed a personalized diagnostic system for an assessment in the unit called “Oriented-object programming design concept.” Feedback from only a small learning scope may hinder the acquisition of connected concepts across successive units and adversely affect the ability to solve well-structured problems. In other words, how to provide feedback for large learning scope (e.g., a series of online assessments) is still blurred at present.

This work provides diagnostic feedback across units in a subject and assigns one assessment to each unit. This work further developed the novel Across-Units Diagnostic Feedback System (AUDFS). The proposed diagnostic feedback scheme detects weak concepts within an assessment and across assessments. One assessment can solidify fundamental knowledge of the current unit, paving the way for the next unit. As arranged assessments in a course schedule increases, students can cement fundamental knowledge across units, gradually understanding correlations among units and the inner concepts of such units, and thus enhance their ability to solve well-structured problems. The recommended remedial learning path and material are regarded as scaffolding feedback for students, which motivates a learner’s interest and reduces frustration (Shute, 2008).

The remainder of this paper is organized as follows. The proposed approach is presented, after which the system development is elaborated. Then the educational evaluation is presented and finally the conclusion is delivered.

The proposed approach

An explicit representation of the structure of conceptual knowledge is constructed by capturing key knowledge concepts and their relationships in a visual format. This visual knowledge structure is a cognitive roadmap that facilitates the knowledge construction and high-level thinking of online learners (Wang et al., 2011). According to Zhuge and Li (2006), sequential and subtype relationships are the two relationships between knowledge concepts. A sequential relationship indicates that one concept should be learned before another, and is based on Ausubel’s assimilation theory (Ausubel, 1968), meaning that the role of prerequisite knowledge is essential for learning new concepts. A subtype relationship means that one child concept belongs to one parent concept, which conforms to Fisher, Gleitman, and Gleitman (1991) who states that such a relationship forms a tree structure, where the scope of the parent concept is broad, while that of a child concept is specific (Liu, Chen, and Chang, 2010). Hwang (2003) argued that a hierarchical structure offers an overall cognition of subject content from a top-down perspective.
Consequently, subject content can be presented in a tree structure, comprising of constituent units (e.g., chapters or sections) and key concepts within a unit. In reality, such a structure of conceptual knowledge is frequently used in a science textbook index. An assessment is assigned to each unit, which contains a number of key concepts. An assessment evaluates a student’s familiarity with key concepts in a unit. According to Wang (2011), the proper number of key concepts in an assessment should not exceed ten, so that learners may have manageable and measurable milestones and not feel overwhelmed. This work considers only the unit level and does not consider the sub-unit level. Figure 1(a) shows a simple example of a conceptual knowledge structure for a subject with three units (i.e., $U_1$, $U_2$, and $U_3$) in which assessment 1 ($A_1$) belongs to $U_1$ and contains six key concepts, $C_1$-$C_6$. Within each assessment, the sequential relationship is set for inner key concepts (Fig. 1(a)). Additionally, sequential relationships should be established among units (Fig. 1(b)).

Notably, a subject typically contains numerous concepts. Without considering the subtype relationship (i.e., considering only concepts without units), all concepts will be mixed up and placed in Level 1, directly below the root (i.e., subject) of the tree. Thus, one assessment will contain all subject concepts, which contradicts the position of Wang et al. (2011) who suggested that the number of concepts in one assessment should not exceed ten. Additionally, the structure of conceptual knowledge will become disorganized and unreadable. Using the subtype relationship, one can organize and present numerous concepts in a systematic manner.

\[
MC(1) = \begin{bmatrix}
X & 0 & 1 & 0 & 1 & 0 \\
0 & X & 0 & 1 & 0 & 1 \\
0 & 0 & X & 0 & 0 & 0 \\
0 & 0 & 0 & X & 0 & 0 \\
0 & 0 & 0 & 0 & X & 0 \\
0 & 0 & 0 & 0 & 0 & X
\end{bmatrix}_{6 \times 6}
\]

\[
MC(2) = \begin{bmatrix}
X & 0 & 1 & 1 & 0 \\
0 & X & 0 & 0 & 1 \\
0 & 0 & X & 0 & 1 \\
0 & 0 & 0 & X & 0 \\
0 & 0 & 0 & 0 & X
\end{bmatrix}_{4 \times 4}
\]

**Figure 1.** (a) Left part: A conceptual knowledge structure of a subject, which presents subject content in a tree structure and identifies sequential relationships among inner concepts in a unit (b) Right part: Identifying sequential relationships among units.

Instead of asking learners to construct the structure of conceptual knowledge alone, students can use expert knowledge to support their thinking and learning based on a solid foundation, reducing the likelihood of misconception (Wang et al., 2011). Experts have acquired a great deal of well-organized content knowledge, and their organization reflects a deep understanding of subject matters (Bransford, Brown, and Cocking, 2000). Thus, in this work, a teacher determines the subtype and sequential relationships to construct the structure of conceptual knowledge.
The proposed across-unit diagnostic algorithm

For facilitating algorithm computing in practice, sequential relationships can be represented in a matrix. $MU = [mu_{j,k}], 1 \leq j, k \leq M$ represents the sequential relationship between all units within a subject in which $M$ is the number of units in a subject. An element is binary, indicating the presence (1) or absence (0) of a relationship. The element in the diagonal is not defined. If $k$ is a succeeding unit of $j$ (i.e., $j$ is a preceding unit of $k$), $mu_{j,k} = 1$, else $mu_{j,k} = 0$. Similarly, $MC(i) = [mc(i)_{j,k}], 1 \leq j,k \leq N_i$ represents the sequential relationship between all concepts within the $i$-th unit in which $N_i$ is the number of concepts in the $i$-th unit. Fig.1 (b) shows an example of $MU$ and Fig.1 (a) shows the examples of $MC(1)$ and $MC(2)$.

$MU$: The matrix of the sequential relationships among all units within a subject
$MC(i)$: The matrix of the sequential relationships among all concepts within the $i$-th unit
$QU$: The queue of unit remedy path within a subject
$QC(i)$: The queue of concept remedy path within the $i$-th unit

```
GenerateQU (current_unit) {
    for (j = 1; j <= current_unit; j++) {
        if (j is a weak unit) and (Visited[j] is false) {
            Add (j) to QU;
            GenerateQC(j);
            for (k = j+1; k <= current_unit; k++) {
                if (k is a weak unit) and (mu_{j,k}=1) {
                    Add (k) to QU;
                    Visited[k] = true;
                    GenerateQC(k);
                }
            }
        }
    }
}

GenerateQC (i-th_unit) {
    for (j = 1; j <= last_concept; j++) {
        if (j is a weak concept) and (Visited[j] is false) {
            Add (j) to QC(i-th_unit);
            for (k = j+1; k <= last_concept; k++) {
                if (k is a weak unit) and (mc(i)_{j,k}=1) {
                    Add (k) to QC(i-th_unit);
                    Visited[k] = true;
                }
            }
        }
    }
}
```

Figure 2. Pseudo code of the across-unit diagnostic algorithm

Figure 2 shows the pseudo-code of the proposed across-unit diagnostic algorithm. The main aims of this algorithm are to identify weak concepts and units and further generate the across-unit remedial paths. $QU$ is the queue of the unit remedy path within a subject. After the algorithm is executed completely, sequentially outputting the elements in the $QU$, which has the property of First In First Out (FIFO), can acquire the unit remedy path and priority. Similarly, $QC(i)$ is the queue of concept remedy path within the $i$-th unit. Sequentially outputting the elements in the $QC(i)$ can acquire the concept remedy path and priority for the $i$-th unit.

Both inner pseudo-codes in the $GenerateQU()$ and $GenerateQC()$ functions are similar. The difference between these two functions is that $GenerateQU()$ is used to generate unit remedial paths between the first unit and the current testing unit while $GenerateQC()$ is used to generate concepts remedial paths within a given unit.
The GenerateQU() function checks each unit from the first unit to the current testing unit. If the \( j \)-th unit is weak, the function first adds the unit into \( QU \) and then calls the GenerateQC() function to generate concept remedial paths within the \( j \)-th unit. Furthermore, it also checks the \( j+1 \)-th element to the current unit-th element in \( j \)-th row of the \( MU \) (i.e., \( mu_{j+1}, mu_{j+2}, \ldots, mu_{current\_unit} \) ) to identify whether there are succeeding units of the \( j \)-th unit which are weak. If existing (e.g. the \( k \)-th unit), these succeeding weak units are sequentially added into \( QU \) and the GenerateQC() function is individually called to generate concept remedial paths for every succeeding weak unit. A succeeding weak unit (e.g. the \( k \)-th unit) which has been visited will never be visited again. The above procedure processes repeatedly until the last unit has been finished.

The GenerateQC() function checks each concept from the first to the last within a given unit, the \( i \)-th unit. If \( j \)-th concept is weak, the function first adds the concept into \( QC(i) \) and further checks the \( j+1 \)-th element to the last element in \( j \)-th row of the \( MC(i) \) (i.e., \( mc(i)_j, mc(i)_{j+1}, \ldots, mc(i)_{last\_concept} \) ) to identify whether there are existing succeeding concepts of the \( j \)-th concept which are weak. If existing (e.g. the \( k \)-th concept), these succeeding weak concepts are sequentially added into \( QC(i) \). A succeeding weak concept (e.g. the \( k \)-th concept) which has been visited will not be visited again afterward. The above procedure processes repeatedly until the last concept has been finished.

**The method of identifying weak concepts within a unit through an assessment**

The assessment result is the main factor in determining a learner’s performance (Zhuge and Li, 2006). Each question in an assessment may be associated with certain concepts (Lee et al., 2009; Zhuge and Li, 2006). The method to identify weak concepts within an assessment is based on the study by Hwang (2003), and has been slightly modified as follows. Herein, 1 represents a question associated with a concept, whereas 0 represents no association. Instructors define the association of questions with concepts through authoring interfaces. Suppose \( A_j \) (Assessment 1 in Fig. 1) consists of six questions (\( Q_1-Q_6 \)) and six concepts (\( C_1-C_6 \)) that are associated with the questions (Table 1).

Suppose a student incorrectly answers \( Q_2 \) and \( Q_5 \). The student can get the Correct Rate (CR) for concepts 1–6 of 100%, 66%, 100%, 66%, 80%, and 75%, respectively, where \( CR \) is calculated as \( CR = \frac{\text{Correct Amount}}{\text{Total Amount}} \).

<table>
<thead>
<tr>
<th>Questions</th>
<th>Concepts</th>
<th>C_1</th>
<th>C_2</th>
<th>C_3</th>
<th>C_4</th>
<th>C_5</th>
<th>C_6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q_1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Q_2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Q_3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Q_4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Q_5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Q_6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1. The association between questions and concepts in Assessment 1 (A_j) in Figure 1**

A concept \( CR \) of a student below the average is classified as a Low-score Cluster (\( LC \)); conversely, that above the average is classified as a High-score Cluster (\( HC \)). After all participants finish an assessment, the average \( CR \) for \( C_1-C_6 \) can be calculated. By comparing a concept \( CR \) with the average, a student can obtain his/her learning status (i.e., which cluster a concept is located) for all concepts within an assessment. A concept located in \( LC \) is deemed as a weak concept and needs to be remedied. If more than two weak concepts are exposed in an assessment, the student should focus on weak concepts that precede other weak concepts. With \( A_1 \) (Fig. 3) as an example, if \( C_2 \) and \( C_4 \) are the two weak concepts, the student is advised that \( C_2 \) has a highest priority to be remedied because it precedes \( C_4 \) in a sequential relationship.
Across-unit diagnostic feedback

Using the above-mentioned methods, one can acquire the learning status for any key concept within a unit when the corresponding assessments are completed. For instance, once $A_1$ is completed, the learning status of key concepts $C_1$–$C_6$ and unit 1 ($U_1$) can be obtained. This is the same scenario for $A_2$. If a student completed $A_1$ and $A_2$ (Fig. 2) and the scores in the $LC$ are marked in red, the suggested remedial learning paths should be addressed before entering the next unit, $U_3$. At this time, applying the proposed across-unit diagnostic algorithm can obtain the following feedback. First, the weak unit $U_1$ is added into $Q_1$ and then generates the $QC(1)$ for $U_1$ which contains two weak concepts: $C_2$ and $C_4$. Then, the weak unit $U_2$ is added into $Q_2$ and then generates the $QC(2)$ for $U_2$ which contains two weak concepts $C_3$ and $C_5$. Finally, sequentially outputting $QC(1)$ obtains $U_1$ and $U_2$, which means the unit remedy path is $U_1$ → $U_2$. Sequentially outputting $QC(2)$ obtains $C_2$ and $C_4$, which means the concept remedy path for $U_1$ is $C_2$ → $C_4$. Sequentially outputting $QC(3)$ obtains $C_3$ and $C_5$, which means the concept remedy path for $U_2$ is $C_2$ → $C_4$. More specifically, these results can be explained as follows. $U1$ should be remedied first because a poor $U_2$ likely results from a poor $U_1$ according to their sequential relationship. Thus, the first recommended concept remedy path is $C_2$ → $C_4$ within $U_1$ because $C_2$ precedes $C_4$. The second recommended concept remedy path is $C_3$ → $C_5$ within $U_2$ because $C_3$ precedes $C_5$.

![Scope in LC](image)

**Figure 3.** Identifying remedial learning paths and their priority from a top-down perspective

In addition to the textual information like above, the proposed system also provides the corresponding graphical information, as shown in Fig. 3. Miller and Miller (1999) explicated the focal tasks of a web-based learning system as follows: 1) present course content in a manner that hierarchically structures the sequence of information; and 2) obtain student feedback to insure accuracy of understanding. In our proposed framework, students can understand the structure of the entire course, perceive weaknesses, and conduct remedial learning from the bottom up. Once these weak key concepts are remedied by remedial learning, high-level units will be improved accordingly. Liu et al. (2010) stated that a learner must link and align the relationship of child concepts under the same parent concept to comprehend this parent concept; this knowledge construction process is bottom-up comprehension. Notably, as a course schedule is ongoing and assessments are gradually conducted, students can retrieve diagnostic feedback and do not have to wait until semester end. This allows students to identify relationships among concepts and units, generate semantic networks, and conduct high-level thinking for review and reflection (Wang et al., 2011).
System development

The operational procedure

Teachers must finish the following two tasks ahead of class. The first task is to establish the tree structure of a subject according to course content and build sequential relationships among units and within each unit (Fig. 1). The second task is to establish a Question Bank Database for each assessment. The assessments are based on textbook material and handouts. In this work, most questions are multiple choices. The operational procedure (Fig. 4) consists of the following steps.

- Step 1: Students learn a new unit in class.
- Step 2: Students take an assessment and obtain their results, including their answers, correct answers, and question restatements, to monitor their learning status for the unit.
- Step 3: Students check every incorrectly answered question. Meanwhile, students can retrieve diagnostic feedback.
- Step 4: If all units are finished, the operational procedure terminates; otherwise, return to Step 1 for a new unit and start again.

![Flowchart of the operational procedure](image)

Students can log in at will to review their current or historical assessments, receive diagnostic feedback, and complete an assessment that they did not complete previously. Providing students with across-unit diagnostic feedback may facilitate students with the acquisition of systematic knowledge, sustain the students’ connections to the system, and help students obtain learning skills to cope with well-structured problems. Therefore, using the proposed approach, students can take advantage of the system in the following scenarios. Scenario 1: A student who fails to attend an assessment session will receive an email reminder, notifying the user to complete the missed assessments to keep up with the class. Scenario 2: When a student who missed preceding assessments finishes conducting the current assessment, he/she will only receive the diagnostic feedback for the current unit. To obtain complete remedial support (i.e., across-unit diagnostic feedback), he/she has to complete the missed assessments. This scenario may engender his/her interest to complete the assessments to remedy weaknesses. These two scenarios demonstrate how the system raises a student’s awareness of his/her learning status, which may facilitate a student’s sustained connection to the system.
The developed system

The proposed system has two parts: a front-end User Platform and a back-end Administrator Platform. The front-end platform is available to students for assessments and diagnostic feedback, while the back-end platform is only available to administrators (teachers) for managing student-related activities.

Back-end platform

The Administrator Platform has several functions for managing course structure, sequential relationships, and assessments. Other functions are remote user management and statistical analysis. Some graphical user interfaces are shown in Fig. 5. The top-left window shows the construction of the course structure while the bottom-left window is to construct sequential relationships among units. The top-right window shows the assessment edition. By using the multimedia editor, embedded by the free software component FCKEditor (http://www.fckeditor.net), a teacher can edit multimedia questions by inputting text, images, tables, flash objects, and hyperlinks and further arrange the layout. The bottom-right window sets the association between questions and concepts within an assessment (Table 1).

Front-end platform

A student takes an assessment through the front-end platform. The assessment is generated by an administrator through the back-end platform. The student can review personal learning profiles from the assessment list and past results (Fig. 6).

Figure 7 shows snapshots of the across-unit diagnostic feedback that students can obtain. In the left window, the upper part shows the diagnostic result among units in a tree structure, while the lower part shows weak units and the remedial learning sequence for units. The top-right window shows weak concepts within a unit and the remedial learning sequence for inner concepts. Students can click a weak concept and receive the corresponding remedial material, as shown in the bottom-right window. The remedial material presented in a multimedia content format, including audio, video, PowerPoint presentations, and file downloads, and it is created using PowerCAM software (http://www.powercam.com.tw/). Students can select or replay the desired material through operational interfaces.
Educational evaluation

Objectives

This evaluation compares the proposed system, the AUDFS, with two other systems, the Single-Unit Diagnostic Feedback System (SUDFS) and Traditional Feedback System (TFS). Once an assessment is completed, all three systems give students performance results, including questions, answers, correct answers, and scores. Notably, students using the AUDFS can receive across-unit diagnostic feedback (Fig. 7). Students using the SUDFS only receive diagnostic feedback for one unit. Students using the TFS do not receive diagnostic feedback, but only information about whether their answer was correct. The three systems were used by three first-year classes at a university, with an approximate enrollment of 160 students. The first class, consisting of 52 students, used the
AUDFS. The second class, comprising 52 students, used the SUDFS. The third class, consisting of 55 students, used the TFS.

This work (1) compares the learning achievements of these classes, and (2) compares student retention time among classes and within the AUDFS class. Finally, a questionnaire was administered and 12 students were interviewed.

Research tools and procedure

The experiment course is a computer science course called “Database Theory and Application—Microsoft Access 2007.” The course goal is to teach students about databases and Microsoft Access 2007. The experiment subject has five units: “Relational Database Concept Introduction”; “Access Object Introduction”; “Entity-Relationship Diagram (ERD)”; “Building a Relational Access Table”; and “Query for Relational Tables.” The upper part (a) of Fig. 8 shows the sequential relationships among the five units within the experiment subject while the lower part (b) of Fig. 8 shows an example of the sequential relationships among the concepts within a unit, the Unit 3. In total, 12, 13, 15, 14, and 13 multiple-choice questions were generated for units 1–5, respectively, to establish question banks. The test content in each assessment primarily originates from teaching materials.

This work adopted a quasi-experimental design. The experiment treatment was 2 months, 2 hours weekly. Before the experiment started, all classes were trained and practiced using their designated system for 1 week. Students in the AUDFS class were advised to utilize its novel functions (Fig. 7), especially its capabilities to identify weaknesses in current and prior units and provide corresponding remedial materials. During the experiment, all classes adhered to the same teaching and assessment schedules. Upon completion of one unit, all classes took the same assessment of that unit. The posttest content came from teaching materials and assessments. Students who did not complete assessments had no historical records in their designated system. Before the posttest, they also had no historical records of assessments in their personal learning profile to review. This motivated students to attend each assessment.

To verify whether pretest scores of students in these classes differed significantly, a pretest of background knowledge was conducted before the evaluation. To assure pretest validity and reliability, two experts reviewed the pretest, which was then tested by 56 students. Inappropriate questions were removed according to their difficulty and discrimination levels, leaving 16 multiple-choice questions with a Cronbach’s α of 0.83.
The validity and reliability analyses of the posttest were the same as those of the pretest, leaving 22 questions with a Cronbach’s $\alpha$ of 0.82. Most questions in the posttest were well-structured problems; for instance, a question was: “How do you transfer an existing ERD of a small business to relational Access tables?” The question was followed by four options, of which only one was correct. To answer this question, one must have an understanding of connected concepts across several units, such as the concepts of “Access Objects” and “Table Object” in Unit 2, “Entity,” “Attribute,” and “Relationship” in Unit 3, and “Creating Tables,” “Assigning the Primary Key,” and “Establishing Relationships” in Unit 4. Thus, completing a posttest question required knowledge across different units.

**Data collection and analysis**

To analyze participant preferences, all systems recorded participant activities as logged data, including login time, activity types (i.e., reviewing or testing), source IP (i.e., locations such as home or school), and stay period (i.e., the time a visitor spends in the system). The SPSS program was used for statistical analysis.

One-way analysis of variance (ANOVA) was applied to examine the effects of these systems on each assessment (Table 2). No significant differences existed in mean scores for units 1 and 2 among the three classes. This is reasonable because units 1 and 2 are both preceding units (Fig. 8) of the subject. At this time, the function of the AUDFS is the same as that of the SUDFS, as the AUDFS can perform for only one unit. For Unit 3, although the result does not reach significance ($F = 1.39$, $p > .05$), the mean score of the AUDFS class is higher than that of the other two classes. At this time, the positive effect of the AUDFS starts to emerge. For Unit 4, the mean score of the AUDFS class is significantly higher than that of the TDS class. Additionally, the mean score of the AUDFS class is higher than that of the SUDFS class, even though the result is not significant. For Unit 5, the mean score of the AUDFS class is significantly higher than that of the SUDFS and TDS classes. This is because addressing Unit 5 demands a larger scope, requiring concepts from units 1–5, compared with Unit 4, which requires concepts from units 1–4. The AUDFS gradually showed its effectiveness in the latter units because understanding concepts in these latter units required more concepts than preceding units. The AUDFS identifies weak concepts within the current unit and within preceding units, further offering remedial material. These features benefit learners in solving well-structured problems, which require the comprehension of a series of concepts in different units. In contrast, the SUDFS identifies only weak concepts within one unit, which may limit learner capability to solve well-structured problems. The TFS is the weakest system because it does not provide feedback support.

**Table 2. Results of one-way ANOVA**

<table>
<thead>
<tr>
<th>Assessment #</th>
<th>Class</th>
<th>Mean</th>
<th>$SD$</th>
<th>$F$</th>
<th>Post Hoc$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit 1</td>
<td>AUDFS</td>
<td>37.34</td>
<td>17.87</td>
<td>1.10</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>SUDFS</td>
<td>33.38</td>
<td>15.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFS</td>
<td>36.93</td>
<td>12.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 2</td>
<td>AUDFS</td>
<td>40.07</td>
<td>20.56</td>
<td>0.79</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>SUDFS</td>
<td>40.51</td>
<td>17.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFS</td>
<td>44.24</td>
<td>20.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 3</td>
<td>AUDFS</td>
<td>52.58</td>
<td>21.18</td>
<td>1.39</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>SUDFS</td>
<td>48.31</td>
<td>19.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFS</td>
<td>45.88</td>
<td>20.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 4</td>
<td>AUDFS</td>
<td>54.89</td>
<td>19.65</td>
<td>3.77*</td>
<td>AUDFS &gt; TFS</td>
</tr>
<tr>
<td></td>
<td>SUDFS</td>
<td>49.02</td>
<td>18.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFS</td>
<td>44.13</td>
<td>28.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 5</td>
<td>AUDFS</td>
<td>54.42</td>
<td>19.19</td>
<td>3.88*</td>
<td>AUDFS &gt; SUDFS</td>
</tr>
<tr>
<td></td>
<td>SUDFS</td>
<td>45.33</td>
<td>20.18</td>
<td></td>
<td>AUDFS &gt; TFS</td>
</tr>
<tr>
<td></td>
<td>TFS</td>
<td>42.07</td>
<td>20.94</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Multiple comparisons: LSD.  
$^* p < .05$.  

148
Analysis of learning achievement among classes

The paired-samples t-test was applied to compare learning achievement among classes. All students in these three classes significantly improved on their posttest (Table 3). One-way ANOVA was further applied to the pretest to verify whether background knowledge of the three classes differed significantly. No significant difference existed in background knowledge among the three classes ($F = 1.25, p > .05$).

<table>
<thead>
<tr>
<th>Class</th>
<th>Test</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUDFS</td>
<td>Pretest</td>
<td>52</td>
<td>37.42</td>
<td>18.03</td>
<td>-16.73*</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>78.29</td>
<td>12.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUDFS</td>
<td>Pretest</td>
<td>52</td>
<td>32.94</td>
<td>14.95</td>
<td>-14.14*</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>70.08</td>
<td>17.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFS</td>
<td>Pretest</td>
<td>55</td>
<td>36.13</td>
<td>11.12</td>
<td>-9.77*</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>64.42</td>
<td>18.47</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$p < .05$.

One-way Analysis of Covariance (ANCOVA) was also applied to compare learning achievement among classes. The analysis regarded the experiment treatment as the independent variable, posttest score was the dependent variable, and pretest score was the covariate. Before analyzing covariance, homogeneity of regression coefficients was tested to examine whether intra-group homogeneity existed. The SPSS analysis demonstrates that the F value of regression coefficients was 1.16 ($p > .05$); thus, the hypothesis of homogeneity was accepted. Thus, covariance analysis was further conducted.

Posttest scores were adjusted by removing the influence of the pretest from posttest scores. Pretest scores had a significant effect on the posttest scores ($F = 8.42, p < .05$) (Table 4). Learning achievement of the three classes differed significantly ($F = 9.61, p < .05$), indicating that a great difference in learning achievement existed among these three classes.

(1) Students in the AUDFS class performed significantly better than students in the SUDFS class and TFS class. For a well-structured problem, an error in one step can carry over to subsequent steps, and, consequently, to the final solution (Corbalan, Paas, and Cuypers, 2010). That is, a mistake can propagate over an entire problem-solving process. The concepts required to solve a well-structured problem usually span different units. The SUDFS can only identify misconceptions within one unit; in contrast, the AUDFS can identify all misconceptions in different units as well as the source of the misconceptions, and provide recommended remedial paths and material. Students using the AUDFS can ideally obtain a deeper understanding, progressively comprehend the correlation across different units, and acquire a better ability to solve well-structured problems.

(2) Students using the SUDFS performed better than those using the TFS. Wang (2008) stated that web-based assessment would be more effective if it could provide strategies and feedback to learners. The TFS only provides the correct or incorrect answer, which is insufficient for overcoming learning barriers, limiting students’ problem-solving ability, and encouraging rote memorization.

Table 4. One-way ANCOVA on posttest scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Class</th>
<th>Mean$^a$</th>
<th>SD</th>
<th>$F$</th>
<th>Post Hoc$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>AUDFS</td>
<td>77.81</td>
<td>2.21</td>
<td>8.42*</td>
<td>N/A</td>
</tr>
<tr>
<td>Type of System</td>
<td>SUDFS</td>
<td>70.71</td>
<td>2.22</td>
<td>9.61*</td>
<td>AUDFS &gt; SUDFS *</td>
</tr>
<tr>
<td></td>
<td>TFS</td>
<td>64.26</td>
<td>2.15</td>
<td></td>
<td>SUDFS &gt; TFS *</td>
</tr>
</tbody>
</table>

$^a$ Covariates appearing in the model are evaluated at Pretest = 35.51.

$^b$ Adjustment for multiple comparisons: LSD (equivalent to no adjustments).

Analysis of retention time for the three classes

Retention time of each student, which was the total time a student spent on his/her system during the evaluation period, was computed. This value was calculated by accumulating the staying time after every login. Table 5 shows
that retention time result is significant \((F = 4.06, p > .05)\). The mean retention time in the AUDFS class (mean = 150.02) is significantly higher than that of the TFS class (mean = 93.72). This is likely because the AUDFS offers more learning assistance and enables a sustained connection between students and the system than the TFS. However, the retention time of the AUDFS class (mean = 150.02) is not significantly higher than that of the SUDFS class (mean = 118.15). This is because students using the SUDFS also feel that their system help their learning.

<table>
<thead>
<tr>
<th>Retention Time (in Minutes)</th>
<th>Class</th>
<th>Mean</th>
<th>SD</th>
<th>(F)</th>
<th>Post Hoc*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUDFS</td>
<td>150.02</td>
<td>124.56</td>
<td>4.06*</td>
<td>AUDFS &gt; TFS *</td>
</tr>
<tr>
<td></td>
<td>SUDFS</td>
<td>118.15</td>
<td>101.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFS</td>
<td>93.72</td>
<td>83.85</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( ^* p < .05. \)

In the AUDFS class, students whose posttest scores were above the average were allocated to the Posttest High-cluster (PostHC) group; otherwise, they were allocated to the Posttest Low-cluster (PostLC) group. The independent samples \(t\)-test was applied to comparison results (Table 6). For retention time, the mean of the PostHC group (mean = 150.02) is significantly higher than that of PostLC group (mean = 84.35) \((t = -2.40, p < .05)\). Moreover, the Pearson coefficient was generated to assess the strength of the correlation between the posttest scores and retention time. A significant positive correlation existed (correlation coefficient = 0.36, \(p < .01\)), showing that retention time is moderately conducive to learning achievement.

<table>
<thead>
<tr>
<th>Retention time (Minutes) within the AUDFS class</th>
<th>Group*</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PostHC</td>
<td>27</td>
<td>262.14</td>
<td>371.77</td>
<td>-2.40*</td>
<td></td>
</tr>
<tr>
<td>PostLC</td>
<td>25</td>
<td>84.35</td>
<td>74.46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Because two students’ posttest scores were the same, the two groups did not have the same number of people.

\( ^* p < .05. \)

**Questionnaire and interview**

To understand student satisfaction, a questionnaire with a Likert scale ranging from 5 for “strongly agree” to 1 for “strongly disagree” was given to all classes at study end. Among 52, 52, and 55 students in the AUDFS, SUDFS, and TFS classes, 49, 50, and 51 valid questionnaires were collected for data analysis. A questionnaire with all answers empty or all answers the same was deemed invalid. After completing the questionnaire, short interviews of 12 randomly selected students were conducted to elicit their feelings.

Table 7 and figure 9 show the questionnaire questions and results, respectively, revealing that all systems received positive feedback for most evaluated aspects. For questions 1–3, most students felt their designated system was stable and convenient. However, the remaining answers (question 4-10) show that the AUDFS class was more satisfied than the other two classes, likely because the AUDFS has more functions for learning assistance. The following focuses on questionnaire and interview results for the AUDFS class.

<table>
<thead>
<tr>
<th>#</th>
<th>Question Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The designated system provided a convenient environment ((e.g.,\ user-friendly interfaces)).</td>
</tr>
<tr>
<td>2</td>
<td>The designated system had good stability ((e.g.,\ quick response)).</td>
</tr>
<tr>
<td>3</td>
<td>You are satisfied with the operation of the designated system.</td>
</tr>
<tr>
<td>4</td>
<td>Assessments helped you identify your weaknesses.</td>
</tr>
<tr>
<td>5</td>
<td>The designated system assisted and reinforced your weaknesses.</td>
</tr>
<tr>
<td>6</td>
<td>The designated system benefited learning new concepts.</td>
</tr>
<tr>
<td>7</td>
<td>The designated system helped you understand the course scope and correlation among units.</td>
</tr>
<tr>
<td>8</td>
<td>You usually logged in to the designated system for review in your daily life.</td>
</tr>
<tr>
<td>9</td>
<td>You logged in to the designated system for review before the post-test.</td>
</tr>
<tr>
<td>10</td>
<td>The designated system significantly improved your learning.</td>
</tr>
</tbody>
</table>
Fig. 9. Questionnaire results among the three classes

Answers to questions 1–3 show that most students were satisfied with the AUDFS. For example, one interviewee stated, “The user interfaces were simple and friendly. As a novice, it was easy for me to catch on by first or second use.” Another interviewee expressed, “The system was pretty stable and convenient and I could review at school or home.”

Answers to Question 4 show that most students felt that the AUDFS is a practical auxiliary tool for identifying learning weaknesses. For example, one interviewee stated, “The system helped me identify where my weaknesses were.” Two interviewees said that they listened to material in class and the system told them which parts of unit they did not understand.

Answers to Question 5 show that the AUDFS identified weaknesses. Informing learners of their weaknesses and providing coaching are important (Jonassen, 1997). One interviewee stated, “For every weakness, I can instantly look at the corresponding PowerCAM remedial material and replay it if necessary. The remedial material was easy to understand because it was illustrated by video.” One interviewee said, “Although the system can spot my weak concepts in each unit, it was hard for me to figure out the better remedial path, especially when there were too many weak concepts or the learning scope was huge. Thanks to the AUDFS, I can follow the provided guide to strengthen my weaknesses step by step.”

Answers to questions 6 and 7 show that the AUDFS benefited learning new concepts, understanding the learning scope, and correlation among units. One interviewee commented, “Sometimes I tried to strengthen my weak concepts in the latter unit (e.g., Unit 5) by directly reading the teaching material of that unit, but I was still confused. However, the AUDFS unveiled which of my weaknesses in the preceding units should be remedied first. I followed the instructions and they worked!”

When investigating whether the AUDFS stimulated students to spend time in their daily life or before the posttest (questions 8 and 9), some said they reviewed in daily life, while some admitted that they reviewed only before the posttest. Although students welcomed the AUDFS, some suffered from excessive loads from other subjects, resulting in insufficient time for review. One interviewee stated, “I regularly logged in to the system to review after school. I cannot stand reviewing at the last moment (i.e., only before the posttest), because I have to prepare too many subjects at that time.” Another one stated, “I reviewed only before the posttest because this strategy gave me a deeper impression and better memory.” From answers to Question 8, one can see that students were willing to spend time on the system.

Question 10 obtained the highest score, directly reflecting the fact that the students in the AUDFS class felt that the overall system significantly improved learning.
Conclusions

This work provides diagnostic feedback for a large learning scope through a series of online assessments that detect weak concepts within an assessment and across assessments. The deployment of assessments is based on the conceptual knowledge structure for a subject. Through the recommended remedial learning path and material, students can solidify their foundations for concepts in the current unit and across other units, which can enhance the students’ ability to solve well-structured problems.

Experiment results show that the AUDFS is conducive to learning subsequent units. The mean post-test score in the AUDFS class is significantly higher than that in the SUDFS and TFS classes, indicating that the AUDFS significantly enhances student learning and ability to solve well-structured problems. The mean retention time in the AUDFS class is significantly higher than that of the TFS class, indicating the AUDFS enables a sustained connection between students and the system. Within the AUDFS class, mean retention time of the Post-HC is significantly higher than that of the Post-LC. Retention time and post-test score were positively correlated. The questionnaire and interview results reveal that most students using the AUDFS have positive attitudes toward all evaluated aspects.

The proposed system should fit such subjects as mathematics and programming courses for the following two reasons. First, these courses have fixed learning sequences and a clear concept structure, which can represent course content into a tree structure (Fig. 1). Second, typical questions in these subjects are well-structured problems, requiring across-unit concepts to solve. However, this assumption needs further verification.

As the online learning environment is characterized as autonomous, the ability of learners to engage in self-regulated behaviors is construed as a crucial factor for successful online learning (Barnard, Lan, To, Paton, & Lai, 2009). A self-regulated learner can play an active role in learning, setting task-oriented and proper goals, taking responsibility for their own learning, monitoring their own learning, and maintaining their own learning motivation (Heikkilä and Lonkab, 2006; Wang, 2011). This ability for self-regulation is important, because online learning requires learners to be more disciplined; they also require considerable persistence and determination (Shea and Bidjerano, 2012). Kauffman (2004) and Wang (2011) also asserted that the learners in an online learning environment must be highly self-regulated; otherwise their learning effectiveness may be low. Thus, in the future, this study will further verify whether the self-regulation level of a student significantly influences his/her learning behavior and learning achievement in the AUDFS.

References


Knowledge Sharing among University Students Facilitated with a Creative Commons Licensing Mechanism: A Case Study in a Programming Course

Chen-Chung Liu¹, Chia-Ching Lin*, Chun-Yi Chang² and Po-Yao Chao³

¹Graduate Institue of Network Learning Technology, National Central University, Taiwan // ²Department of Computer Science and Engineering, Yuan-Ze University, Taiwan // ³Department of Information Communication, Yuan-Ze University, Taiwan // ccliu@cl.ncu.edu.tw // cclin.edu@gmail.com // j1202748@gmail.com // poyaochao@saturn.yzu.edu.tw

*Corresponding author

(Submitted April 11, 2013; Revised September 16, 2013; Accepted December 6, 2013)

ABSTRACT

Creative Commons (CC) mechanism has been suggested as a potential means to foster a reliable environment for online knowledge sharing activity. This study investigates the role of the CC mechanism in supporting knowledge sharing among a group of university students studying programming from the perspectives of social cognitive and social capital theories. By gathering 40 university students’ feedbacks and their behaviors, this study found that, in terms of trust, sharing self-efficacy and outcome expectations, the students had a more positive perception of the CC-integrated platform than that of conventional online sharing. In addition, the students were more likely to favor the CC-integrated sharing platform in which they believed that individual identification and profit were more effectively sustained and protected in the process of knowledge sharing. The students’ activities on these platforms and their perception of knowledge sharing also revealed relationships that differed from those found in their performance of programming tasks on other online sharing platforms. This study suggests that the employment of the CC mechanism could be helpful in promoting a more positive perception of and willingness to engage in knowledge sharing, thus ensuring the effectiveness of peer communication and collaboration in stimulating programming performance within the Internet-based collaborative context.

Keywords

Knowledge sharing, Creative commons, Programming, University students

Introduction

To encourage knowledge sharing has become the key challenge and one of the important research issues in online learning contexts (Ma & Yuen, 2011). However, previous studies also indicated that most community participants tend to be unwilling to share their knowledge because of a social dilemma associated the sharing behaviors (e.g., Cabrela & Cabrela, 2002; Ridings, Gefen, & Arinze, 2006). Therefore, an increasing number of studies have devoted themselves to research into what interpersonal relationships and environmental support would facilitate online knowledge sharing in virtual communities based on social cognitive and social capital theory (Chou, 2010; Hsu, Ju, Yen, & Chang, 2007; Kimmerle et al., 2007; Zhou, 2008). The rationale behind these studies is that it may well be that a sense of insecurity toward sharing knowledge online may result in limited participatory engagement, as knowledge sharing activities bring about iterative social interaction with peers and involve distributing one’s intellectual property. Furthermore, previous studies indicate that the absence of workable norms to establish authorship and to trace the history of shared content creation may hinder the participant’s willingness to share knowledge with unknown users in online settings (Hsu et al., 2011; Kanawattanachai & Yoo, 2002). Ambiguous ownership of shared content may reduce student reliance on the engaged environment and in interpersonal relationships among peers, and in turn may impede the development of virtual communities.

Based upon the perspectives of social cognitive and capital theories, knowledge sharing could be regarded as a process of mutual exchange between individual cognition and inter-personal relationships in the social networking context. Taking consideration of both personal needs for and environmental influences on knowledge sharing would be helpful to the advancement of knowledge sharing activity. This issue could be addressed by adopting the Creative Commons (CC) licensing mechanism as it provides flexible protective licensing options for regulating how content creators and users publish, share and reuse online shared materials. Using CC principles, validating the originality of shared content may significantly relieve tension arising from ambiguous contributions to and ownership of intellectual property in knowledge sharing activities (Liu et al., 2013). The objective of this study thus is to design an
online learning environment to support programming learning in a formal education setting with the CC mechanism because cooperative programming activities such as referencing others’ works are helpful to students’ cognitive development of programming learning (Hwang et al., 2008). To examine the role of the CC mechanism in supporting students’ programming learning within the online learning environment, the following questions were posited and investigated:

- How do student knowledge sharing perceptions associated with programming study differ between online learning environments with and without CC?
- Are students more likely to share knowledge in an online learning environment with CC than one without CC?
- How do student knowledge sharing perceptions relate to their sharing/non-sharing intention?
- How do student knowledge sharing perceptions relate to their programming performance?

Literature review

The social dilemma of knowledge sharing

Extensive related studies have been conducted to investigate how individuals have shared knowledge in communities of practice (CoP) (Chiu, Hsu & Wang, 2006). In particular, much literature exists from studies that have looked into the factors that influence the knowledge sharing process in online communities, as the Internet may serve as an online platform that facilitates the sharing process. However, many knowledge management organizations have not successfully facilitated the sharing activity through online platforms due to the reluctance of individuals to share knowledge (Connolly, Thorn & Heminger, 1992; Jian & Jeffres, 2006). The use of online platforms does not necessarily enhance reciprocal knowledge sharing activities.

In a literature review of knowledge sharing cases by Cabrera and Cabrera (2002), they indicated that individuals may tend not to share their knowledge in a virtual community because such sharing may incur many types of costs, including not only that of relinquishing the knowledge that is shared but also those involved in the process of making that knowledge available to others and the social dynamic changes associated with the sharing that knowledge. Therefore, knowledge sharing in a virtual community has been considered a social dilemma, characterized by “paradoxical situations in which individual rationality -- simply trying to maximize individual payoff -- leads to collective irrationality” (Cabrera & Cabrera, 2002, p. 692). In such dilemmatic cases, refusal to share becomes the most advantageous strategy in securing personal benefit within a knowledge sharing context (Dawes, 1980). As more and more learning systems adopt the Internet as the platform to support learning in CoP (Fischer, 2001; Kafai & Peppler, 2011), there arises a need to resolve this dilemma and to enhance and encourage knowledge sharing in online platforms.

A social cognitive approach to facilitating knowledge sharing

Previous studies have investigated the factors that may contribute to the enhancement of willingness on the part of individuals to share knowledge in virtual communities. Kimmerle et al. (2007) identified specific psychological and structural solutions to the social dilemma. The structural solution mainly involves changing the pay-off structure to increase the personal pay-off of knowledge sharing behaviors (Cress & Martin, 2006; Cress et al., 2007). Such a solution is a cooperation-contingent transformation approach (Cabrera & Cabrera, 2002) in which a selective reward is offered to encourage sharing behavior in the individual.

Such an approach relies mainly upon the basic social cognitive theory which asserts that individual cognition and behavior are influenced and shaped in large part by the social networking context (Bandura, 1989). In particular, outcome expectations and self-efficacy are deemed to be the major cognitive forces guiding behavior (Chiu, Hsu & Wang, 2006). Bandura referred to self-efficacy as those capabilities that an individual perceives as producing designated levels of performance (Tsai et al., 2011) and an outcome expectation is a likelihood of the consequence such a performance will produce (Bandura, 1997). As individuals are not confident with their capability and possess low outcome expectation from the community, they may exhibit avoidance behavior to alleviate unexpected consequences of sharing knowledge with community members. Thus, to build up individuals’ confidence of and expectation form knowledge sharing could be critical to their increasing willingness to perform knowledge sharing activity. It has been confirmed by many studies that self-efficacy has a positive effect upon students’ behavioral
intention to use online learning course websites (Tsai et al., 2011). Furthermore, in the study by Hsu et al. (2007), they confirmed that outcome expectation and knowledge sharing efficacy profoundly influence knowledge sharing behavior. Such a result can also be explained by the value expectancy theory asserting that an individual’s behavior is determined by the perceived likelihood, that this behavior will produce a valued outcome (Kalman, Monge & Fulk, 2002).

The literature adduced above suggests that outcome expectation and self-efficacy may play an important role in knowledge sharing activity. Although, the above studies provide a framework for the investigation of knowledge sharing activity from the perspective of social cognitive theory, these studies analyze user perceptions primarily in informal settings. However, research into knowledge sharing in a formal education setting needs to consider also the role that specific pedagogic strategies may play (Bloom, 1984), so that educators may understand how pedagogy may be applied in classrooms. Therefore, the present study aimed to explore how Creative Commons may work as a mechanism in a school class setting to facilitate knowledge sharing and to investigate how students may react to the mechanism in terms of outcome expectation and self-efficacy regarding knowledge sharing.

A social capital approach to facilitating knowledge sharing

The other approach to enhancing knowledge sharing in virtual communities is a psychological solution. The psychological solution does not change the pay-off structure, but instead enhances psychological factors such as trust and the identification of the contributors interacting with others in the virtual community. The solution does not directly manipulate the incentive of contributors but instead increases willingness to share knowledge by enhancing the knowledge sharing mechanism (Kimmerle et al., 2007). For instance, in the study by Kimmerle et al. (2007), they found that identifying and bestowing distinction upon highly cooperative members of the virtual community may encourage participants to trust one another and may marginalize the effect of the fear of being exploited. The results of Fang and Chiu’s study (2010) also indicates that procedural and informational justice correlate significantly with a perception of trust in management which in turn contributes to active participation in knowledge sharing in a virtual community.

The above psychological solution to enhancing knowledge sharing relies mainly upon the social capital theory which asserts that an individual’s resources comprise not only those possessed by the individual but also those embedded within the social network to which the individual can connect (Nahapiet & Ghoshal, 1998). It has been indicated by Nahapiet & Ghoshal (1998) that, in addition to social network ties, social capital involves inter-personal relationships which manifest themselves as trust and reciprocity norms in social network relationships. Of the inter-personal factors present, trust has come to be considered the prevailing factor exerting influence upon the willingness of individuals to share knowledge in virtual communities. In the literature, trust is “a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another” (Rousseau et al., 1998, p. 395). More specifically, Mayer et al. (1995) defined trust as a set of specific beliefs related to benevolence, competence, and integrity (Mayer et al., 1995). Researchers have also generalized the concept of trust into a generalized trust that rests on behaviors that is generalized to a social unit as a whole rather than resting with a specific individual (Putnam, 1993). In the study by Kankanhalli, Tan and Wei (2005), they referred to trust as the belief in the good intent, competence, and reliability of individuals with respect to contributing and reusing knowledge through EKRs. Similar to such a generalized notion, Hsu et al. (2007) analyzed the influence of trust with respect to a detailed framework of trust including economy-based, information-based, and identification-based trust (Hsu et al., 2007) in a virtual community. Because such a framework of trust provides clear categories of trust related to online sharing platform, the present study adopted Hsu et al.’ definition and questionnaire to evaluate students’ perception in different online sharing platforms.

Trust on the part of community members is being considered crucial to the promotion of knowledge sharing in virtual communities (Hsu et al., 2007; Hsu, Chang & Yen, 2011; Chiu et al., 2006). Empirical studies indicate that trust has a positive impact on knowledge sharing intention (Chai & Kim, 2010), self-efficacy and expectation (Li, Hu, & Zhang, 2010) and also on the quality of knowledge sharing that takes place (Chiu et al., 2006). In other words, when a sense of insecurity is experienced in the online learning environment in which participants are engaged, participant intention and willingness to share knowledge, and to participate and interact with others in the virtual community may be reduced. As most Internet-based learning environments require students to learn through iterative
interaction and communication, it is important to foster student trust within these environments, trust which, in turn, shapes their intentions and behavior and engenders a positive attitude toward knowledge sharing.

The role of CC in knowledge sharing in classroom settings

It has been indicated by Cabrela and Cabrela (2002) that one way to increase knowledge sharing is to establish mechanisms by which individuals receive feedback whenever others use their contributions. Therefore, many studies have employed an Internet-based environment based on CC principles in support of knowledge sharing for collaborative learning purposes on the Internet (Monroy-Hernández & Resnick, 2008; Tanaka, Tokui, & Momeni, 2005). CC concepts may function as such a mechanism in a school class setting to facilitate a sharing atmosphere from the perspective of social cognitive and social capital theory. From the social cognitive perspective, the CC approach may increase student outcome expectations associated with knowledge sharing in their coursework. This is due to the fact that CC enforces an identity preservation strategy which requires the users of a work to acknowledge the original creators of the work, thereby increasing the visibility of the knowledge sharer in the community.

From a social capital perspective, the CC mechanism may well increase the sense of trust between students in a formal class setting. Previous research findings reveal that adopting the CC mechanism is helpful in shaping student perception of and interaction with the environment they are engaged in. For example, through integrating the CC mechanism into a collaborative drawing system, the findings of Liu et al.’s (2013) study reveal that the students manifested a higher level of mutual support under a CC-integrated collaborative environment. The empirical findings shed light on the previous suggestion that the principles and features of the CC mechanism are helpful in supporting a reliable sharing environment wherein student perception of knowledge sharing and behavior toward it are heightened. However, in most previous studies, the CC mechanism has mainly been employed to support informal online learning contexts such as the online discussion forum. In other words, the benefit of integrating the CC mechanism together with knowledge sharing activities in a formal learning course (e.g., classroom setting) was left unclear. In such a classroom setting, the social network ties may be different from those in informal context as students know each other in person. Furthermore, students are engaged in the same learning task and sharing of course work may influence their learning performance assessment. Therefore, how CC may play a role to facilitate knowledge sharing in the classroom setting is worthy of investigation. In this study, this concern was of particular importance and learning and course work are supported by the CC mechanism within a formal educational context in a programming course.

Methodology

Participants

The participants in this study were 40 freshmen majoring in Information and Communications in northern Taiwan (21 males and 19 females). They were enrolled in an introductory course in programming languages, and had the basic programming skills needed to perform the programming assignments given them. In addition, these were students who spent an average of 41.8 hours a week using the Internet. In other words, the participants had considerable experience using the Internet and performing tasks involved in research such as information uploading and exchanging. These participants may therefore help us to understand how CC may play a role in facilitating knowledge sharing among Internet users.

Instruments

The purpose of this study was to understand how students share knowledge, i.e., their programming code, in a programming course. Therefore, the online learning environment of this study involved two components: a construction component, which supported students in coding a program, and a sharing component that allowed them to demonstrate their work and to access that developed by others. As shown in Figure 1, the construction component allowed them to build a railway system using building blocks such as track and bridge modules and to apply the programming language (Java) to program the behavior of the railway system. One of the purposes of this study was to understand how the CC mechanism may facilitate the sharing experience. Two different versions of sharing
components were implemented: one of the versions was the conventional online sharing platform, while the other was a CC-integrated online sharing platform. On both platforms, the students could upload their programming works, and were allowed to decide whether or not to share with others. With works shared on the conventional sharing platform, the authors’ name, upload time and date were displayed, but the status of how the shared works were retrieved and reused was not visible (Figure 2).

By contrast, the CC-integrated online sharing platform adopted the CC mechanism to enhance the knowledge sharing experience. This study used the Share-alike option of the CC mechanism with which the students could revise and distribute shared programming works with an identical license indicating the original creators’ contributions. For a shared work that was declared sharable in this platform, as shown in Figure 3, the platform would display the creator’s name and uploading time and date, along with all programming works that were built by modifying this work. In this context, the students would always know who originally created and shared a work, the authorship of the shared work, and who had used this work to create a new one. Consequently, whether or not the CC mechanism was used determined how the two online sharing platforms functioned (i.e., conventional versus the CC-integrated online sharing platforms) in many distinctive aspects.
To explore student perception of knowledge sharing, a set of questionnaires by Hsu et al. (2007) were employed to examine the students’ degree of trust, their outcome expectations and their self-efficacy toward knowledge sharing in this study. The definition of each factor conducted to assess student perceptions of knowledge sharing is listed as follows:

- Economy-based trust: To examine students’ trust in the sharing platform due to the costs and benefits in time, knowledge and advantage.
- Information-based trust: To assess students’ trust in the sharing platform due to information privacy and openness.
- Identification-based trust: To examine students’ trust in the sharing platform regarding inter-dependency and shared identity among community members.
- Knowledge sharing self-efficacy: To assess individuals’ perceived self-efficacy of knowledge sharing in the platform.
- Personal outcome expectation: To examine individuals’ expectations about their personal gains from sharing in the platform.
- Community-related outcome expectation: To examine individuals’ expectations about the impact of their sharing on other members in the platform.

The questionnaires contained a total of 31 items, categorized into six dimensions of knowledge sharing which included “Economy-based trust” (4 items, e.g., By joining this online community, I will save time and effort in learning programming and sharing.), “Information-based trust” (4 items, e.g., This online community protects personal works from unauthorized access.), “Identification-based trust” (6 items, e.g., If I share my works, I know others will respond constructively and caringly.) “Personal outcome expectations” (5 items, e.g., If I share my works with other community members, I will get more cooperation and benefit in return.), “Community-related outcome expectations” (5 items, e.g., My work would help this community accumulate or enrich this website.) and “Knowledge sharing self-efficacy” (7 items, e.g., how confident are you in providing your works?). The original reliability coefficient of each dimension ranged from 0.59 to 0.75. By assessing 40 students’ responses to the questionnaire, the reliability coefficient of each dimension was 0.83, 0.81, 0.76, 0.92 0.93 and 0.94, which serves as a reliable instrument for assessing the student perception in knowledge sharing in this study.

To assess students’ perceptions of knowledge sharing toward a particular sharing platform in which they were engaged (i.e., conventional or CC-integrated), the original questionnaire was modified by offering specifically instructional guidance to highlight the sharing platform particularly conducted at each stage. That is, the implement of the questionnaire was conditional on the conducted sharing platforms at different stages. The students’ responses to the questionnaires could represent their perceptions of knowledge sharing toward conventional and CC-integrated sharing platforms respectively for later comparisons.
Research procedure

This study was conducted in a formal classroom setting with regular course scheduling. Since the students were enrolled in the same programming course, to divide them into two groups with different treatments would be conflict with the school policy because of unequal conditions for course assessment. In this regard, the within-subjects design was adopted and both experimental treatments were viewed as instructional assistants to support students’ programming learning in this study. However, this study did not only compare the CC-integrated platform with the conventional platform, but also compare the perceptions of the students’ who shared their works with those who did not in the two platforms. Therefore, the results of the study may be helpful to reveal the factors that influence student knowledge sharing willingness in the two different platforms.

The research procedure consisted of two stages over a period of six weeks. By using the construction component and the online sharing platform, all of the participants were assigned the task of designing and implementing a solution to the given programming problems at two different stages within the schedule of the programming curriculum.

The first stage lasted 3 weeks, during which the participants were first introduced to the construction component, the online sharing platform and the programming task. Then the participants were asked to construct a first railway model using the construction component, and upload their works to the online sharing platform without applying the CC mechanism. Their activities on the platform were logged for further analysis. By the end of this stage, they needed to complete the questionnaire—an instrument assessing their perceptions of knowledge sharing in the online sharing platform without the CC mechanism.

The second stage also lasted 3 weeks. During this stage, the participants were introduced to the concept of the CC mechanism. Then they were asked to construct another railway model, and upload their works to the online sharing platform, this time incorporating the CC mechanism. Finally, the students completed the same questionnaire assessing the participants’ perceptions of knowledge sharing within the sharing activity in the platform using CC.

Since this study adopted the within-subjects design, ones may consider that the students may be aware of the instructor’s intentions of the following treatment, and then displayed involuntary behaviors in response to the instructor’s expectation while introducing the CC mechanism. In this respect, this study introduced the CC mechanism as a kind of technical assistance in providing the students the alternative of information sharing. Such an introduction did not act as an instructional strategy to stimulate the students’ behaviors and perceptions of sharing programming works. Therefore, only the characteristics of rather than the benefits of the CC mechanism about information sharing were introduced by the class instructor. The instructor did not offer additional assistance in the completion of programming work except for technical support in resolving unexpected crashes in the system.

Data analysis

The participants’ activity logs and their perceptions of knowledge sharing obtained from the questionnaire were used as analytical resources in this study. As reciprocally interactive activities should be initialized by one’s sharing voluntariness, this study considered the students’ sharing decision making (share versus not-to-share) as an initial step of nurturing consequential interactions between each other. In addition, from the technical aspects, only when a student shared her/his program work, the online sharing platform could display the works to others. Therefore, explicit sharing orientation (share versus not-to-share) is necessary to design the online sharing platform. Given these reasons, this study adopted sharing orientation as a dichotomous variable. To understand the role of the CC mechanism in affecting their perceptions of knowledge sharing, a series of dependent t-tests were employed to examine the differences between perceptions of knowledge sharing in the two online sharing platforms (the conventional versus the CC-integrated). The relationship between the perceptions of knowledge sharing and sharing orientations on the two online sharing platforms, were also examined in order to achieve a better understanding of the factors influencing the students’ intention to share their works on the two platforms. In addition, the method of the McNemar test was used to analyze the participants’ decisions on sharing/non-sharing knowledge in the two online sharing contexts.

Four indicators were selected to represent the students’ participatory interactions with others in the platforms. The frequency of viewing/downloading personal/peer programming works was calculated through analysis of log data,
because the four activities could reveal the participants’ efforts in sharing their programming works on the platforms. These sharing behaviors recorded on log data may represent seed (uploading programing works) or follow-up contributions among the students (e.g., viewing or downloading peers’ works). While the students reviewed the works shared by their peers, they may appreciate the merits of the shared works and turn to revise one’s own works. However, in the CC mechanism, some seed contributions may be original or derived by using other works. It becomes complicated to distinguish clearly seed and follow-up contribution. More fine-grained and specific ways such as sequential analysis or social networking analysis may be more applicable to analyze reciprocally dynamic process. Since this study aimed to reveal the difference of behaviors and perceptions of knowledge sharing between conventional and the CC-integrated sharing platforms, these behavioral records were analyzed at a general level without distinguishing the seed and follow-up contribution to avoid losing the foci of this study.

A set of evaluative criteria was employed to grade the students’ performance on the programming tasks from the aspects of the correctness, structure, and complexity involved in the railway model for each programming task. Three qualifying standards in the number of train and railway (40 points), the length and modeling of the railway (40 points), the completeness and correctness of program execution (40 points) would be considered for the grading of the railway model. The students would obtain a total score ranged from 0 to 120 for each programming task.

The evaluation criteria for the programming task were introduced to the students before entering into the research procedure. They were also informed that their sharing activities on both platforms would not be included in the grading of the programming performance. Because this study would focus on understanding the students’ perception facilitated by the CC mechanism as a psychological solution to enhance sharing. In other words, this study would not change the pay-off structure associated with the knowledge sharing behavior. Therefore, the students’ decision on whether to share or not would not be taken into account in determining their scores and grades in the course. Only the average score of two programming works would represent 10% of each student’s total achievement this semester.

Finally, for each of the two different online sharing platforms, the Spearman’s correlation analysis was used to examine the relationships between participatory activities on the sharing platforms and the perceptions of knowledge sharing to programming task performances. By contrasting the relative results in different online sharing platforms, the patterns of student interaction with and perceptions of knowledge sharing in the different online sharing platforms were identified and interpreted.

**Results and discussion**

**The perceptions of knowledge sharing in conventional and the CC-integrated online sharing platforms**

A series of paired t-tests were employed to explore the role of the different online sharing platforms in the perceptions of knowledge sharing. Table 1 displays the results of the students’ perceptions of knowledge sharing toward the two different online sharing platforms. Except for the dimension of information-based trust, the students’ responses to the perception of knowledge sharing toward the CC-integrated online sharing platform were significantly higher than those of the conventional online sharing platform. The results revealed that the students had a more positive perception of knowledge sharing in the CC-integrated platform than in the conventional online sharing platform. Such results support that students can establish a higher level of trust and confidence in sharing knowledge on the CC-integrated platform as it affords a protection mechanism. Students may therefore have heightened expectations of personal and interpersonal gain when interacting with that platform.

<table>
<thead>
<tr>
<th></th>
<th>Conventional online sharing platform</th>
<th>CC-integrated online sharing platform</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ET</td>
<td>Mean / S.D.: 4.16/0.85</td>
<td>Mean / S.D.: 4.41/0.74</td>
<td>-2.865</td>
</tr>
<tr>
<td>IT</td>
<td>4.40/0.78</td>
<td>4.58/0.83</td>
<td>-1.365</td>
</tr>
<tr>
<td>DT</td>
<td>4.25/0.64</td>
<td>4.51/0.60</td>
<td>-2.767**</td>
</tr>
<tr>
<td>SE</td>
<td>4.06/0.86</td>
<td>4.41/0.85</td>
<td>-3.808***</td>
</tr>
<tr>
<td>PE</td>
<td>3.44/0.96</td>
<td>3.70/1.06</td>
<td>-2.161*</td>
</tr>
<tr>
<td>CE</td>
<td>3.98/0.95</td>
<td>4.39/0.92</td>
<td>-3.101**</td>
</tr>
</tbody>
</table>

**Note.** Economy-based trust (ET), Information-based trust (IT); Identification-based trust (DT); Knowledge sharing self-efficacy (SE); Personal outcome expectation (PE); Community-related outcome expectation (CE).

*p < 0.05. **p < 0.01. ***p < 0.001.
The relations of the perceptions of knowledge sharing to sharing/non-sharing decision on the conventional and the CC-integrated online sharing platforms

This further examined the differences in knowledge sharing participation level between that of students who shared (Sharing) and those who did not share their works (Non-Sharing) in the two different sharing platforms. As shown in Table 2, the students oriented toward sharing generally evinced more positive perceptions of knowledge sharing than those with non-sharing orientation in both of the two platforms. Furthermore, the t-test results reveal that there were no significant differences found in the conventional online sharing platform. By contrast, the students with different sharing orientations in the CC-integrated online sharing platform exhibited a significant difference in their perceptions of knowledge sharing. More specifically, those students with sharing orientation manifested a significantly higher level of identification-based trust and personal outcome expectation than those disinclined to share. In other words, those students with sharing orientation tended to appreciate the merits of applying the CC mechanism in support of knowledge sharing activities to augment trust among community members and personal gain from the sharing activities.

The findings of this study offer some evidence that the perception of knowledge sharing correlated positively with student decisions regarding sharing with online community members. The students who actually shared their compositions with peers on the CC-integrated online sharing platform exhibited a significantly higher degree of the perceptions of knowledge sharing (i.e., identification-based trust and personal outcome expectation) than did those unwilling to share knowledge.

Table 2. The differences in student perception of knowledge sharing -- between sharing and non-sharing decisions – for the two different online sharing platforms

<table>
<thead>
<tr>
<th></th>
<th>Conventional online sharing platform</th>
<th>CC-integrated online sharing platform</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean / S.D.</td>
<td>Mean / S.D.</td>
</tr>
<tr>
<td>Sharing</td>
<td>N = 23</td>
<td>N = 32</td>
</tr>
<tr>
<td>Non-Sharing</td>
<td>N = 17</td>
<td>N = 8</td>
</tr>
<tr>
<td>ET</td>
<td>4.33/0.76</td>
<td>4.50/0.59</td>
</tr>
<tr>
<td></td>
<td>-1.50</td>
<td>4.03/1.15</td>
</tr>
<tr>
<td>IT</td>
<td>4.50/0.74</td>
<td>4.65/0.83</td>
</tr>
<tr>
<td></td>
<td>-0.937</td>
<td>4.28/0.82</td>
</tr>
<tr>
<td>DT</td>
<td>4.40/0.59</td>
<td>4.62/0.53</td>
</tr>
<tr>
<td></td>
<td>-1.811</td>
<td>4.08/0.70</td>
</tr>
<tr>
<td>SE</td>
<td>4.21/0.88</td>
<td>4.48/0.69</td>
</tr>
<tr>
<td></td>
<td>-1.264</td>
<td>4.13/1.34</td>
</tr>
<tr>
<td>PE</td>
<td>3.52/0.90</td>
<td>3.90/0.89</td>
</tr>
<tr>
<td></td>
<td>-0.663</td>
<td>2.88/1.34</td>
</tr>
<tr>
<td>CE</td>
<td>4.23/0.90</td>
<td>4.50/0.66</td>
</tr>
<tr>
<td></td>
<td>-2.027</td>
<td>3.93/1.58</td>
</tr>
</tbody>
</table>

* Shapiro-Wilk test on the distribution of students making the decision to share or not to share in the different online sharing platforms

Table 3. McNemar test on the distribution of students making the decision to share or not to share in the different online sharing platforms

<table>
<thead>
<tr>
<th></th>
<th>Sharing</th>
<th>Non-sharing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional online sharing platform</td>
<td>23 (57.5%)</td>
<td>17 (32.5%)</td>
<td>40</td>
</tr>
<tr>
<td>CC-integrated online sharing platform</td>
<td>32 (80%)</td>
<td>8 (20%)</td>
<td>40</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 5.4; \text{ p-value } = 0.035. \]

To examine how conventional and CC-integrated online sharing platforms support knowledge sharing, McNemar’s test was applied to make comparisons of sharing orientations between the two platforms, given two matched subjects. The results of the test showed that the students’ sharing orientations changed significantly from the conventional online sharing platform to the CC-integrated online sharing platform (\( p < .05 \)). As shown in Table 3, the students were more likely to share their work in the CC-integrated online sharing platform (80%) than they did in the conventional online sharing platform (57.5%). The findings support the conclusion that students were more favorable toward knowledge sharing activities in the CC-integrated online sharing platform.
Statistics of activities and programming task performance on the different online sharing platforms

Table 4 reveals the students’ participatory interactions and programming performance in the two different online sharing platforms for learning programming. Seven students who were enrolled in the programming construction activity in the first stage were absent from the knowledge sharing activity in the second stage. Therefore, only the 33 students completing both stages were included in the comparative analyses. As shown in Table 4, the students demonstrated a higher frequency of downloading personal programming works and viewing each other’s shared programming works while engaging in the CC-integrated online sharing platform than they did in the conventional online sharing platform.

Table 4. Paired samples test among two sharing platforms for participatory interactions and programming performance

<table>
<thead>
<tr>
<th>Participatory interactions</th>
<th>Conventional sharing platform</th>
<th>CC-integrated sharing platform</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of viewing personal works</td>
<td>2.24/2.11</td>
<td>1.79/1.65</td>
<td>1.15</td>
<td>0.29</td>
</tr>
<tr>
<td>Frequency of downloading personal works</td>
<td>0.15/0.44</td>
<td>0.45/1.77</td>
<td>3.73</td>
<td>0.06</td>
</tr>
<tr>
<td>Frequency of viewing peer’s shared works</td>
<td>0.88/1.85</td>
<td>1.91/4.38</td>
<td>2.61</td>
<td>0.12</td>
</tr>
<tr>
<td>Frequency of downloading peer’s shared works</td>
<td>2.09/3.48</td>
<td>2.48/2.11</td>
<td>0.37</td>
<td>0.55</td>
</tr>
<tr>
<td>Performance of programming works</td>
<td>107.30/12.12</td>
<td>113.18/13.51</td>
<td>4.08</td>
<td>0.05</td>
</tr>
</tbody>
</table>

A repeated measure ANOVA was conducted on four indicators of participatory interaction and programming performance in the two online sharing platforms. The results of multivariate test by within-subjects design revealed Wilks’ $\lambda = 0.97, 0.97, 0.93, 0.99, F(1,32) = 0.29, 0.89, 2.61, 0.37, p < .05$ for the four participatory interactions respectively, and Wilks’ $\lambda = 0.89, F(1, 32) = 4.08, p < .05$ for the programming performance between the online sharing platforms. Thus, there were no significant paired differences in the students’ participatory interactions and programming performance toward different online sharing platforms.

Because the ANOVA overall test did not yield a significant result, this study further examined the relations of the students’ perceptions and activities of knowledge sharing to their programming performance in both of the two online sharing platforms. The contrast of the interrelationships between the two online sharing platforms may shed light on the important factors profoundly related to programming performance in the two different online sharing platforms.

The relationship of activities and of the perception of knowledge sharing to programming task performance on the different online sharing platforms

Correlation analyses were conducted to understand the relationship of activities and of the perception of knowledge sharing to programming task performance in two online sharing platforms. As shown in Table 5, in two participatory interactions (i.e., viewing and downloading personal works), identification-based trust and knowledge sharing self-efficacy were correlated significantly with student performance of programming works in the conventional online sharing platform. By contrast, there were no significant relationships between participatory interactions and the student performance of programming works in the CC-integrated online sharing platform. However, three dimensions of knowledge sharing perception, i.e., economy-based trust, knowledge sharing self-efficacy, and community-related outcome expectations, revealed a significantly positive correlation with programming work performance. This correlation between the two different sharing platforms reveals that knowledge sharing self-efficacy plays a critical role in programming work performances in both of the two platforms. This correlation
between community-related outcome expectation and programming work performance was highest in the perceptions of knowledge sharing in the CC-integrated online sharing platform. Such results may imply that, within the CC mechanism, students may perform better when they expect their contributions to be able to enrich the learning community through sharing their works with other members.

Table 5. The relation of participatory activities and perception of knowledge sharing to programming performances on different online sharing platforms

<table>
<thead>
<tr>
<th></th>
<th>Performance with conventional online sharing (n = 33)</th>
<th>Performance with CC-integrated online sharing (n = 33)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interaction activities in the online platforms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of viewing personal works</td>
<td>0.34*</td>
<td>0.17</td>
</tr>
<tr>
<td>Frequency of downloading personal works</td>
<td>0.39*</td>
<td>0.01</td>
</tr>
<tr>
<td>Frequency of viewing peer’ shared works</td>
<td>0.24</td>
<td>0.21</td>
</tr>
<tr>
<td>Frequency of downloading peer’ shared works</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Perceptions of knowledge sharing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economy-based trust</td>
<td>0.01</td>
<td>0.41*</td>
</tr>
<tr>
<td>Information-based trust</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td>Identification-based trust</td>
<td>0.36*</td>
<td>0.30</td>
</tr>
<tr>
<td>Knowledge sharing self-efficacy</td>
<td>0.36*</td>
<td>0.40*</td>
</tr>
<tr>
<td>Personal outcome expectation</td>
<td>0.08</td>
<td>0.23</td>
</tr>
<tr>
<td>Community-related outcome expectation</td>
<td>0.29</td>
<td>0.59**</td>
</tr>
</tbody>
</table>

*p < 0.05. **p < 0.01.

In contrast, on the conventional online sharing platform, students’ participatory interaction (i.e., the frequency of viewing personal works and frequency of downloading personal works), as well as their perceptions of knowledge sharing regarding identification-based trust and knowledge sharing self-efficacy corresponded significantly with programming work performance. Therefore, in the conventional online sharing platform, students usually focused on their own programming works even if they were willing to share their works with others on the platform. They may have intended to gain identity recognition through the knowledge sharing activity rather than through contributing to the community. In contrast, the students’ perceptions of knowledge sharing (i.e., economy-based trust, identification-based trust, knowledge sharing efficacy, and community-related outcome expectation) became a significantly prominent factor in their programming work performances on the CC-integrated platform.

Discussion and conclusions

The present study explored the role of the Creative Commons (CC) in student perceptions of knowledge sharing in a university programming course. In this study, the findings revealed some relationships among student learning behaviors, perceptions of knowledge sharing and their programming task performances. For example, the activity involving student self-examination of their programming works was a significant factor in programming performance on the conventional sharing platform, but not in the CC-integrated platform. In other words, student performance was more a function of individual learning activities in the conventional sharing platform. However, student perception of knowledge sharing showed a stronger correlation with programming performance on the CC-integrated than the conventional platform. The above results support the contention that the CC mechanism transformed individual leaning on the conventional sharing platform into a social learning activity.

Furthermore, in addition to assessing the student willingness to share knowledge (e.g., Hsu et al., 2007; Hung & Cheng, 2013), this study examined students’ sharing behaviors with reference to their actual sharing or non-sharing decisions. On both of the online sharing platforms, the students who chose the sharing option generally showed positive responses to perception of knowledge sharing than did the others who made the decision not to share. However, the students showed different patterns of interrelationships between the perceptions of knowledge sharing and sharing willingness on the two platforms respectively. In contrast with the previous finding that trust could be predominant in knowledge sharing behaviors (e.g., Hsu et al., 2007), these comparative results provide some evidence that the students’ identification-based trust and expectation of personal profit may be a salient factor in their subsequent engagement in knowledge sharing activities mediated by the CC-integrated platform.
In addition, significant transition from the non-sharing to the sharing decision reveals that the students were more willing to share their works with peers on the CC-integrated sharing platform. Hence, this study suggests that the online sharing environment integrated with the CC mechanism not only provides students with a reliable tool for knowledge sharing, but also develops their perceptions of knowledge sharing and willingness to share learning works. In other words, instructors and system developers could consider the CC mechanism to be valuable in creating a secure and comfortable Internet-based environment supportive of knowledge sharing in educational settings.

Consistent with findings in previous studies reporting significant correspondence between self-efficacy and learning performance (Moos & Azevedo, 2009), student knowledge sharing self-efficacy still played a critical role in programming performance in both online sharing platforms. In addition to the knowledge sharing self-efficacy, student trust in identification was significantly related to their programming performance on the conventional platform, whereas their economy-based trust (i.e., trust in receiving benefit from sharing) and community-based outcome expectation (i.e., belief in their contributions to the growth of the community knowledge base) became critical to their programming performance on the CC-integrated platform. These results suggest that the student perception of knowledge sharing, as supported by the CC mechanism, would be crucial to learning performance. Furthermore, in addition to the research into online participation and interaction for collaborative learning (e.g., Chou & Min, 2009; Lang, 2010; Lin & Tsai, 2012), the employment of the CC-integrated platform may serve as a reliable mechanism to facilitate a sense of trust among students and to satisfy individual expectations in such a social learning context.

The present study is only based on a local interest and concern since the sample has been collected from only Taiwanese class. Future study could try to generalize the study's findings with extended sample representing non-Taiwanese subjects. Instead of confirming the effect of the CC-integrated platform, this study turned to differentiate the students' perceptions and activities of knowledge sharing toward the sharing platforms. In this regard, this study performed the within-subjects comparisons to offer some evidence on how the students perceived and interacted with the sharing platforms for knowledge sharing. Since ones may consider the learning effect bias caused by the within-subjects design, this study suggests that experimental design could be conducted to examine the effectiveness of the CC-integrated platform on the students' learning through knowledge sharing. Further research with a larger sample may be needed to validate the causal effects of the CC mechanism on student perception and knowledge sharing behaviors. An intensive investigation into the correlation between the perceptions of knowledge sharing and the process of programming works within a CC-integrated online sharing platform may inform our understanding of how individual differences may alter programming learning trajectories in a knowledge sharing context.

Acknowledgements

This research was partially funded by the Research Center for Science and Technology for Learning at National Central University and the National Science Foundation under 101-2511-S-008 -006 -MY2 and 101-2511-S-008 -005 -MY3.

References


The Role of Community Trust and Altruism in Knowledge Sharing: An Investigation of a Virtual Community of Teacher Professionals

Hsiu-Ling Chen1*, Hsueh-Liang Fan2 and Chin-Chung Tsai1

1Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology. No. 43, Sec. 4, Keelung Rd., Taipei, 106, Taiwan // 2Department of Business Administration, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 106, Taiwan // shirley@mail.ntust.edu.tw // sam.fan@gmail.com // cctsai@mail.ntust.edu.tw

*Corresponding author

(Submitted May 07, 2013; Revised October 07, 2013; Accepted November 12, 2013)

ABSTRACT

The knowledge sharing process within a virtual community of teacher professionals is viewed as a social exchange process in that the knowledge sharing intention and behavior of individuals are influenced by the exchange relationship among members. However, relatively little research has focused on this approach to exploring the factors that enhance the knowledge sharing of individuals. This study aims to explore the mechanism of how individuals’ knowledge sharing intention and behaviors might be altered with a sample of 332 teachers within a virtual community of teacher professionals (i.e., SCTNet) in Taiwan. After controlling for the technology support factors (i.e., perceived usefulness and perceived ease of use), we found that community trust impacts knowledge sharing intention, which in turn improves knowledge sharing behavior. Furthermore, altruism augmented the relationship between community trust and knowledge sharing intention. The positive relationship between community trust and knowledge sharing intention was stronger when teachers perceived a high level of altruism than it was with a low level of altruism.

Keywords
Knowledge sharing, Community trust, Altruism

Introduction

In a rapidly changing educational environment, students’ learning and achievement depend upon the professional development of teachers. In response to their needs, educational research and practice has begun to emphasize the importance of teacher professional communities for improving the professional growth of teachers. The Internet enables groups of people to interconnect through virtual communities (Blanchard, 2008). Teachers within virtual communities share and exchange personal experiences of and resources for teaching (i.e., knowledge sharing); this would help their professional development and enhance their teaching quality. SCTNet is short for Smart, Creative Teachers Network, and aims to construct a teacher professional community through which members can share information, knowledge, and teaching experience (Lin, Lin, & Huang, 2008). SCTNet has become the biggest virtual community of teacher professionals in Taiwan. However, virtual communities are of limited value without rich knowledge (Chiu, Hsu, & Wang, 2006). This presents the new challenge of how to effectively motivate knowledge sharing among community members. Scholars have suggested that we could support knowledge sharing through social exchange relationships (Blanchard, 2008). Thus, the objective of this study is to adopt social exchange theory as a way of exploring the underlying mechanism that motivates knowledge sharing by community members.

Social exchange theory explains interactive relationships among individuals across various disciplines (Cropanzano & Mitchell, 2005). The theory suggests that social exchange involves various interactions that generate obligations. The interactions among individuals are also viewed as interdependent and are contingent on the actions of others (Blau, 1964; Cropanzano & Mitchell, 2005). Knowledge sharing processes within a virtual community are viewed as a social exchange relationship in which voluntary acts of individual knowledge sharing are motivated by the returns they are expected to bring to and receive from other members (Ma & Yuen, 2011). When considering the reciprocal relationship among people, community trust plays a fundamental role in knowledge sharing in virtual communities. Trust is embedded in a person or in a social network (Granovetter, 1985). When reciprocal acts occur in social interaction, individuals may trust each other and are likely to share personal feelings, information, and knowledge. In a virtual community, the members were not previously familiar with each other, and thus community trust is required for successful interaction and knowledge sharing to take place (Hsu & Lin, 2008; Ridings, Gefen, & Arinze, 2002).

Prior literature has indicated that a person’s behavioral intention influences subsequent behavior (Ajzen & Fishbein,
1974; Bock, Zmud, Kim, & Lee, 2005; Davis, 1989; Hsu & Lin, 2008; Webb & Sheeran, 2006). We thus hypothesize that community trust enhances knowledge sharing behavior through knowledge sharing intention.

As community trust plays a fundamental role in enhancing knowledge sharing, exactly when the influence of community trust on individuals’ knowledge sharing intentions would be augmented needs to be well understood. We suggest that altruism, a kind of helping or sharing behavior (Hoffman, 1979), may augment the relationship between community trust and knowledge sharing intention. Altruism represents an individual’s willingness to benefit the well-being of others on a voluntary basis without the anticipation of any form of return (Chai & Kim, 2010; Deci, 1975; Kankanhalli, Tan, & Wei, 2005). When individuals perceive an atmosphere of trust in the virtual communities, those with a high level of altruism are more likely to freely share information or discuss personal experiences in the community than those with a low level of altruism. We thus hypothesize that altruism is a moderator which augments the relationship between community trust and knowledge sharing intention.

In order to rule out extraneous influences on the relationship among community trust and altruism in knowledge sharing intention as well as knowledge sharing behavior, we have to take into account other possible confounding variables. Prior literature has indicated that technology support factors are related to knowledge sharing (Hsu & Lin, 2008; Hung & Cheng, 2013; Reychav & Te'eni, 2009). For example, Hsu and Lin (2008) found that a user’s beliefs in blog usage (i.e., perceived usefulness and perceived ease of use) may affect the individual’s attitudes toward participating in blog activities such as sharing information and knowledge. Hung and Cheng (2013) also found that perceived usefulness and perceived ease of use regarding a virtual community facilitated knowledge-sharing intentions. To account for this relationship, perceived usefulness and perceived ease of use were controlled in our data analysis. The control variables in our study are statistical controls rather than experimental controls; we entered the control variables into the hierarchical regression model before other independent variables (Atinc, Simmering, & Kroll, 2012). Our theoretical framework is presented in Figure 1. A detailed discussion of the proposed framework is presented below.

Community trust as an antecedent of knowledge sharing intention

Knowledge sharing intention refers to the willingness of individuals within a community or an organization to share with others the knowledge they own (Bock et al., 2005), while knowledge sharing behavior refers to the individuals’ behaviors of sharing knowledge. According to the Theory of Reasoned Action (Ajzen & Fishbein, 1974), a person’s behavior is determined by his or her intention. Prior literature suggests that a person’s behavioral intention influences subsequent behavior across different contexts such as knowledge sharing (Bock et al., 2005; Chen, Chen, & Kinshuk, 2009; Hsu & Lin, 2008), the usage of information technology (Davis, 1989), and changes in social and health behavior (Webb & Sheeran, 2006). However, as outlined earlier, the availability of a virtual community does not necessarily automatically encourage individuals to share knowledge (Usoro, Sharratt, Tsui, & Shekhar, 2007). We propose that community trust plays a fundamental role in knowledge sharing.

The nature of community trust is similar to that of trust. When considering different targets, scholars use different nouns to describe trust, such as organizational trust, team trust, and community trust. Community trust represents a
psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behaviors of another (Rousseau, Sitkin, Burt, & Camerer, 1998). The degree of vulnerability is enhanced in situations where the parties are interdependent such that the interest of one party cannot be achieved without the other (Aryee, Budhwar, & Chen, 2002). Prior literature on trust has shown that it is associated with enhanced cooperation, knowledge sharing, information sharing, problem solving, and collaborative learning (Bulu & Yildirim, 2008; Jameson, Davies, & de Freitas, 2006; Granovetter, 1985; Levin & Cross, 2004; Lewicki & Wiethoff, 2000). For example, Levin and Cross (2004) supported Granovetter’s (1985) suggestion that trust is a kind of social capital which is embedded in an exchange relationship. In a face to face context, they found that when trust exists, individuals are willing to share knowledge as well as to acquire knowledge from others.

In a virtual community, the members were not previously familiar with each other, and therefore this requires building an atmosphere of trust within the community (Ridings et al., 2002). Although scholars have introduced the social exchange theory to understand individuals’ motivational mechanisms behind knowledge sharing in Internet contexts such as blogs (Chai & Kim, 2010; Hsu & Lin, 2008), online learning platforms (Ma & Yuen, 2011), and virtual communities (Chiu et al., 2006; Ridings et al., 2002; Wasko & Faraj, 2005), relatively little research has provided empirical evidence of the influence of community trust on knowledge sharing within virtual communities of teacher professionals.

We hypothesize that community trust enhances knowledge sharing behavior through knowledge sharing intention. When individuals perceive a strong trust in the virtual community, they would perceive that the community provides a meaningful rationale for engaging in knowledge sharing behavior. Consequently, an individual’s trust is ultimately related to his/her knowledge sharing behavior, and this linkage is mediated by the knowledge sharing intention. The above relationship indicates that knowledge sharing intention mediates the relationship between community trust and knowledge sharing behavior.

Hypothesis 1: After controlling for perceived usefulness and perceived ease of use, knowledge sharing intention mediates the extent to which community trust enhances knowledge sharing behavior.

The moderating role of altruism

The hypothesized links between community trust and knowledge sharing intention shown in Figure 1 are expected to be moderated by altruism, i.e., altruism augments the community trust—knowledge sharing intention relationship. Kankanhalli et al. (2005) suggested that although there are few examples of absolute altruism (absolute lack of self-concern in the motivation for an act), relative altruism (individuals’ desire to help others) is more prevalent. In our study, we adopt Kankanhalli et al.’s concept of relative altruism as our definition of altruism and use their “enjoyment in helping others scale” to measure altruism.

Altruism is not only viewed as an important value in relationship-oriented societies such as those in the Chinese-speaking world, but there is also considered to be an increasing need for the acquisition of altruistic behavior in educational systems (Baytiyeh & Pfaffman, 2010; Etxebarria et al., 1994) and volunteer communities such as religious communities and public welfare communities (Fischer & Schaffer, 1993). Prior literature has suggested that individuals’ altruistic behavior can be motivated by the enjoyment of helping others. For example, in an organizational context, Kankanhalli et al. (2005) found that employees may share knowledge because they enjoy helping others (i.e., altruism). In a study of electronic communities of practice, Wasko and Faraj (2000) reported that members were willing to share personal knowledge and experiences because they wanted to contribute to the overall welfare of the community.

As mentioned earlier, the interactive relationship among individuals within the virtual community is viewed as a social exchange relationship. Individuals are not familiar with each other in such a community and thus a basic foundation for encouraging knowledge sharing is to build an atmosphere of trust. As people who trust each other, individuals who enjoy helping others are likely to exhibit greater knowledge sharing intention. For example, they would have a greater willingness to share information or discuss personal experiences in the community. Thus, we propose that altruism can enhance the association between community trust and knowledge sharing intention. Individuals who highly trust the virtual community would demonstrate a greater knowledge sharing intention with a
high level of altruism. Exploring the moderating effect of altruism offers a novel explanation of why altruism strengthens the relationship between community trust and knowledge sharing intention. According to our knowledge of the prior literature, few studies have investigated this relationship.

Hypothesis 2: After controlling for perceived usefulness and perceived ease of use, altruism will moderate the relationship between community trust and knowledge intention in such a way that the higher the frequency of altruism, the more pronounced the positive association between community trust and knowledge sharing intention will be.

Method

Research context and data collection

The SCTNet virtual community for teachers (http://sctnet.edu.tw) was investigated in this study. SCTNet was established in March 2000 as a result of the joint efforts of academic institutes and government departments in Taiwan. Based on Wenger’s (1998) definition of communities of practice, SCTNet is viewed as a virtual community of practice wherein teachers engage in knowledge sharing and learning processes. Currently, additional sponsorship of the platform comes from the “achieving excellence project” and the Ministry of Education. To date (2013/06), there are 157,604 teachers registered in the community. More than 5,900 teaching resources have been uploaded, such as information regarding instructional design, personal teaching experiences, teaching cases, and instructional aided media.

For this study, teachers were asked to respond to a questionnaire through SCTNet, administered with the website manager’s assistance. In total, 401 respondents completed the questionnaire survey, of which 69 were eliminated due to missing responses or the members not being teachers, leaving a total effective sample of 332. The results of the descriptive analysis show that 33.1% of the respondents were male; 6.6% were kindergarten teachers, 65.7% were elementary school teachers, 20.2% were junior high school teachers, and 7.5% were high school teachers. In addition, 79.8% of respondents had used SCTNet for more than one year; 56.1% of respondents had visited SCTNet once a week, while 21.4% had visited it once every two weeks.

Instruments

We followed Hinkin’s (1995) suggestions for administering the research instruments. First, the research instruments for this study were adapted from previously published studies which have considerable validity and reliability. Second, each measure included at least 4 items, which is the minimum number recommended for summing by Norman (2010). Third, anonymous respondents were asked to rate their perception of each item using a 5-point Likert scale (1 = “strongly disagree” to 5 = “strongly agree”). Such 5 point scales have been commonly used in prior studies as they can generate sufficient variance among respondents for subsequent statistical analysis (Hinkin, 1995). Finally, the instruments were originally designed in English; thus, the back translation method was used to avoid cultural bias and to ensure validity. The English versions were translated into Chinese by a university professor competent in both languages, and then the Chinese versions were translated back into English by a language professional, and special attention was paid to detecting misunderstandings due to translation.

Community trust

A six-item scale was adapted from Ridings et al.’s (2002) scale, which was initially used to assess the integrity and benevolence component of trust among members of a bulletin board. The items in this study were “I feel members of SCTNet have reciprocal trust and a reliable relationship.” “The members of SCTNet could usually fulfill the commitments made to others.” “The members of SCTNet would not knowingly do anything to disrupt the conversation of other members,” “The members of SCTNet behave in a consistent manner,” “The members of SCTNet treat others sincerely,” and “The members of SCTNet don’t take advantage of others even though they may have opportunities to do so.”
Altruism

This scale of 4 items was taken from Kankanhalli et al.’s (2005) enjoyment in helping others scale, and was used to evaluate individuals’ perceptions of the pleasure obtained from helping others through knowledge sharing within the community. The items were: “I enjoy helping others,” “I enjoy answering questions to help others,” “I enjoy helping others by sharing my knowledge through SCTNet,” and “Sharing my knowledge with others through SCTNet gives me pleasure.”

Knowledge sharing intention

We measured knowledge sharing intention using a 4-item scale modified from Moon and Kim (2001), which included three items initially used to elicit people’s behavior intentions of technology use (e.g., I will frequently use WWW in the future). We modified the scale’s items so that it was suitable for measuring the respondents’ intention to share knowledge via SCTNet. We added one additional item for assessing knowledge sharing because of the recommendation of a minimum of four items (Norman, 2010). The items are: “I am willing to share my knowledge with other members through SCTNet,” “I will use SCTNet on a regular basis to share knowledge in the future,” “I will strongly recommend that others use SCTNet to share knowledge,” and “It’s worthwhile to share knowledge on SCTNet.”

Knowledge sharing behavior

This scale with 4 items adapted from Yu, Lu, and Liu’s (2010) knowledge sharing behavior scale was used to measure people’s knowledge sharing behavior via a blog. We changed the target media to SCTNet. The items are: “I often share my knowledge on SCTNet,” “I often actively share my knowledge with other members on SCTNet,” “I often respond to other members’ comments on SCTNet,” and “I often participate and respond to different discussion topics on SCTNet.” We used a self-report scale of knowledge sharing behavior for three reasons: (1) all responses were anonymous meaning that we had difficulties collecting objective data from SCTNet; (2) knowledge contributors have a more accurate view of what kinds of knowledge they share than others; and (3) it is difficult to observe or count people’s knowledge sharing behavior from a website (Yu et al., 2010).

Control variables

As mentioned earlier, prior literature has suggested that perceived usefulness and perceived ease of use positively affect knowledge sharing (Hsu & Lin, 2008; Hung & Cheng, 2013; Reychav & Te’eni, 2009). To take these factors into account, we controlled for these two variables which were adapted from Davis’s (1989) research. The five items of the perceived usefulness scale are: “Using SCTNet for knowledge sharing would improve my work performance,” “Using SCTNet for knowledge sharing would be beneficial for learning professional knowledge,” “Using SCTNet for knowledge sharing would be beneficial for solving problems in my job,” “Using SCTNet in my job would enable me to accomplish tasks more quickly,” and “I would find SCTNet useful in my job.” The five items of the perceived ease of use scale are: “Learning to operate SCTNet would be easy for me,” “I would find it easy to share knowledge through SCTNet,” “Using SCTNet for knowledge sharing doesn’t take too much of my time or energy,” “I seldom require assistance when using SCTNet for knowledge sharing,” and “I would find SCTNet easy to use.”

Data analysis

The instruments in this study were validated employing several approaches. Internal consistency analysis was used to calculate the reliability of each scale. Factor analysis was used to examine the factorial validity of the scales used in this study. Next, we used correlation analysis to explore the relationships among the variables in this study. Hierarchical regression analysis was used to test the research hypotheses, i.e., the mediating effect and moderating effect.
Preliminary analysis

We used the principal component analytical method to examine the factorial validity of perceived usefulness, perceived ease of use, community trust, altruism, knowledge sharing intention, and knowledge sharing behavior. As Table 1 shows, the factor loading of each item was over the acceptable value of .70 (Tabachnick & Fidell, 2007), and the variances extracted were 67.41, 63.87, 69.11, 74.49, 75.44, and 86.53. In addition, the eigenvalues were 3.37, 3.19, 3.46, 2.98, 3.02, and 3.46. Table 1 also shows that the Cronbach’s α values for the instruments ranged from .85 to .95 which are above the acceptable level of .70 (DeVellis, 1991). With the acceptable factorial validity and reliability, we proceeded to explore the correlation among variables and to test the two research hypotheses. We also used confirmatory factor analysis to evaluate the Composite Reliability and Average Variance Extracted of the measurement models of altruism, knowledge sharing intention, and knowledge sharing behavior. The results indicated that the factor loading of each observed item was statistically significant and over .60, and the AVE values were .94, .94, .94, which are all over the suggested level of .50 (Anderson & Gerbing, 1988). In addition, the CR values were .67, .67, and .82, which all exceeded the acceptable level of .50 (Tabachnick & Fidell, 2007). With a well-fitting measurement model, we proceeded to test the proposed structural models.

Table 1. Factorial validity and Cronbach’s α

<table>
<thead>
<tr>
<th>Factor</th>
<th>Item no.</th>
<th>Factor loading</th>
<th>Eigenvalue</th>
<th>Variance Extracted</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived usefulness</td>
<td>item 1</td>
<td>.84</td>
<td>3.37</td>
<td>67.41</td>
<td>.88</td>
</tr>
<tr>
<td></td>
<td>item 2</td>
<td>.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 3</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 4</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 5</td>
<td>.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived ease of use</td>
<td>item 1</td>
<td>.82</td>
<td>3.19</td>
<td>63.87</td>
<td>.85</td>
</tr>
<tr>
<td></td>
<td>item 2</td>
<td>.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 3</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 4</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 5</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community trust</td>
<td>item 1</td>
<td>.82</td>
<td>3.46</td>
<td>69.11</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td>item 2</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 3</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 4</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 5</td>
<td>.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 6</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altruism</td>
<td>item 1</td>
<td>.77</td>
<td>2.98</td>
<td>74.49</td>
<td>.89</td>
</tr>
<tr>
<td></td>
<td>item 2</td>
<td>.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 3</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 4</td>
<td>.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge sharing intention</td>
<td>item 1</td>
<td>.84</td>
<td>3.02</td>
<td>75.44</td>
<td>.89</td>
</tr>
<tr>
<td></td>
<td>item 2</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 3</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 4</td>
<td>.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge sharing behavior</td>
<td>item 1</td>
<td>.91</td>
<td>3.46</td>
<td>86.53</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>item 2</td>
<td>.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 3</td>
<td>.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>item 4</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The means, standard deviations, and correlations among variables are presented in Table 2. Most of the zero-order correlations fell in the expected directions. For instance, knowledge sharing behavior was significantly correlated with community trust ($r = .49, p < .001$), altruism ($r = .60, p < .001$), and knowledge sharing intention ($r = .65, p < .001$), while knowledge sharing intention was significantly correlated with community trust ($r = .57, p < .001$) and altruism ($r = .66, p < .001$).
To test hypothesis 1, we argued that knowledge sharing intention mediates the relationships between community trust and knowledge sharing behavior. The mediation effect was tested following Baron and Kenny's (1986) three-step regression approach. First, the predictor (i.e., community trust) and the mediator (i.e., knowledge sharing intention) must be significantly related to the dependent variable (i.e., knowledge sharing behavior). Second, the predictor must be significantly related to the mediator. Finally, both the predictor and the mediator must be included in the regression model. When the predictor remains significant but the beta value is decreased when the mediator is included, this model represents a partial mediation model. On the contrary, if the beta value for the relationship between the predictor and the dependent variable becomes non-significant when the mediator is included, full mediation is present. As the correlation matrix (Table 2) shows, community trust, knowledge sharing intention, and knowledge sharing behavior are significantly related to each other. The results satisfy the first two criteria specified by Baron and Kenny. Below, we further test the mediation effect.

The results of the mediation analyses are presented in Table 3. In model 1, the control variables, perceived usefulness (β = .38, p < .001) and perceived ease of use (β = .25, p < .01) were significantly related to knowledge sharing intention. In model 3, perceived usefulness (β = .22, p < .001) and perceived ease of use (β = .18, p < .01) were significantly related to knowledge sharing. As models 2 and 4 indicate, after controlling for the control variables, community trust was found to be significantly related to knowledge sharing intention (β = .38, p < .001), and knowledge sharing behavior (β = .35, p < .001). Model 5 includes both the mediator and the predictor. The results indicate that the strength of the beta value for community trust decreased (β = .35 → β = .14) when knowledge sharing intention was included, but remained significant after controlling for the control variables. This result reveals that knowledge sharing intention partially mediates the relationship between community trust and knowledge sharing behavior. Furthermore, model 5 explains more variance (R2 = .44) in knowledge sharing behavior than model 3 (R2 = .20) and model 4 (R2 = .28), which indicates that it is meaningful to consider the influence of community trust and knowledge sharing intention on knowledge sharing behavior. Therefore, hypothesis 1 was supported.

Next, we argue that altruism moderates the relationship between community trust and knowledge sharing intention in Hypothesis 2. A variable is a moderator if the relationship between a predictor (community trust) and the outcome variable (knowledge sharing intention) is a function of the level of the moderator (altruism) (Howell, 2002). As Table 4 shows, altruism was introduced in model 3. After controlling for the control variables, the result indicates that altruism was significantly correlated with knowledge sharing intention (β = .46, p < .001), resulting in a 14% increase in R2 (F = 103.63, p < .001). In model 4, the interaction term is created after mean centering community

**Table 2.** Means, standard deviations, and correlations (n = 332)

<table>
<thead>
<tr>
<th>Variables</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Perceived usefulness</td>
<td>4.05</td>
<td>.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Perceived ease of use</td>
<td>3.95</td>
<td>.57</td>
<td>.55***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Community trust</td>
<td>3.78</td>
<td>.58</td>
<td>.53***</td>
<td>.41***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Altruism</td>
<td>3.97</td>
<td>.61</td>
<td>.44***</td>
<td>.41***</td>
<td>.50***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Knowledge sharing intention</td>
<td>3.97</td>
<td>.61</td>
<td>.52***</td>
<td>.46***</td>
<td>.57***</td>
<td>.66***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Knowledge sharing behavior</td>
<td>3.43</td>
<td>.84</td>
<td>.42***</td>
<td>.36***</td>
<td>.49***</td>
<td>.60***</td>
<td>.65***</td>
<td>-</td>
</tr>
</tbody>
</table>

***p < .001.

**Table 3.** Regression analysis for mediation

<table>
<thead>
<tr>
<th>Variables</th>
<th>DV = Knowledge sharing intention</th>
<th>DV = Knowledge sharing behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>.38***</td>
<td>.22***</td>
</tr>
<tr>
<td>Perceived ease of use</td>
<td>.25***</td>
<td>.18**</td>
</tr>
<tr>
<td>Predator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community trust</td>
<td>.38***</td>
<td>.35***</td>
</tr>
<tr>
<td>Mediator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge sharing intention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>.31</td>
<td>.41</td>
</tr>
<tr>
<td>F</td>
<td>74.99***</td>
<td>77.85***</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
trust and altruism (Aiken & West, 1991). As shown in Table 4, the interaction term was significantly associated with knowledge sharing intention ($\beta = .08, p = .034$), indicating that altruism moderated the relationship between community identity and knowledge sharing intention. Thus, Hypothesis 2 is supported.

**Table 4. Regression analysis for moderation**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived usefulness</td>
<td>.38***</td>
<td>.22***</td>
<td>.14**</td>
<td>.15***</td>
</tr>
<tr>
<td>Perceived ease of use</td>
<td>.25***</td>
<td>.18**</td>
<td>.09*</td>
<td>.09*</td>
</tr>
<tr>
<td>Predictor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community trust</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altruism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction term</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community trust × Altruism</td>
<td></td>
<td></td>
<td></td>
<td>.08*</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>.31</td>
<td>.41</td>
<td>.55</td>
<td>.56</td>
</tr>
<tr>
<td>F</td>
<td>76.74**</td>
<td>79.21***</td>
<td>103.63***</td>
<td>84.55***</td>
</tr>
</tbody>
</table>

*p < .05. ** p < .01. *** p < .001.

In Figure 3, to identify the nature of the interaction, the relationship between community trust and knowledge sharing intention is plotted for high and low altruism. We also conducted a simple slopes analysis (Aiken & West, 1991). The simple slope of the regression of community trust onto knowledge sharing intention within high altruism was significant (simple slope = .388, $t$ (335) = 6.955, $p < .001$). Within high altruism, the relationship between community trust and knowledge sharing intention was also significant (simple slope = .257, $t$ (335) = 4.960, $p < .001$). This result supports our hypothesis that the relationship between community trust and knowledge sharing intention was positive and stronger in the situation of a high level of altruism than it was with a low level of altruism.

![Figure 3. Simple regression lines for the effect of community trust on knowledge sharing intention at two levels of altruism](image)

**Discussion**

We tested the underlying mechanism that motivates the knowledge sharing of members within a virtual community of teacher professionals while taking into account the perceived technology support factors (i.e., perceived usefulness and perceived ease of use). Our findings suggest that community trust is fundamental to individuals’ knowledge
sharing intention, and therefore enhances their knowledge sharing behavior. Moreover, altruism moderated the relationship between community trust and knowledge sharing intention. These findings offer some important theoretical implications and practical suggestions for virtual communities of teacher professional development.

Theoretical implications

First, although scholars have introduced the social exchange theory to understand individuals’ motivational mechanisms behind knowledge sharing in the context of the Internet such as blogs (Chai & Kim, 2010; Hsu & Lin, 2008), online learning platforms (Ma & Yuen, 2011), and virtual communities (Chiu et al., 2006; Ridings et al., 2002; Wasko & Faraj, 2005), relatively little research has provided empirical evidence of the influence of community trust on knowledge sharing within virtual communities of teacher professionals. Our research reveals that community trust enables knowledge sharing behavior through knowledge sharing intention. The results contribute to the literature on knowledge sharing and virtual communities of teacher professionals.

Second, our results indicate that community trust was more strongly associated with knowledge sharing intention within the group of respondents who reported a high level of altruism than for those with a low level of altruism. This result implies that altruism acts as a potential facilitator of knowledge sharing in a virtual community of teacher professionals. Although prior literature has suggested that individuals who enjoy helping others are likely to be voluntarily motivated to respond to others’ needs (Kankanhalli et al., 2005), few studies have investigated the moderating role of altruism in the relationship between community trust and knowledge sharing intention in a virtual community. Therefore, these findings extend the current research on virtual communities by highlighting the important role of altruism in the relationship between community trust and knowledge sharing. Individuals who enjoy helping others in a trusting, reliable, and respectful community are likely to demonstrate a high level of knowledge sharing intention.

Practical suggestions

Rich knowledge from community members is a critical component of virtual communities. Teachers could promote their professional growth by engaging in the knowledge sharing process. Based on the social exchange theory, our study demonstrates that individuals are likely to share knowledge if they perceive a high level of trust and altruism toward the virtual community. Both trust and altruism (Brewer, 2003) are considered as types of social capital embedded in a social network (Granovetter, 1985). An institution or an organization could mobilize such social capital through effective management practices, meaning that community members’ trust and altruistic behaviors toward the virtual community could also be mobilized by adopting various strategies. From a practical standpoint, we offer several suggestions for improving community trust and altruism.

First, to encourage individuals’ knowledge sharing within a virtual community so as to promote the professional development of teachers, the community management teams need to establish an atmosphere of trust. As prior literature has suggested (Levin, Whitener, & Cross, 2006; Williams, 2001), community management teams could clearly define the community’s values and vision, explain to the community members the meaning of why sharing is important for teacher professional development, establish a shared understanding among members, and enrich the content and information in the community. When members realize the community’s value or recognize the community as an important platform for acquiring necessary information, they are likely to visit it often and gradually feel trust in the community.

Individuals’ justice perceptions in the virtual community can also lead to fewer conflicts among members and will thus enhance their willingness to share knowledge (Colquitt & Rodell, 2011; Fang & Chiu, 2010). The community management teams may need to design an appropriate mechanism for protecting the personal benefits of each individual in the knowledge sharing processes. When community members perceive that they could freely share personal experiences or discuss information related to teacher development without fear of criticism from other members, this would reduce the conflicts among community members and would be beneficial for engaging in knowledge sharing behavior.
Our results also suggest that altruism enhances the positive relationship between community trust and knowledge sharing intention. Thus, the second practical implication of our findings is to enhance teachers’ altruistic motivation and behaviors through different strategies. Individuals’ altruistic behavior could be improved through empathy training programs (Etxebarria et al., 1994; Grant & Berry, 2011), so that teacher education institutions could add such programs into existing teacher training programs. When individuals learn the ability to accept another person’s point of view, they may come to understand individual differences and realize the value of respecting others. This would enhance their helping behaviors.

Finally, although we controlled for the influences of technology support factors on knowledge sharing, technology itself may serve as an important facilitator in enhancing trust. We suggest that community designers demonstrate what others are doing using a visual representation scheme. As community members efficiently and accurately recognize what others are doing and sharing, this may motivate their trust in the community.

Limitations and future directions

Although our results contribute to the literature, there are still several limitations which point to directions for shaping future research. First, the research on knowledge sharing in online communities is interdisciplinary, so there are various taxonomies and multi-domain studies based on different theories such as personality theory, motivation theory, social cognitive theory, game theory, and system design. Research based on these different theories could provide implications for knowledge sharing and virtual communities. Thus, we suggest that future research be based on a variety of theories to identify factors (e.g., personality or contextual factors) that influence knowledge sharing, or that it explore factors that motivate altruistic behavior such as task characteristics or personality types. Furthermore, we employed self-report scales in our study, which may have some limitations in terms of the inferences which could be made. We suggest that future research use different research designs (e.g., longitudinal design or experimental design) or analytical techniques (e.g., social network analysis) to identify stronger causal inferences. Finally, we provide practical implications. The suggestions are considered as important antecedents of community trust and altruism by the prior literature. However, we still cannot be certain how people perceive community trust and act altruistically because we did not test the antecedents. Thus, we encourage future research to examine the antecedents of community trust and altruism.

Conclusions

The value of a virtual community of teacher professionals is created with the rich knowledge of community members. Engaging in the knowledge sharing process can enhance teachers’ professional development and teaching quality. Teachers’ knowledge sharing intention and behaviors can be promoted in many ways. By introducing the social exchange theory indicating that social exchange relationships support the knowledge sharing process, we found that teachers who report higher community trust have higher willingness to share knowledge, which in turn increases the likelihood of knowledge sharing behavior. We also found that altruism augments the role of community trust in knowledge sharing intention. Together, these results enrich our understanding of the mechanisms by which community trust and altruism can shape the knowledge sharing processes of teachers in a virtual professional community.

Acknowledgements

The authors would like to thank the two anonymous reviewers for their thoughtful comments and Sufen Chen for her valuable suggestions. Our special thanks also go to Chien-fang Huang for her assistance in collecting data. This work was financially supported by National Science Council, Taiwan (NSC 96-2511-S-011-002-MY3 and NSC 98-2511-S-011-003).

References


Visualisation in Applied Learning Contexts: A Review

Adrian Twissell

School of Education, Oxford Brookes University, Oxford, UK // atwissell@brookes.ac.uk

(Submitted June 12, 2013; Revised October 21, 2013; Accepted January 16, 2014)

ABSTRACT

This literature review explores visualisation within the context of learning in design, engineering and technology education. The investigation first defines visualisation, providing examples of activities that utilise visualisation skills within an applied field. Then exploration of the mental mechanisms of visualisation used to engage with those activities is placed within the context of learning in applied fields. The discussion leads to consideration of the role of visualisation in relation to learning abstract technological concepts. The paper concludes that the application of particular visualisation mechanisms is dependent on task and context. Approaches to instruction and the design of learning materials are also considered within the context of visualisation. Finally the review of literature highlights a number of research questions which relate to the individual differences and socio-cultural aspects of learning and using visualisation.

Keywords

Cognition, Visualisation, Engineering and technology, Technological concepts, Thinking skills

Introduction

This paper explores visualisation within the context of learning in design, engineering and technology education. These fields can be defined as ‘applied’ learning contexts, as they apply knowledge and understanding in ways that require a physical relationship with a practical, material vehicle providing a focus for learning through doing. Visualisation has been noted as supporting creative thinking (Arnheim, 1970), visual thinking in design education (McKim, 1978) and thinking within science and technology education (Mathewson, 1999). It is suggested that visualisation provides a holistic, “synchronous” (Paivio, 1986, p. 60) approach to thinking, the benefits of which are absent from logical-mathematical and verbal thinking styles (McKim, 1978; Mathewson, 1999). Visualisation ability has been found to indicate the possible future success of students in STEM (Science, Technology, Engineering and Maths) careers (Lohman & Lakin, 2006; Webb, Lubinski & Benbow, 2007) and within the context of further education, has been described as an essential aid to learning about relationships between theory and practise (Haight, 2012). In addition, (Gardner, 1993) has revealed strong individual differences in thinking styles, with some employing largely linguistic methods (e.g., Freud) and others visual-spatial and logical-mathematical means (e.g., Einstein), suggesting that individuals adopt an affinity with certain mental mechanisms in their thinking.

An exploration of KS3 (Key Stage Three in England) students’ logical-mathematical, verbal and non-verbal MidYIS (Middle Years Information System) test scores, revealed that spatial ability may be one indicator of high ability within technology education, when compared with performance on subject specific outcome measures at secondary school (Twissell, 2011). Personal experiential evidence suggests that students’ ability to visualise and use image based systems of representation promotes improved subject specific thinking and learning through an improved ability to make effective inferences from those representations (Larkin & Simon, 1987). The specific mental mechanisms involved in visualisation would therefore appear to warrant further investigation as to the efficacy of learning and teaching in this mode and the potential benefits to students’ learning outcomes.

This investigation first defines visualisation, providing examples of activities that utilise visualisation skills within an applied field. Then exploration of the mental mechanisms of visualisation used to engage with those activities is placed within the context of learning in applied fields. The review of literature is used to identify future research opportunities, emerging patterns, uncertainties and gaps in knowledge (Robson, 2011).

Defining visual-spatial thinking

Visual perception represents a main pathway to the world of experience. Visual information received by the eye is transferred to the occipital cortex and interpreted by the visual association area; here meaning is attached to the
received visual stimulus (Martini, Nash & Bartholomew, 2012). Visualisation however is conceived in the literature as a mental process. Hoffler (2010) for example defines visualisation as “any kind of non-verbal illustration (both symbolic, such as graphs, and pictorial, such as realistic diagrams, pictures, or animations)” (p. 246). Lohman (1993) defines “the ability to generate, retain, retrieve, and transform well-structured visual images” (p. 3). Kosslyn (2005) makes a useful distinction between visual perception (viewing a stimulus) and visual mental imagery (an internal process of visualisation drawing on memory in the absence of a stimulus). Van Garderen (2006) posits “visual imagery” (object representation: shape and colour) and “spatial imagery” (spatial relationships between parts of objects, their spatial location and movement) (p. 497). Visual-spatial thinking has been linked with memory as a mechanism to combine perceiving and visualisation processes, which are thought to aid the rapid processing of information (Gegenfurtner, Lehtinen & Saljo, 2011; Mathewson, 1999; Smith, Ritzhaupt & Tjoe, 2010). Thus a distinction emerges between perceiving and processing images, and constructing images as part of a process of cognition in conjunction with memory. In this paper I refer to viewing an external stimulus as perception and thinking with images, their spatial relation and transformation, as visualisation.

Applying visualisation to design, engineering and technology

Within applied fields, visualisation can be conceived in a number of ways, dependent upon the context of the thinking task. For example producing a conceptual design for a new electronic product requires the designer to imagine the form and features of the product, often without an external visual stimulus. The designer may mentally rotate the design during the exercise, imagining its facets from different viewpoints. Preparation of the concept for manufacture may involve mentally transforming the design from one image type (three dimensional pictorial view) to another (two dimensional plans and elevations), or the designer may imagine a cross section chosen to show hidden detail. Alternatively designing may take the form of technological development and involve visual thinking associated with the application of abstract technological concepts during problem solving. Here concepts may be represented by visual analogy and metaphor, employed to aid the manipulation of, for example, electronic system concepts. Thus a number of visual mechanisms can be identified which aid the designer’s task, including: visual imagination, image rotation, image transformation, visual analogy and visual metaphor. Technology education in particular often draws upon several, if not all, of these mechanisms in the development of product designs and their realisation. In the following section each mechanism is explored in relation to its application to design and the process of learning using these modes of thinking.

Mechanisms of visual thinking

Spatial relation, rotation and transformation

Hoffler’s (2010) meta-analysis identifies two commonly explored factors: spatial visualisation and spatial relation, describing the mental rotation and object relation ability of individuals accordingly. Zacks (2008), exploring neuroimaging, supports the concept of analogue spatial representations, which describe the mental representation and rotation of objects by the brain, as if they were being physically rotated. Correlation between degree of rotation on an experimental task and amount of brain activity is reported to support the use of spatial analogues by the brain to perform these tasks (Zacks, 2008). It has been suggested that these spatial abilities represent innate cognitive processes (Shepard, 1978). However research suggests that the interpretation of visual information is dependent upon early childhood experiences which develop cognitive mechanisms to support that interpretation (Eysenck, 1996). This may include the early play experiences of the child (Vygotsky, 1978) and particularly play with physical objects (Shepard, 1978), in the above analogue representation example (Zacks, 2008).

Gardner (1984) suggests that these abilities represent a separate intelligence, within his multiple intelligence (MI) theory, which draws on experimental psychology and neurology. A range of visualisation abilities are posited which are considered to reinforce one another and work together, such as spatial relation and object rotation. Figure 1 shows a typical spatial relation task which draws on object recognition skills to compare a variety of figures with a target form.
Shepard and Metzler’s (1971) experiment (Figure 2) develops this task, the mental rotation activity introducing the use of imagination to create the new object view when rotated through the picture plane. Participants mentally rotate the object to check whether both items match; thus in ‘c’ the test items do not match, a conclusion reached following the rotation of the left hand figure and comparing (spatially relating) this mental view with the figure on the right.

However using neuroimaging to monitor brain activity, Zacks (2008) suggests that this type of study indicates only where brain activity can be observed, rather than an indication of specifically what the brain is doing. Zacks’ (2008) discussion supports the assertion that mental rotation is enhanced by corresponding manual rotation performed in the same direction, linking mental and sensori-motor facets of brain activity; and supporting the commonly held belief that seeing and doing combine to enhance individuals’ learning experiences.

Within their stage theory of child development, Piaget and Inhelder (1971) believe that the ability to anticipate or imagine a new object view arises only after the child has developed an understanding of object permanence, the gradual discovery of permanence within the environment occurring during the first, sensori-motor stage of development. Piaget and Inhelder (1971) classify imagery in two ways: reproductive images representing events or objects known to the viewer and anticipatory images representing imagined events not yet experienced, but constructed by the viewer. This classification is supported by a framework describing the development of imagery within the distinct age related stages; thus as children advance through the stages, they gradually develop the ability for thought beyond the use of perception and static images (pre-operational stage) and begin to perform (around the beginning of the concrete operations stage) mental transformations and use a developing imaginal ability (anticipation). Thus mentally rotating a figure such as that shown in Figure 2 is only possible following a period of mental maturity. Criticism of this view (Bruner, 1977) however suggests that, although a number of stages can be observed, individual differences and experiences represent a more significant role than stage theorists suggest.
Memory is closely linked with experience and has been posited as an important component of visualisation (Williams, 2012). Memory has also been considered to make a significant contribution to the process of object recognition (Bruner & Postman, 1947) and the creation of new visual imagery in conjunction with stored experiences during child’s play (Vygotsky, 1978). Some researchers have reported, however, that KS3 students often fixate on known imagery during the design process (Nicholl & McLellan, 2007), revealing a negative focus in this context. Others (Newcomb, 2007) have found that visualisation strategies (orthographic drawing) improve students’ spatial ability and problem solving skills, beyond memory and direct observation. Sibbet (2008), a proponent of visualisation as “process,” that is using universal patterns of visual representation in conjunction with experience, describes intelligence as the ability to use memory to draw on these patterns during visualisation and make “inferences” toward an “appropriate action” (p. 124) on the basis of these constructed patterns. Thus mentally rotating an object may facilitate several sequential thought processes, beginning with object recognition, the creation of a mental analogue and its commitment to memory, rotation of the analogue and finally anticipation of the new, unseen object image.

Transformation

Kosslyn’s (2005) neuroimaging studies have shown that brain activity responsible for perceiving is also largely responsible for visualisation. Kosslyn’s (2005) theory locates visualisation within a connected “visual buffer” (p. 336) which uses multiple areas of the brain to process information related to visual thinking. However as Kosslyn (2005) concedes, it is not clear from results precisely when and how mental images are used to aid cognition. In a related study (Thompson et al., 2009), progress toward this anomaly suggests that two forms of visual mental activity can be identified, on the basis of neuroimaging experiments, including spatial relation processing and spatial transformation processing each undertaken by different parts of the brain. This, it is suggested, has important implications for specific skills training in particular domains, such as “navigating an environment” and “learning surgical techniques,” which draw on these two abilities to varying extents (Thompson et al., 2009, p. 1252). In general, researchers agree that “like-modality perception” (i.e., visual, auditory, motor) is strongly connected with visual mental imagery, which in turn is used in conjunction with the body’s motor and affective systems (Kosslyn, Ganis & Thompson, 2001, p. 641).

In applied fields spatial relationship and spatial transformation abilities are commonly used in the preparation of manufacturing drawings. This often necessitates the transformation of visual information from one mode of representation such as a three dimensional (3D) external view, to an alternative two dimensional (2D) view of one side of the object, or a cross section showing part of the object as if it were “cut” away. Thus transformation requires an additional ability, beyond object rotation, which draws on imagination to “see” or anticipate the object from a different viewpoint and represent it in a different way. Figure 3 illustrates such a task as a test item, which requires the learner to visualise the cutting plane intersection (right hand image) to allow hidden or otherwise awkward to reveal details to be shown (within the left hand ‘Pointed Bracket’ model).

![Figure 3](image.png)

*Figure 3. Example spatial relation/transformation exercise (“Pointed Bracket” after Hsi et al., 1997, p. 155)*
Piaget and Inhelder’s (1971) stage theory would suggest that a series of developmental steps be completed before a learner is able to successfully perform the task in Figure 3. Transformation ability widens the range of representational forms used by the learner, beyond those associated with object recognition and rotation. The sequential relationship between 2D and 3D representations has been explored by Akasah and Alias (2010), within an engineering student skill-development context. Instruction beginning with 3D forms, rather than 2D, was found to accelerate students’ visualisation skills due to the existence of a more recognisable referent during mental transformations (Akasah & Alias, 2010). The 3D to 2D visual sequence model has also been found to support students’ conceptual learning in physics and mathematics (Basson, 2002). However this approach, the “whole-to-parts” (3D to 2D) rather than “parts-to-whole” (2D to 3D) approach might be predicted to yield accelerated learning, due to its reduced demand on visualisation skills (specifically visual rotation and transformation).

So far visualisation has been conceived as a number of mechanisms which may be an innate part of human physiology, a separate “intelligence,” but which have also been shown to involve an element of learning and development (Piaget & Inhelder, 1971). Clearly well-formed visualisation abilities, such as those required for the tasks in Figures 1 and 2, support the designer producing a new product concept. However Gardner (2003) has suggested that the concept of individual intelligences needs revising, on the basis of new genetic and neuroscientific research, which indicates a conception based on wider cognitive interconnectivity. The next section considers the role of verbal reasoning in connection with visualisation as a multi-modal form of cognition.

**Verbal and non-verbal processing - dual coding theory**

Blakemore and Frith’s (2005) overview of neuroimaging and brain functioning, focuses on the connection between visualisation and verbal learning, particularly in support of learning the meaning of concrete and abstract words. Drawing on research using brain injured patients, Blakemore and Frith (2005) support the notion that imagery and language areas of the brain usually conspire, in conjunction with memory, to aid object and image based recognition. Thus neuroscience is suggesting that recognition and recall of concrete words makes use of visualisation, in support of dual coding theory (DCT) (Paivio, 1978). Conversely abstract words draw on alternative language based regions of the brain, which may also include auditory processing (Blakemore and Frith, 2005).

However Arnheim (1970), an opponent of this view, believes visualisation occupies a more prominent role within cognition, which has as its basis perception leading to thought processes grounded in imagery. Arnheim (1970) makes a powerful claim for visual thinking, believing that without it “productive thinking is impossible in any field of endeavour” (p. 3). Verbal language is viewed as a “one dimensional sequence” (Arnheim, 1970, p. 232), not capable of the type of manipulation needed to deal with complex reasoning and problem solving; its origin considered to follow closely perceptual experience. Nevertheless language (both auditory and sign codes) is acknowledged as providing a number of useful functions, including: categorical naming, differing levels of abstractness, thought organisation and the ability to attach a communicable code to an object or event (Arnheim, 1970).

DCT (Paivio, 1978) on the other hand posits a strong psychological relationship between verbal and nonverbal processes, which are considered to be separate, but associated functions of thinking which aid memory in encoding information in different ways. DCT conceptualises the representation and processing of information in two ways (Paivio, 1978; Paivio, 1986; Clark & Paivio, 1991). Firstly words (sequentially processed arbitrary symbols) and images (synchronously processed visual imagery) are represented as codes which form associative networks; that is, they associate through the referential process of naming images and representing words with images (Clark & Paivio, 1991). Secondly associative connections make links within the coding representations, such as connecting words with other associated words (or events or emotions) and images with other visual imagery related to the image, object, sound, smell or previous experience.

Two key features of DCT emerge. First, and importantly for the student/designer who may be working with physical models, is the acknowledgement of wide ranging sensory input (visual, auditory, tactile and kinaesthetic) and the role of emotion during the thinking process. Clark and Paivio (1991) describe a process of arousal which affects the prominence of any representation and its subsequent associative potential. This also includes the influence of past experience, representing a constructivist perspective for the theory which indicates that individual experiences can affect the processing and storage of the same sensory input among individuals who may recall these experiences in
different ways. An example may include different individuals learning or performing a task in different environments, who may consequently attach different emotions to the task and its recall. Second is the facility for imagery to represent multi object perspectives, which can aid memory and their potential for “dynamic spatial transformations,” providing the facility for imaginary thinking not possible with words alone (Clark & Paivio, 1991, p. 152). This personalised approach to thinking and the application of images and words is illustrated by Piaget and Inhelder (1971) with their number sequence representations and Hoffler’s (2010) counting strategy example.

However Randhawa (1978) questions the efficacy of DCT to adequately explain specifically how verbal or nonverbal systems relate to cognition. Shepard (1978) questions verbal means of cognition as an adequate mode of representation for, or thought about, complex ideas during the learning process suggesting that it maintains “established ideas and entrenched traditions” (p. 156). In addition Steiner (1974), using multiple ordering tasks, found that training using an “ikonic” (p. 892) medium (wooden blocks of varying sizes and colour), as opposed to a symbolic medium (words representing the blocks: “grey,” “tiny” (Steiner, 1974, p. 897)), significantly improved performance, possibly due to the clarity of the visual material which avoided the need for interpretation through reading and verbalisation. Nevertheless a recent study by Griffin and Robinson (2005), claims support for DCT, on the basis of an exploration into “visuality” and “spatiality” (p. 24). Their study found that icons (visual representations) were more effective at encouraging recall of textual features, when compared with recall based upon spatial positioning (spatiality). Consequently Griffin and Robinson (2005) conclude that text recall makes most effective use of visually iconic (mimetic) information, rather than spatially related information and thus support the concept of DCT, within the context of Geography education.

Symbols, codes and multimodal thinking

The use of recognisable referents in the form of symbol systems has been suggested as important to spatial ability; for example allowing the child who, able to negotiate the environment using well-formed spatial skills during school age years, finds it very difficult to apply symbol systems to recreate or communicate that environment (Gardner, 1984; Vygotsky, 1978). This indicates a link between learning and the development of linguistic and symbolised modes of thinking in relation to the use of spatial skills; thus support for multimodal thinking. Relatedly Piaget and Inhelder (1971) propose a “system of imaginal symbols” (p. 381); for thought to occur, be internally processed and communicated, personal images (imaginal symbols) are created to represent the child’s words. These symbols are then used as referents to replace verbal means when memorising, evoking and thinking about a perception or event. In applied learning contexts recognisable referents often take the form of symbols and codes in addition to those used verbally. For example learning about technological concepts, such as the electronics example in Figure 4, involves understanding meaning represented by electronic component symbols, which are in turn used to aid spatial understanding about current flow.

Hoffman (2012) describes learning generally as “the development of cognitive systems [which] depend on signs and representations as mediators” (p. 185) and which in turn are facilitators of communicable knowledge, personal to the individual. Signs and representations have been considered to evolve, from signals to symbols, and through socio-cultural complexity, to complex grammar and language (Sinha, 2004). Therefore visual thinking within this view is closely associated with the socio-cultural practises of communication structures and the construction of personal knowledge, concurring with Piaget and Inhelder (1971). Bruner’s (1977) conception of learning specifically links the use of media symbols (ikonic and symbolic) with cognitive processing as a means to construct and facilitate new knowledge. Hoffman’s (2012) discussion makes a persuasive case for the tactile engagement with “concrete objects and representations” (p. 193) in the development of cognitive ability, within a mathematics learning context, reinforcing the use of concrete symbols to aid cognition. Thus this perspective closely interlinks perception, visual thinking using symbols and learning through the manipulation of the symbols (Bruner, 1977), imagery and connected verbal representations. Continuing the example in Figure 4, this equates to recognition of the symbols, the use of the symbols to mediate the behaviour of the electronic circuit and the use of spatial ability to visualise the relationship between the system’s input, process and output.

Developing a technological understanding, particularly in electronics, often necessitates an understanding about mathematics, which uses its own taxonomy of codes and symbols accompanying their verbal counterparts. Developing mathematical competency begins with the child’s understanding of object permanence, evolving into the use of complex symbols to enable reasoning with abstract concepts (Eysenck, 1996; Gardner, 1984). Research in this
field has considered object recognition, perceptual organisation and structural representation, identifying student difficulties with movement between these phases and beyond object recognition, when problem solving (Gal & Linchevski, 2010). Assel et al. (2003) found that children’s early experience with visual-spatial problem solving through play enhances later mathematical ability. Bobis (2008) describes the use of “subitising,” (p. 6) that is the recognition of spatial structure (features and relationships) utilising visual means to calculate the number of items within an object arrangement, thereby reducing the need for a mental calculation. Here verbal codes are also suggested as supporting the child’s construction of “pattern-name associations” (ibid) which increase the speed of cognitive processing. Thus this line of reasoning would suggest that visualisation is a means to link quantity with instruction directing its manipulation, within a system of symbolic representations.

van Garderen’s (2006) and Edens and Potter’s (2008) studies make a distinction between pictorial and schematic imagery within mathematics education. The studies found that high mathematical achievers held a preference for using schematic imagery (spatial relationships encoded within the imagery), while low ability achievers preferred pictorial imagery (visual appearance encoded within the imagery). Thus visualisers focused on either the appearance of the objects within the problem situation (pictorial) or the spatial relation and overall concepts within the problem (schematic). Positive correlation was identified between the use of schematic imagery and successful mathematics problem solving. Within electronics education Chen et al. (2011) exploit these alternative symbols, using both external representations (pictorial) and concept models (schematic) and conclude that combining these into one simulated visualisation enhances the learning of abstract concepts as the learner can more easily “verify and clarify the existing knowledge” (p. 269) using both familiar referents and symbolic representations.

In summary the use of codes and symbols appears to improve the speed of cognitive processing, beyond that which may be possible with words and images alone. They may be applied to allow thought about complex and abstract concepts and their communication to others. The choice of code or symbol may depend upon context and the type of knowledge being manipulated. In many cases, such diagrammatic illustrations, symbolised forms of representation are quicker to recall and explain (Larkin & Simon, 1987).

**Analogy and metaphor**

Abstract technological concepts, such as that shown in Figure 4, form an integral part of the work of designers in many applied fields. Thought about these concepts is enhanced with the use of visual analogy (Petrucci, 2011), which allows a visual likeness or explanation in combination with verbal means to represent the complex abstract concept; verbal means alone often not allowing adequate engagement with concepts or facilitating the correction of misunderstandings. Figure 4 represents an analogy which depicts the behaviour of an electronic capacitor using hydraulics. Interpretation requires an understanding of the symbols used, draws on both spatial relation and transformation ability, but crucially offers an explanation difficult to provide verbally.

*Figure 4. Example visual analogy using hydraulics (after Hughes and Smith, 1995, p. 90)*

186
Similarly Mathewson (1999), exploring visualisation within a scientific and technical context, suggests “higher-order visual spatial thinking is inherently analogic” (p. 38) and relies on mental representations and comparisons which often arouse visual metaphors to enhance their potential. Thus in the hydraulic analogy (Figure 4) the metaphor ‘pressure’ might be used to replace ‘current flow’ to explain more successfully how an electronic capacitor charges and discharges. In another way, Figure 4, a “graphic” metaphor, allows the learner to connect with “foundation” metaphors (Sibbet, 2008, p. 122); early experiences which are used for future interpretations, in this case the early experience of the behaviour of liquids.

Geake (2008) captures the essence of the discussion around cognitive functioning in research on “fluid analogizing” (p. 187) in relation to the presentation of new visual stimuli. While recognising the role of perception in analogy generation, Geake (2008) supports the concept of fluid analogizing which describes the brain’s ability to draw upon “myriad functional modules,” (p. 191) with working memory acting as a “dynamic workspace” (ibid) for communication between modules, leading to divergent thinking beyond concrete solutions to problems.

Visualisation in applied learning contexts

Research within applied learning contexts has focused on the trainability of visualisation, often within engineering as a means to improve 3D spatial awareness (Potter et al., 2009). Improved interaction and personal learning of abstract engineering concepts has been reported, through ICT facilitated graphics (Nguyen & Khoo, 2010). Hsi, Linn and Bell (1997) report the use of specific problem solving strategies by students, which relate to domain specific instruments (orthographic projection, isometric views and section drawings) and parallel Mathewson’s (1999) visualisation strategies. Pulé and McCardle (2010) suggest the use of strategy can be linked with a preferred learning style (visualisation method preference) and memory, on the basis of research within the context of electronics education and the use of symbol systems. Chen et al. (2011), also exploring learning within electronics, believe a visualisation based learning model which incorporates ICT facilitated graphic manipulation and a reflective stage enhances conceptual learning and achieves a “higher level of cognition” (p. 269) on the basis of verification and clarification of existing knowledge. Conjointly these studies support the notion that visualisation skills can be acquired, honed and developed, and are not “fixed or culturally exclusive abilities, but respond to instruction and mediation” (Potter et al., 2009, p. 109).

Discussion

Researchers have highlighted the significance of perception and memory in the child’s early years of cognitive development (Piaget & Inhelder, 1971), but the existence of distinct stages has been questioned with an alternative, individualistic viewpoint proposed on the basis of environmental influences and differing methods of representation used by children to develop their thinking (Bruner, 1977). The socio-cultural context of these processes has been considered, emphasising a link between an individual’s learning and the context of the learning (Vygotsky, 2004) and the interactions made between the participants within the learning context (Fler & Richardson, 2008). Consequently here a constructivist, socio-cultural perspective is accepted as underpinning the development of visualisation and its application to learning.

The review reveals a number of mechanisms which aid visual cognition and a number of more general learning concepts. The overlapping nature of some of the terms and concepts is particularly notable, perhaps reflecting the contextual relationship between mental mechanisms and use of terminology. The wide range of visualisation conceptions may also indicate a relationship between concept and learning context. Thus spatial relation and mental rotation ability are clearly of benefit to students of engineering, while analogy and metaphor and the use of icons and symbols, may enhance the understanding of technological concepts. Figure 5 provides an example of an abstract concept represented in three ways, taken from electronics education. Each symbol system has its own method of representation and consequently involves different modes of thinking in its understanding of an electronic logic system. Further exploration may explain how individuals construct an understanding of such a concept in this context, how imagery is used across the three symbol systems and the relative use of different modes such as verbal language, icons and symbols within each coded system of representation; the translation from one mode to another perhaps reinforcing conceptual connections through reinterpretation (Suh & Moyer-Packenham, 2007). This type of visualisation ability appears to link with individual learning and the construction of personal knowledge systems,
where it has been shown that preferences include pictorial, animated or schematic representations for thinking about complex abstract concepts (van Garderen, 2006; Edens & Potter, 2008).

<table>
<thead>
<tr>
<th>Circuit Symbol</th>
<th>Truth Table</th>
<th>Boolean Algebra</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

$F = A \cdot B$

*Figure 5. AND gate represented by three symbol systems*

DCT provides a model for thinking about how individuals achieve the interaction between verbal and visual means, including the utility of iconic and symbolic imagery, to aid the construction of personal knowledge, analogy and metaphor. Further exploration may develop an understanding of the relative relationship between concreteness, abstraction and imagery outlined by Clark and Paivio (1991), which may be explained with greater clarity on the basis of empirical evidence from applied field contexts.

**Conclusions and further research**

Two key areas have been discussed: the mechanisms of visual thinking and their application to learning in applied fields. These have emerged from a body of research which in general posits the positive role of visualisation as a means to enhance learning, problem solve, memorise and recall information, represent and communicate. These skills can be fostered (Potter et al., 2009) and the design of instruction in visualisation should aim to reinforce the mechanisms which support the process of visual thinking. Wu, Krajcik and Soloway (2001) offer a useful framework for thinking about instruction, based around three levels of representation. These include macroscopic (observable-including engagement through practical action), microscopic (atoms, particles-abstract and invisible elements) and symbolic (symbols, numbers, formulas and equations) representations providing an approach to the design of instruction which, when each level is satisfied, may contribute to increased understanding through the development of a “higher-order” visual metaphor (Mathewson, 1999, p. 38). In addition, offering opportunities for students to translate between representations has been shown to reinforce conceptual development (Suh & Moyer-Packenham, 2007). Similarly, the design of instructional materials may also benefit from an approach which includes multi-modal representations, encouraging engagement with and translation between knowledge presented in different forms, particularly in light of research which suggests multiple representations are of benefit to students’ conceptual understanding (Ainsworth, 2006) and the importance of visualisation as a processing mechanism (Wu et al., 2001). This review also reveals the significance of practical application i.e., sensori-motor engagement, to learning using visualisation which has been shown to enhance visualisation processes (Kosslyn et al., 2001; Shepard, 1978; Vygotsky, 1978; Zacks, 2008).

A number of research questions emerge from the review, which may be of importance to educators in improving subject specific thinking and learning. For example how specifically do students construct their understanding of abstract technological concepts using the visual means available to them? What is the relative use of icon, symbol, code, analogy and metaphor when constructing these concepts? How do individuals construct these concepts differently?

Developing Paivio’s (1978) associative networks and associative connections concepts, four specific research questions emerge in relation to each, within applied learning contexts:

- How does image attachment vary among individuals when thinking and learning about abstract technological concepts?
- What is the nature of any connected foundation metaphor?
- What visual links do individuals make while translating between the elements of abstract technological concepts when developing their understanding of these?
- How do individuals construct their personal understanding, using additional means such as verbal and sensori-motor modes, within applied learning contexts?
A significant gap in the literature relates to the individual and social context of learning using visualisation. Drawing on Bruner (1977) and Vygotsky (1978) respectively, these two themes warrant further investigation in relation to individual differences and the influence of environmental factors on learning, including the effects of socio-cultural interactions (Fleer & Richardson, 2008). Answers to these questions may support the development of curriculum materials and approaches to learning which better support individuals during their learning of complex technological concepts.

Acknowledgements

I am very grateful to Ms Georgina Glenny, Senior Lecturer in Education at Oxford Brookes University, Oxford, UK for comments on an earlier draft of this manuscript.

References


190


A Model for the Design of Puzzle-based Games Including Virtual and Physical Objects

Javier Melero* and Davinia Hernández-Leo

Universitat Pompeu Fabra, Barcelona, Spain // javier.melero@upf.edu // davinia.hernandez@upf.edu

*Corresponding author

(Submitted July 30, 2013; Revised October 2, 2013; Accepted December 12, 2013)

ABSTRACT

Multiple evidences in the Technology-Enhanced Learning domain indicate that Game-Based Learning can lead to positive effects in students’ performance and motivation. Educational games can be completely virtual or can combine the use of physical objects or spaces in the real world. However, the potential effectiveness of these approaches largely depends on the pedagogical design behind the game and to what extent this design is aligned with the requirements of specific educational situations. This paper focuses on puzzle-based games, as a particular strategy that can foster students’ problem solving, analytical and memory skills, and on the role of teachers as designers of meaningful games. The paper presents a model (conceptual model and XML binding) to represent puzzle-based games including virtual and physical elements. The expressiveness of the model is shown with several cases to illustrate that the model covers a wide range of significant puzzle-based games. The results from an exploratory user study show that, except two elements of the model that require revision, teachers understand the model and can apply it to solve design tasks.

Keywords

Game-based learning, Puzzle games, Game design, Conceptual model, Information model, Exploratory study

Introduction

Educational games are being backed up in the Technology-Enhanced Learning domain as strategies that can lead to worthy learning outcomes. These games emerge as an option to address the learning and motivation requirements of the current generation of students (Prensky, 2001). Empirical evidences support the positive effects of computer games as learning tools. The evidences indicate that games implementing pedagogical designs can strengthen and support school achievement, cognitive abilities, motivation towards learning, reflection, attention and concentration (Jenkins, 2002; McFarlane et al., 2002; Bottino et al., 2007). Furthermore, the nature of educational games is very varied, from videogames to mainstream games (Mitchell & Savill-Smith, 2004). Also, the majority of educational games are completely virtual, allowing students to interact with virtual representations of concepts that are difficult to access in the real world (Melero et al., 2011). However, there are other relevant games that include the use of physical objects (with tangible interfaces or embedding sensing technologies) that address specific educational needs (e.g., manipulation of physical objects vs. symbolic representations, physical interaction for particular therapies, etc.) (Li et al., 2008). Moreover, other games are played in physical spaces considering real objects and their locations to facilitate contextualized learning (Avouris & Yiannoutsou, 2012).

Despite the potential benefits of educational games, teachers do not broadly adopt them in formal learning settings (Williamson, 2009). The reasons behind the low adoption include that the available games do not often fulfill the requirements of particular educational situations, and that teachers do not have advanced technological skills to create or adapt their own educational games (Melero et al., 2011; Tornero et al., 2010; Yang, 2005). Diverse research efforts are being devoted to provide easy-to-use game editors, such as <e-Adventure> (Torrente et al., 2010), Alice (Conway et al., 2000), Squeak (Ingalls et al., 1997), and GameMaker (Overmars, 2004). These authoring tools have been developed to allow teachers to design educational games. Unfortunately, these tools can be still too complex for some instructors (Tornero et al., 2010), hard to adapt to individual courses and require much time for development (Yang, 2005; Tornero et al., 2010).

This paper focuses on puzzle-based games as a particular educational strategy to feasibly involve teachers as designers of the games. The nature and duration of this type of games is typically equivalent to other types of learning tasks for the classroom or field trips activities (Michalewicz & Michalewicz, 2008; Falkner et al., 2010). Indeed, the proposed games could be used in laboratory sessions, as a complement to traditional approaches. But also, students could play these games as a homework assignment to reinforce the concepts learned in classroom.
Besides, teachers can use puzzle-based games to engage their students in the subject topics, while at the same time foster students’ problem solving, analytical and memory skills (Huang et al., 2007; Bottino et al., 2007). Puzzle games have simple rules around the basis of a challenge that requires interrelating pieces over a given board (Huang et al., 2007). Their rules can be defined independently from content and clues (Crawford, 1982) and, therefore, they can be applied to multiple subject matters. Besides, the concept of puzzle games is not tied to a specific platform or type of technology. Indeed, current advances in technologies and the generic characteristics of puzzle-based games enable the consideration of virtual and physical objects in the design of the puzzle game (as pieces or slots).

Therefore, this paper proposes a model to support the design and computational representation of puzzle-based games including virtual and physical objects. The main aim is to enable teachers the design of this type of games. First, we discuss relevant aspects around the design of puzzle-based games, followed by the proposed conceptual model, and its XML binding. Second, we illustrate how the model covers a variety of significant puzzle-based games including virtual and physical objects. Third, an exploratory user study is presented to analyze the extent to which teachers understand and are able to use the model to complete game-design tasks. Finally, we discuss the main conclusions to highlight the contributions of the paper.

Design of puzzle-based games for education

Several studies have identified different factors to consider when designing educational games (Fisch, 2005; Kirriemuir & McFarlane, 2004; Sandford & Williamson, 2005; Squire & Jenkins, 2003). The factors suggest that games should be based on constructivist learning theories, promote active learning and metacognition; clearly define the learning goals and tasks, be challenging and progressively increase the level of difficulty; provide immediate feedback and task-related supportive learning; and, offer help and hint structures to scaffold players along the game flow (Hjert-Bernardi et al., 2012). In this context, scaffolding refers to support mechanisms (e.g., hints or supportive learning material) that thoroughly guide students towards the successful completion of the learning activities proposed in the game (Wood et al., 1976).

Concerning game design models and frameworks, a broadly recognized useful approach is the 4-dimension framework proposed by de Freitas & Oliver (2006). This framework aims at helping teachers to evaluate the potential of using games- and simulation-based learning (Figure 1). We adopt this framework to analyse and extract the main characteristics to consider when supporting teachers in the design of puzzle-based games.

![Figure 1. Framework for evaluation educational games (de Freitas & Oliver, 2006)](image-url)
The first dimension of this framework focuses upon the particular context where play/learning takes place. In that sense, mobile technologies make possible to extend the learning environment far beyond classroom walls and school schedules (Liu, 2007). Also, mobile learning enhances traditional educational methodologies with greater portability and flexibility (Chang & Sheu, 2002). That means mobile games provide a way for learning activities that requires physical motion, problem solving, inquiry and collaboration (Spikol & Milrad, 2008; Avouris & Yiannoutsou, 2012). For this reason, it seems to make sense to consider mobile learning and games-based learning not as distinct experiences, but as experiences that could fruitfully be combined (Facer et al., 2004).

The second dimension focuses upon attributes of the particular learner or group. This may include the age and educational level, as well as specific components of how students learn including their learning background, styles and preferences.

The third dimension focuses upon the internal representational world of the game (i.e., the mode of presentation, the interactivity, the levels of immersion and fidelity used in the game). In this line, we differentiate between virtual and physical objects to represent both the different pieces and boards of a puzzle game. Interacting with virtual representations allow students to comprehend and manipulate abstract concepts that cannot be accessed in real world. Whilst there are learning situations in which physical representations have clear benefits: physical objects can be more easily understood than more symbolic ones, it is easier to demonstrate knowledge with physical actions, and solving problems with concrete objects can be easier than using symbolic representations (O’Malley & Stanton-Fraser, 2004; Li et al., 2008).

The fourth dimension promotes the practitioners’ reflection upon methods, theories, models and frameworks used to support learning practice. Several models have been used to describe the activities’ flow of educational games. One example is the Game Achievement Model (Amory & Seagram, 2003) that takes in consideration the learning objectives for the game and the storyline that encompasses these objectives and is defined as actions or activities. Also, Educational Modelling Languages have been studied for making digital games for learning purposes, and how aspects of gaming and learning have been combined with educational standards. In that sense, IMS Learning Design (Koper & Olivier, 2004), a specification for describing sequences of activities associated to user roles has been proposed as an option for designing educational games (Kelle et al., 2011).

We argue that the previously discussed design factors and dimensions can be used as the basis for a model that enables teachers to design their own puzzle-based games according to the educational needs. In particular, we focus on educational jigsaw puzzles. The objective of any jigsaw puzzle is the arrangement of a set of given pieces into a single, well-fitting structure that interrelates the pieces. Puzzle should be interesting because its result is not immediately intuitive; they can foster students’ problem solving, analytical and memory skills (Huang et al., 2007). Besides, there are several reasons that make puzzle-based games interesting approaches to involve teachers as designers. Puzzles usually have simple game rules (simplicity); their rules can be defined independently from content and therefore they can be applied in a wide range of subject matters (generality, independence); and their nature and duration is typically equivalent to other types of learning activities for the classroom or field trips (Michalewicz & Michalewicz, 2008; Falkner, 2010).

A Model to design technology-supported puzzle-based educational games

Conceptual model

Aligned with the factors and characteristics of games and puzzle-based design discussed in the previous section, Figure 2 adapts the (de Freitas & Oliver, 2006) framework to describe an overview of our model. Overall, a design of a puzzle-based game contains a learning flow consisting, at least, of a gaming objective, and a story structured by levels (which may relate to difficulty degrees or other issues). Different levels form the story of the game, and each level of the game presents either a single activity or a group of activities. Each activity includes a puzzle, and players have associated specific activities depending on their role within the game. Also, the context is a factor to consider in the design of the game since it can be done indoors or outdoors, allowing situated experiences. Each puzzle is represented solely by pieces, but also could be based on a board with slots where to place the different pieces. This means, puzzles could be solved on the one hand, by relating pieces among them; or, on the other hand, by relating...
pieces with the corresponding slots of a board. Both pieces and slots can represent virtual objects or computer-
recognized physical objects.

Figure 2. Overview of puzzle-based game design including virtual and physical objects

The entities depicted in Figure 2 can be conceptualized in a set of elements and their relationships. Figure 3
represents the different relationships that can be established between the different conceptual elements of our
proposed model for puzzle game design. The story of the game specifies which role has to perform which activity
and at what moment in the gaming process. For a player role, outcomes are stated as gaming objectives. These
outcomes are achieved by performing learning activities, within the levels of the game, with the help of hints that
scaffold the learning process if needed. In order to solve the learning activities, a set of (virtual or physical) pieces is
provided to the players to propose their own solutions. Pieces and their relations could provide mechanisms, such as
hints, questions or supportive learning material, to scaffold the gaming process as well. Besides, pieces could be part
of a board in which players should relate each piece with its corresponding slot. Otherwise, players should relate
pieces between them to propose the solutions. Depending on the puzzle game design (e.g., a virtual puzzle game, a
puzzle game for a museum, a puzzle game designed to be played in a city), either pieces or slots could have
associated virtual, tagged or geo-located positions.

Figure 3. Conceptual model of puzzle-based game design

Information model

The conceptual model previously explained can be computationally represented by an XML binding as shown in
Figure 4. More specifically, the computational representation for a puzzle game design includes:
- The title for a puzzle game design as well as the game objectives.
- The story of the game that contains the levels to be carried out sequentially.
- Each level defines who has to perform which activity or set of activities. That means, levels link each player-role to a specific activity or activity flow (i.e., group of activities).
- Activities appear as single activities or they could be grouped in an activities flow. The activity provides a description about the problem that has to be solved by a player performing a concrete role. An activity also contains a reference to a metadata that defines the puzzle intended to solve the proposed problem. Each activity also can provide different scaffolding mechanisms to assist the player during the game as well.
- Puzzles are used when a player performs a concrete activity, but they do not form part of the activity description itself. Thus, each puzzle is linked to each activity of the game.
- Scaffolding is used by player roles when asking for some type of help while performing an activity. Scaffolding does not form part of the activity description itself neither.

![Diagram of puzzle game design](image)

*Figure 4. The main elements of a puzzle game design*

The different puzzles associated to each activity are also defined by its corresponding XML document in order to be computationally represented as well (Figure 5).

![Diagram of single puzzle](image)

*Figure 5. The main elements of a single puzzle associated to an activity*
• We differentiate two types of relations within a puzzle. First, those relations in which the puzzle is not contained into a board. They define only the relations that can be made between the different pieces that form the puzzle (rels-between-pieces). Second, the relations that has to be defined for relating the different puzzle pieces with the slots of the board in which the puzzle is contained (rels-pieces-slot).
• A piece includes its name, the type of piece depending on its representation (i.e., virtual or physical), and scaffolding mechanisms (i.e., hint, question prompt, or supportive learning material).
• A slot is defined by its type depending on its representation (i.e., virtual or physical), information to describe the slot, location (defined by coordinates \{x, y, z\} for virtual representations; or a geo-located coordinate, for in situ interactions), and scaffolding mechanisms (i.e., hint, question prompt, or supportive learning material).
• Each relation, for puzzles that only consider the relations between pieces (rel-between-pieces), is defined by the reference of the different pieces forming the specific relation, a feedback provided to the player depending on whether the relation is correct or not, and a score consisting on positive (correct relationships) or negative (incorrect relationships) points.
• Each relation, for puzzles that consider the pieces and the slots of a board (rel-pieces-slot), is defined by a reference of the piece and the slot forming the relation itself, a feedback provided to the player in order to inform whether the relation is correct or not, and a score consisting in positive points for correct relations, and negative points for incorrect ones.

Implementation guidelines and authoring tool prototype

The implementation guidelines introduced in this section should be understood as a possible approach suggested for the actual enactment of puzzle-based games. It aims to help game developers with the system design process.

A first authoring tool (Figure 6) has been created to generate the XMLs compliant with the information model for designing puzzle games. The tool is implemented as a wizard. First, the teacher specifies the title of the puzzle and the learning goals expected to be achieved by playing the game. Next, the teacher introduces the information associated to the levels for the game. For each level, at least, an activity with its associated puzzle has to be defined. In order to create the puzzle, the teacher selects between creating a puzzle of “relation between pieces” or “relation between pieces and slots”. Depending on this selection, the teacher specifies the demanding information, for instance, the content for the different pieces, slots, hints and the relationships between them. At the end, the different XML documents are automatically generated by the authoring tool.

Figure 6. Some screenshots of the authoring tool

The XML format allows storing the data independently from specific platforms. Depending on the targeted platform, it is needed, an engine able of interpreting (parsing and instantiating) these packages of information compliant with the puzzle-based game model, and a player to execute the live interpretation of the game. Platform-specific engines (e.g., for handheld devices, tangible interfaces…) should read the information contained in the different XML documents and check the different constrains (e.g., to show the activities associated to a player role, to show the puzzle associated to a concrete activity, etc.). Therefore, platform-specific engines need to contain the business logic of the conceptual model for technology-supported puzzle games. The responsible of visualizing and dealing with the
content provided by the engine is the platform-specific player (Figure 7). External interfacing can be included in the architecture by providing gateways to other systems (as similarly achieved in (Santos et al., 09)).

Expressiveness of the model

To show the expressiveness of the model, this section discusses different examples covering a wide range of puzzle-based games. These examples include diverse virtual and physical objects’ (as pieces or slots) combinations and are meaningful to different educational contexts. In concrete, several examples by the authors of the model, as well as third-party examples proposed by other researchers are described. Reporting not only authors’ own examples, but also external ones, adds a stronger validation of the model expressiveness and shows how the model can be applied to a diverse variety of educational contexts.

Examples of puzzle game including virtual pieces and virtual slots

Several puzzle-based games, following a simplified version of the proposed conceptual model, have been developed. In concrete, the games aim at learning computer architecture, programming fundamentals, computing networks (Melero et al., 2012) and chemistry (Hjert-Bernardi et al., 2012). Examples of implemented puzzle pieces includes: logic gates (computer architecture), chunks of pseudo-codes (programming fundamentals), and elements of the periodic table (chemistry). All these games were designed and evaluated in laboratory sessions of different secondary and high schools.

The proposed conceptual model was also used by a secondary school teacher to conceptualize a game consisting in answering questions about different pictures of a contemporary art museum (Melero et al., 2013). This example shows a learning situation in which the implemented game was enacted in a field trip. The game content reflects
concepts of different contemporary art pictures exhibited in the museum. Students, while visiting the museum had to
answer different questions in situ using their smartphones. In order to implement the game compliant with the model:
(a) each picture was virtually represented as a slot, the content of which was the question itself; (b) the possible
options to answer a question were represented as virtual pieces; and (c) in order to indicate whether an option (i.e.,
piece) of a given question is correct or not, we used the “rel-piece-slot” element (see Figure 8).

Figure 8. Screenshots of the game and chunks of its corresponding XML representation compliant with the model

Example of a puzzle game including virtual pieces and physical slots

Lee and Chen (2008) propose a mathematical educational game for young children in an Augmented Reality
environment to practice with the concept of addition and subtraction. This could be an example of a puzzle game in
which virtual pieces and physical slots work together. The player has three dice (two numbered dice and one operator
dice) that have to roll and calculate the outcome. Then, the player moves a piece on the board according to the
outcome. The player who arrives at the finishing point first wins the game. The board is made of markers and pieces
are augmented as 3D models. The players can select different boards and the 3D models player’s pieces would
change as on the selected board. The representation of the game with the model would specify each board and its
corresponding pieces as a different puzzle in the game. Each of the board’s markers would correspond to a physical
slot, and the augmented pieces would be mapped as virtual pieces of the model. Thus, different “rel-piece-slot”
elements would be defined to specify the possible relations that can be established between each piece and slot.

Example of a puzzle game including physical pieces and virtual slots

Two functional prototypes of puzzle-based games using physical pieces and virtual slots have been implemented
(Ponz-Adán, 2012; Cardona-Serra, 2012) compliant with the conceptual model. These games reflect contents of
computer architecture and programming fundamentals. Puzzle pieces were implemented using the ReacTIVision
technology (Kaltenbrunner & Bencina, 2007), an open-source toolkit for developing tangible multi-touch surfaces.
Particularly, the games consisted in matching the physical pieces, representing different computer science concepts,
with their corresponding virtual slots (see Figure 9). Besides, both games were evaluated with secondary education
students in different one-hour session of a laboratory class.
Example of a puzzle game including physical pieces and physical slots

“The Smart Jigsaw Puzzle Assistant” (Bohn, 2004) is a computer application that operates with a physical jigsaw puzzle game that uses miniature RFID tags and a palm-sized RFID scanner to technologically support the interactions with both physical pieces and slots. The application is executed on a computer and monitors the current status of the physical jigsaw puzzle. Whenever the player chooses a new piece of the physical jigsaw game to be added to the previously combined pieces on the table, he or she scans it with a handy RFID reader connected to the computer. The application then automatically recognizes the added piece and updates the status of the jigsaw game on the computer screen. Both piece and slot elements would be specified as physical (setting the attribute type to “tag”) to comply with the conceptual model. The application would recognize the added physical piece and check, using the “rel-piece-slot” element, whether the piece has been placed in a correct position.

Exploratory study with teachers

In order to test whether teachers understand the elements of the model, we carried out an exploratory study in which they have to solve a proposed design task. Before discussing the results, we describe the proposed design tasks and the methodology applied in the study.

Methodology and design task

Teachers were recruited to participate upon an open invitation. Invitations were sent by email to two secondary education schools and one higher education (university) department. A total of 11 teachers replied positively to the invitation and agreed to participate in the user study. As done in (Derntl et al., 11), this study was organized as paper-based sessions. Participants used paper snippets representing the elements of the conceptual model to solve the design task. In particular, the sessions were structured as:

- Pre-test for gathering contextual information and the teachers’ background using technology or educational games in their learning approach (10 minutes).
- Introduction about game-based learning and the key elements of the model needed to build a puzzle game. Some examples using the elements of the model were shown as well (15 minutes).
- Paper-snippets introduction. Every participant received an envelope with paper snippets (see Figure 10), each representing one instance of an element of the conceptual model. Each paper snippet was divided into boxes that represent information about the element in conformance with the model. Besides, each element had its own unique colour in order to guide participants in placing connections between elements (5 minutes).
- Puzzle-based game design task. After the introduction, every participant had the task of representing one scenario (two narratives of puzzle-based game scenarios were defined and randomly offered to participants) (50 minutes). The resulting designs created with the snippets were collected, once the participants finished the design task.
- Post-task survey. Participants were asked to fill out a post-task survey, which aimed to collect additional information about what caused them more problems during the task (10 minutes).
As a task protocol, participants were not offered any help or guidance during the task other than (a) provision of a cheat sheet with a tabular overview of the elements of the conceptual model and (b) personal answers to questions for clarification.

Prototype solutions for each of the two proposed scenarios were created by decomposing the overall task into solution items. Each solution item consisted of an action that needed to be performed with an individual element of the model, or a small group of elements, in order to obtain a correct solution. Each solution proposed by the participants was individually analysed and matched with the prototype solution. Table 1 lists the scenarios and includes sample snippets. The complete collection of snippets composing the specification of the games described in the scenarios is available as an on-line appendix to this article (http://www.dtic.upf.edu/~jmelero/et&s/index.html).

<table>
<thead>
<tr>
<th>Table 1. Scenarios proposed in the design task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario</td>
</tr>
<tr>
<td>1. Imagine a non-native English student wants to improve his/her English vocabulary related to shopping. The first level aims at learning English words for different clothes [s1-1]. Each activity, in this level, consists in forming words of different type of clothes that are shown in pictures. The player, using different letters, has to construct the specific name of clothes shown in the picture [s1-2]. The next level consists in an activity for learning specific shop departments (dressing room, counter, etc.) [s1-3]. To this end, a map of the shop containing different empty spaces is presented to the player. Then, the player has a set of given words to place in a specific area of the map [s1-4]. The purpose is to match each word with its correct empty space in the map. During the game different hints can be provided to the player indicating for instance, the first letter of a word, the last letter, etc.</td>
</tr>
<tr>
<td>2. Using a Smartphone, a student has to find some contemporary pictures in a museum [s2-1]. The Smartphone shows, on one hand, the museum’s map and, on the other hand, a set of pieces that represent specific parts of concrete pictures [s2-2]. The museum’s map also contains different highlighted locations that correspond to specific pictures in the museum. Students have to match each piece with its corresponding picture/location. Each picture in the museum has associated a physical card (e.g., NFC tags, QR-Code, etc.) [s2-3]. The player has to read the physical card to indicate that he/she has reached a specific location. Then, once the player arrives to a concrete location and read the physical card, he/she can select which of the virtual piece correspond to the concrete the picture [s2-4]. Each virtual piece may contain some hints indicating the area of the museum where the student can find the related picture.</td>
</tr>
</tbody>
</table>

To analyze to what extent teachers were able to complete the specification of the games and identify the missed elements of the resulting designs, we use a chart taking into account the aforementioned decomposition of the overall tasks into solution items. Thus, Table 2 summarizes the different data sources considered in the evaluation.

<table>
<thead>
<tr>
<th>Table 2. Data sources for the evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data source</td>
</tr>
<tr>
<td>Pre-test</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Participants Solutions</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Results: Teachers’ profiles, opinions and previous experiences with educational games

Participants’ average teaching experience is 5 years, in different subject areas. Half of them have a technical background. 75% of teachers teach in secondary education; the rest teaches in higher education. All the participants use technology in their classes, with YouTube (63%), Moodle (45%) and Facebook (36%) the most used tools. All participants rated with 4 or more (in a scale from 1 to 6) the need of using technology in education to enrich the students’ learning experiences.

Using a scale from 1 to 6, 70% of the teachers rated with 4 or more the interest of using games in education. Also, 80% of the teachers rated with 4 or more the usefulness of games in education. In this regard, some comments were: “Educational games can be a complement to the students’ learning in other to motivate them” [Pretest-2], “Learning by means of games stimulates and motivates students, but designing appropriate games can be difficult for a teacher” [Pretest-5]. These data indicate that teachers perceive educational games as important approaches to support their lessons because technology and gaming motivate students. However, the teachers also recognize that its design is time consuming and they do not have enough resources to easily create their own games. It seems that for this reason teachers tend to adopt mainstream games in their lessons.

Half of participants had adopted or made some variations from existing games. These games were “collaborative games, avoiding competitiveness, intended to work on retention capacity and strategy” [Pretest-2], “games such as memories or scatergories” [Pretest-4], and “games to form sentences using cards, to guess famous people” [Pretest-11]. The other half of participants do not use educational games because “the devoting time; I prioritize to explain my lessons” [Pretest-2], “There are not games covering the topics of my subject” [Pretest-8] and “I don’t know the technology or the means to include games in my teaching” [Pretest-6].

Results: On the outputs from the design tasks

The resulting designs created by the teachers were compared to the solutions prototyped for each scenario. The different teachers’ profiles did not seem to influence the results of the experiment since their outcomes and opinions were overall quite aligned. Then, the data was aggregately analyzed, and it was just indicated the teachers’ profiles when appropriate. The comparisons are captured in Tables 3 (scenario 1) and 4 (scenario 2). Each table shows whether the participants used correctly the elements of the conceptual model for designing the puzzle-based games.

The percentage of conformity in the total sample for the results of the scenario 1 is 76%, whereby only 1 out of 4 solutions was rated at less than 50%. Despite final solutions, such as [Sol-Sc1-5], were simpler than expected; all participants were able to use and correctly relate the different elements of the conceptual model. Activities’ flow was the more problematic element. In order to include several puzzles in a level, it is necessary to include the element activities flow that contains as many activities as puzzles are designed for the level (one activity per puzzle). However, [Sol-Sc1-2] specified the references of the different activities in the attribute “activity reference” (instead of using the activities flow element), [Sol-Sc1-4] specified for a single activity a group of puzzles instead of creating a puzzle per activity, and [Sol-Sc1-9] used the activities’ flow snipped to directly relate the different puzzles for a level, instead of specifying the single activities included within the activities’ flow and related each single activity to a puzzle. The percentage of conformity was also affected by the time devoted to design the proposed scenario; participants [Sol-Sc1-2] and [Sol-Sc1-4] (from secondary education schools) did not have enough time to finish their solutions. Two isolated mistakes referred to not using the puzzle-based game design snipped to indicate the title, objectives and levels of the games [Sol-Sc1-5]; and duplicate the same player role to different levels of the game [Sol-Sc1-9].
The results for scenario 2 show that there is a generally high conformity of the teachers’ solutions with the prototype solution. The percentage of conformity in the total sample is 85%. Despite final solutions were simpler than expected, all participants were able to properly use the different elements of the conceptual model. There was only one exception, [Sol-Sc2-6], in which the proposed solution was not satisfactory. The participant associated tag pieces for the pictures of the museum, and virtual pieces for showing chunks of some pictures of the museum. The expected solution was using a board with slots (i.e., QR-codes for pictures of the museum) and pieces (i.e., virtual pieces showing chunks of some pictures of the museum). For this reason, the solution was not considered correct since the participant did not use the expected elements. Besides, the attributes ‘virtual’ and ‘tag’, for the pieces (i.e., [Sol-Sc2-1], [Sol-Sc2-7], [Sol-Sc2-11]) and slots (i.e., [Sol-Sc2-3], [Sol-Sc2-6], [Sol-Sc2-7]), were not correctly used in some cases. A concrete error was the [Sol-Sc2-11] that specified both “relation between pieces” and “relation pieces – slots”, when only “relation pieces – slots” was correct. However, this specific error seemed to be more related with the graphical design of the snapshot than with the formulation of the model elements.

Table 4. Participants’ solutions’ conformity with the prototype solution of Scenario 2 (Notation according Table 2)

<table>
<thead>
<tr>
<th>Scenario 2</th>
<th>Elements</th>
<th>Sol-Sc2-1</th>
<th>Sol-Sc2-3</th>
<th>Sol-Sc2-6</th>
<th>Sol-Sc2-7</th>
<th>Sol-Sc2-10</th>
<th>Sol-Sc2-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>[s2-1]</td>
<td>Puzzle game design</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td></td>
<td>Level</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td></td>
<td>Player role</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>[s2-2]</td>
<td>Activity</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td></td>
<td>Puzzle</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td></td>
<td>Piece</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td></td>
<td>Slot</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td></td>
<td>Rel. piece-slot</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>[s2-3]</td>
<td>Slot</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>[s2-4]</td>
<td>Rel. piece-slot</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Use of hints?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Conformity (%)</td>
<td>89%</td>
<td>89%</td>
<td>78%</td>
<td>78%</td>
<td>100%</td>
<td>78%</td>
<td></td>
</tr>
</tbody>
</table>

Results: On the design process

Participants encountered some difficulties in the design process in regards to: (a) the creation of groups of activities for designing different puzzles for a concrete level of a game, (b) referencing the different snippets, and (c) the design of the task protocol.

8 out of the 11 participants thought that they correctly understood the elements for the design of the game. However, participants have troubles with understanding some of the elements. The most problematic element was the activity flow [Postest-4-5-7-9-10-11]. In fact, as showed in the previous section, participants [Sol-Sc1-2], [Sol-Sc1-4] and...
[Sol-Sc1-9] failed at creating activity flows. For instance, some participants indicated “I considered the activity about clothes as a one activity with various puzzles instead of an activity flow” [Sol-Sc1-5], and “I don’t distinguish between activity and activity flow” [Observer 1]. Other problematic elements were the pieces and slots [Postest-2-6-8]. In particular, some participants did not correctly understand the attributes ‘virtual’ and ‘tag’ for these two elements [Sol-Sc2-1-3-7]. These misunderstandings were mostly derived from the terminology: “I don’t understand the meaning of some concepts because they are unfamiliar” [Postest-1], “what is the difference between puzzle and board?” [Observer 1].

10 out of the 11 participants understood how to relate the different elements. Participants appreciated the examples showed in the introduction [Postest-4-8], and the use of different colours helps the participants to correctly design the games [Postest-1-5]: “The colors of the snippets were useful to design the game” [Observer 1]. However, the participants [Sol-Sc1-2], [Sol-Sc1-4] and [Sol-Sc1-6] had troubles when indicating the ‘id’ for each snippet. For instance, participant 9 pointed out: “what does ‘id’ mean in the different targets?” [Observer 1].

Task protocol induced some issues to the participants during the design process. Participants felt lost because we did not offered any guidance during the task: “at the beginning I did not understand anything” [Postest-3], “I have some problems to start designing the game, I wasn’t sure what to put in each snipped” [Postest-2], “I understand the scenario of the game but now, which targets do I take?” [Observer 1]. However, after some clarifications, all of the participants agreed that they were able to propose a solution: “once you understand the mechanics, it is easier” [Postest-4], “the process is through repetitions, so one can easily become familiar with the process” [Postest-5], “It presents a clear structure” [Postest-9].

Finally, all participants agreed that they would like to design their own games and test them with their students. Besides, when asking the participants how much they would devote to designing their own puzzle-based games according to our conceptual model, the average was half a day [Postest-all]. Some comments were: “If it is effective, I would devoted as much time as it was necessary” [Postest-4], “I don’t care, I would expend as much time as I need” [Postest-11].

Conclusions

The literature and the feedback provided by the teachers in the presented exploratory study agree on the importance of the problem addressed in this paper. Despite the potentially effective learning benefits of educational games, teachers do not apply widely these games in their teaching. Only occasionally they use mainstream games, such as mind games or puzzle games. The adoption is not extensive, because game designs are not always aligned with the requirements of the specific educational situations faced by the teachers and because the existing authoring tools are still too complex for teachers (Tornero et al., 2010). Creative teachers do devote time to design paper-based games for their classrooms but they do not typically have the advanced technological skills that would enable them the design of computer games.

As a first step to work towards solutions that tackle this problem, in this paper we focus on puzzle-based games. The simplicity and generality of puzzle-based games makes reasonable for teachers to act as their designers. We contribute with a model (conceptual model and XML information binding) that enables expressing diverse types of puzzle-based games including virtual and physical objects. As shown in the paper, the games that can be computationally represented with the model share the basic rules of games (e.g., scoring and hint mechanisms) and puzzles (e.g., interrelating pieces, considering slots) but allow different types and nature of content (e.g., contextualized in locations, tangible, completely virtual). In this line, the framework proposed by de Freitas & Oliver (2006), has been useful to define the general elements needed to conceptualize puzzle games independent from context and interactions, while considering the player’s role and activity flow. Besides, focusing on activity flow design, both Game Achievement Model (Amory & Seagram, 2003) and IMS Learning Design (Koper & Olivier, 2004) have been worthy examples to structure and interrelated the different elements of our proposed model.

To evaluate to what extent teachers are able to use and understand the different elements of the conceptual model, we carried out an exploratory user study with teachers from secondary and higher education. In particular, two different scenarios were proposed as game design tasks. The overall percentage of conformity for the first scenario (when compared to a prototype solution) was of 76%, whilst in the second scenario was of 85%. The degree of achieved
conformity is valued as satisfactory, given that it was the first time the teachers were using the model (still low familiarity) and considering the limitations associated to the experimental design (snippets vs. proper authoring tool). The aspects of the model that hindered a complete conformity were related to difficulties in understanding the differences between activities’ flows and single activities, and the virtual and tag attributes for pieces and slots. We will consider these details to refine how the model will be implemented in an authoring tool.

References


Strategies for Smooth and Effective Cross-Cultural Online Collaborative Learning

Junfeng Yang1,2, Kinshuk2,4, Huiju Yu1, Sue-Jen Chen3 and Ronghuai Huang4*

1Hangzhou Normal University, China // 2Athabasca University, Canada // 3University of North Carolina Wilmington, US // 4Beijing Normal University, China // yangjunfengphd@gmail.com // kinshuk@athabascau.ca // yhj@hzn.edu.cn // chensj@uncw.edu // huangrh@bnu.edu.cn

*Corresponding author

(Submitted August 28, 2013; Revised April 8, 2014; Accepted April 21, 2014)

ABSTRACT

As the communication between different cultures is becoming more and more frequent, the competence of cross-cultural awareness and collaboration is emerging as a key ability in the 21st century. Face to face communication is the most efficient way to cultivate the competence of cross-cultural awareness and collaboration. However, there are very few opportunities currently available for university students to have such face to face communication. Therefore, cross-cultural online collaborative learning utilizing web 2.0 technologies is proposed in this paper as a way to cultivate students’ cross-cultural competence. The purpose of the study is therefore to elicit strategies for smooth and effective cross-cultural online collaborative learning through a pilot study between the West and the East. Students of a Chinese University and an American University took part in the study. A mixed method research approach using questionnaire, interview and content analysis was used. The findings of the study revealed that students from both sides were interested in each other’s culture, their attitudes to cross-cultural online collaborative learning were positive, and culture had an influence on learning methods. Social interaction played an important role, and students preferred to have more prior knowledge of each other’s cultures and backgrounds. They were also inclined towards more in-depth individual conversations. As a result of this study, several strategies are proposed to facilitate effective implementation of cross-cultural collaborative learning in typical higher education settings.

Keywords

Cross-cultural, Collaborative learning, Online learning, Cultural awareness, Collaborative skills

Introduction

The global village nature of the world has become more salient in recent years (Landis, Bennett, & Bennett, 2003), and people from different cultures are communicating more and more with each other. The collaboration between different cultures has also been increasing dramatically. It has therefore become very important that current generation of students has acquired multi-cultural awareness and cross-cultural collaborative skills before it embarks on to the work environment.

Trilling & Fadel (2009) proposed 21st century skills in the book “21st century skills: learning for life in our times”, with cross-cultural communication competence and collaboration competence being included in those skills. They stressed these skills to be essential for students for better involvement in future society and for having good performance in the future career. For cultivating students’ cross-cultural ability, the first choice is to be immersed in another culture for a long time and have face-to-face communication with local people (Wang, 2011). Although the international activities of universities have dramatically expanded in volume, scope, and complexity over the past two decades, especially the study abroad programs allowing students to learn about other cultures for cultivating their cross-cultural awareness and cross-cultural collaborative skills (Altbach & Knight, 2007), very few students in fact are able to avail such opportunities, and existing opportunities are not sufficient by any means when considering the current notions that intercultural education should be implemented into all levels and forms of education in the future (Harms, Niederhauser, Davis, Roblyer, & Gilbert, 2006; Wang, 2011).

In this context, web 2.0 technologies have emerged as a potentially vast technological solution for global knowledge sharing, construction and distribution across groups, countries, and cultures (Friedman, 2005). When collaboration involves cross-cultural experience, it not only expands students’ view of multiculturalism, but also enhances their self-concept and cross-cultural communication and collaboration competence (Cifuentes & Murphy, 2000). Cross-cultural online collaborative learning utilizing web 2.0 technologies is therefore proposed in this paper as a way to
enable communication and collaboration among students from different cultures. Cross-cultural dimension in this paper refers to the context where students from different cultures take part in online collaborative learning, without any emphasis on essential cultural attributes. Beginning with email, educators have applied technology to increase cultural awareness over the past three decades which has provided an online method for improving students' learning experience in a cross-cultural context (Liaw & Johnson, 2001). Learning management systems, blogs, social media platforms such as Facebook and Twitter, and even synchronous communication methods have been used by researchers to facilitate cross-cultural online learning (Law & Nguyen-Ngoc, 2010; Wang, 2012; Wang & Chen, 2012; Wu, Marek, & Chen, 2013). Researchers have also explored pedagogy, instructional design, task and assessment design for effective cross-cultural online learning (McLoughlin, 2001; Chen, Hsu, & Caropreso, 2006), and collaboration and community building are generally regarded as designs for cross-cultural online learning, which explains the rationale behind the current study on cross-cultural online collaborative learning.

Researchers have argued that learning through cross-cultural online collaboration is not an easy task (Kim & Bonk, 2002; Wang, 2011; Canto, Jauregi, & van den Bergh, 2013), and therefore it is important to identify and explore critical factors for smooth and effective cross-cultural online collaborative learning. “Smooth and effective” are used here to indicate the requirements for the implementing cross-cultural online collaborative learning. “Smooth” indicates that the whole learning process should be conducted by teacher with no difficulty, and “effective” means students from different cultures actually engage in the whole learning process.

Students’ attitude has frequently been reported to be the most critical factor for success within computer-assisted learning environments (Ushida, 2005). Previous studies have also suggested that students’ attitude towards cross-cultural learning plays an important role in the success of cross-cultural online collaborative learning (Lim, 2007; Chen, Hsu, & Caropreso, 2009; Chen, Caropreso, Hsu, & Yang, 2012). Law & Nguyen-Ngoc (2010) used social software to analyze cross-cultural online collaborative learning and pointed out that motivation of the learners as well as the facilitators is a critical success factor for the success of the online collaborative learning environments. Attitude is found in the literature as the foundation for the success of such new initiatives, and knowing the attitude towards cross-cultural online collaborative learning would be of great help in conducting successful cross-cultural online collaborative learning. Attitude is closely related to motivation, but the affective issues or motivation aspects in collaborative learning have largely been neglected in previous research (Dillenbourg, Järvelä, & Fischer, 2009).

The second critical factor is the strategies for implementing cross-cultural online collaborative learning. Online facilitation in a cross-cultural collaborative learning context is increasingly prevalent, but limited international cross-cultural online collaborative learning project implementations have been reported in the literature (Nguyen-ngoc & Law, 2009; Liaw & Bunn-Le Master, 2010; Wang, 2011; Leppisaari & Lee, 2012), partly because it is not an easy task to develop cross-cultural online collaboration (O’Dowd & Ritter, 2006; Hauck, 2007). Instructors and students need to overcome the problem associated with the synchronous communication across time zones and between institutions (Murphy, 2005), and they also need to be aware of differences between groups in terms of the levels of proficiency in using technology. Language may also be a barrier for cross-cultural communication (Wang & Reeves, 2007). It is also time consuming, since the instructors need to communicate and negotiate with teachers from other countries, and there may be difficulties in grading due to the differences in the academic quality standards among different cultures (Braskamp, 2008). Although some general tips for online cross-cultural collaboration have been proposed in the literature, such as being mindful, being comfortable with silence, encouragement for differing viewpoints, avoidance of debates, observations, and normalization of diversity, there is a lack of strategies for implementing cross-cultural online collaborative learning (Wang, 2011).

A case study has therefore been conducted to elicit strategies for smooth and effective cross-cultural online collaborative learning. A detailed plan for the collaborative learning process is proposed, along with a new data analysis framework. The data collected from the questionnaire, the posts in the learning management system and the focus group interviews are then analyzed using the proposed framework. The next section of the paper discusses the research framework, study process, the instruments, and data analysis methods. This is followed by a comprehensive analysis of strategies for smooth and effective cross-cultural online collaborative learning based on the findings of the study. The paper finally concludes with a discussion on the strategies to implement smooth and effective cross-cultural online collaborative learning and an outlook towards future research directions.
Methods

Research framework and theoretical background

Three theoretical models have guided this research: collaborative learning (Johnson & Johnson, 1999), model of “community of inquiry” (Garrison, Anderson, & Archer, 2000), and evaluation methods for cross-cultural collaboration (Law & Nguyen-Ngoc, 2010). Application of these models for cross-cultural online collaborative learning is discussed briefly in the following sections.

Collaborative learning

Collaborative learning is a situation in which two or more people learn or attempt to learn something together (Bruffee, 1993). Johnson & Johnson (1999) proposed positive interdependence, individual accountability, face to face interaction, social skills and group processing as the five basic elements of collaborative learning, and the grouping strategies were essential to meet the needs of the five basic elements. Dillenbourg, Baker, Blaye, & O’Malley (1996) suggested that heterogeneous groups could be beneficial as a condition to trigger conflicts and require negotiation and social grounding, in addition, Setlock, Fussell, & Neuwirth (2004) suggested that experimental groups with homogenous and heterogeneous cultural backgrounds had different perceptions of the study task. With respect to this importance, the rules for grouping such as heterogeneity in a group and homogeneity between groups were decided in terms of students’ features such as gender, age, learning style, knowledge base, and cultural background.

Model of “community of inquiry”

The model of “Community of Inquiry” assumes that learning occurs within the community through the interaction of three core elements: cognitive (construction of meaning through sustained communication); social (ability of participants to project their personal characteristics into the community); and, teaching presence (the design and facilitation of educational experience) (Garrison, Anderson, & Archer, 2000). In cross-cultural context, social presence could be facilitated by designed process of social lounge and cultural orientation (Chang, & Lim, 2002). Strong teaching presence and cognitive presence are important to allow collaborative knowledge construction to occur in communicatively lean environments (Osman & Herring, 2007). As teachers from both sides gather together with students in the collaborative process, teaching presence can be guaranteed and cognitive presence is facilitated in the process of technology integration.

Evaluation methods for cross-cultural collaboration

Law & Nguyen-Ngoc (2010) proposed three dimensions, namely social, cognitive and affective, to analyze the quality of cross-cultural interaction. Social means how well individual group members are collaborating which mainly refers to the participation in collaboration. Cognitive means how well the group performs the task. Affective means how satisfied or frustrated the group members feel during the task. For the cross-cultural online collaboration, cultural identity is also important (Parrish & Linder-vanberschot, 2010). Previous research has suggested that cultural differences could have a negative effect on students’ participation in online courses (Liu, Liu, Lee, & Magjuka, 2010), and language barriers for non-native speakers tend to detract from equal participation (Gunawardena, et al., 2001). Therefore, participation should also be considered as a key element for evaluating cross-cultural collaborative learning. With respect to these dimensions, cultural identity, affect, and participation are identified in this research as the three elements for evaluating cross-cultural online collaborative learning. For each aspect, multiple methods are used to collect data, as shown in table 1.

Table 1. Evaluation methods for cross-cultural online collaborative learning

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Index</th>
<th>Data collection mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultural Identity</td>
<td>Mutual cultural consideration</td>
<td>Posts about culture understanding</td>
</tr>
<tr>
<td></td>
<td>Understanding of each other’s culture</td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td>Culture’s influence on behavior</td>
<td>Interview with students</td>
</tr>
</tbody>
</table>
Based on the three guiding theoretical models, a research framework is designed in this study, as shown in figure 1.

**Participants, group formation and collaborative learning process**

A case study was conducted between a Chinese University and an American University. A judgment sample technique was used to select students from both sides. 28 students (10 males and 18 females) with major in educational technology from the Chinese university and 34 students (6 males and 28 females, none of them from Chinese origin) with major in educational technology from the American university took part in the research. All students at the Chinese university were designated as juniors with age range = 18–23 and average age distribution (mean) = 20.4. Most of the American students were juniors (82.4%), with some sophomores (8.8%) and seniors (8.8%), with age range = 18–44 and average age distribution (mean) = 28.6.

At first, students from China and United States were divided into different groups. In collaborative learning, group formation plays a critical role that affects the acceptance of group activities and the success of the learning process (Isotani, Inaba, Ikeda, & Mizoguchi, 2009). Membership or group composition is a critical component as it influences the motivation of individual learners, whose needs (cognition, social and emotion) determine how they behave in a group (Trautwein, Lüdtke, Marsh W., Köller, & Baumert, 2006). Previous studies have found that three to five members in a group are more effective for cooperative or collaborative learning, and Chun-min Wang (2011) suggested smaller groupings would be preferred by students. However, this research included two classes directed by teachers from China and US, where small groups were formed in each class, and therefore, the combined Chinese
and US students groups were large groups according to the rules for grouping. The participants were assigned into 5 groups, each containing 5 to 6 Chinese students and 6 to 7 American students. When dividing students into different groups, factors of heterogeneity (such as the level of English competence, learning style, learning performance, etc.) within each group and homogeneity (such as total number of students, formation of group, average learning performance, etc.) between the groups were taken into consideration. Each group was assigned a leader within the group who could take charge of the group in order to make sure that the mixed group work was conducted in close cooperation from both sides.

Secondly, the collaborative processes which lasted one month were divided into three successive stages: social lounge lasted for one week, cultural orientation lasted for one week after social lounge, and technology integration lasted for two weeks after cultural orientation. These collaboration processes were started with participants knowing each other, and in a cross-cultural context, knowing each other means knowing each other’s basic information and knowing each other’s cultural background. Social lounge enabled students from both sides to communicate with each other in order to know basic information about each other, such as hobbies, families, and so on. Cultural orientation enabled students to communicate in order to know about each other’s culture. Technology integration forum enabled students from both sides to discuss some major topics with each other, such as what technologies were used in the class, how teacher integrated these technologies into curriculum, etc.

Effectiveness of the group work in such collaborative processes depends on the proactive monitoring and intervention by the instructors during collaborative activities through mechanisms for assistance, feedback, and evaluation (Capdeferro & Romero, 2012). Therefore in each stage, requirements and statements were presented to the students so that they understand their activities clearly. Due to the time zone differences, all these three stages in the online collaborative learning environment were preceded by asynchronous communication between the two sides. Students from both sides were asked to reply to each other’s posts at least twice in the cross-cultural online learning environment. Chinese students gathered together in a classroom from 8:00-10:00 in the morning (China time), which is 19:00-21:00 in the evening in US (Eastern Time Zone). The online learning environment used in the study was Blackboard platform in English, which was selected after analyzing a number of Web 2.0 learning management systems and personal learning platforms, where many of the systems could not be used between China and the United States due to various firewall restrictions.

Instruments and materials

Both the qualitative and the quantitative research methods were used in this research. A questionnaire was developed based on the evaluation methods for cross-cultural collaborative learning. The quantitative data was collected from various close-ended questions of the questionnaire. The qualitative data was collected from the posts in the forum, through open-ended questions of the questionnaire, and from the focus group interviews.

To analyze the participation, the questionnaire contained open-ended questions related to the process of completing group collaboration activities and mutual inspirations to reach consensus. To analyze affect, the questionnaire contained both close-ended questions and open-ended questions related to the formal experience and attitude towards the cross-cultural online collaborative learning and satisfaction with the collaborative learning process. To analyze cultural identity, the questionnaire contained open-ended questions related to cultural differences.

The questionnaire was validated using "content validity ratio" (CVR). Five experts of cross-cultural research and educational technology were invited to give scores on the validity of the questionnaire. Based on the feedback from the experts, the questionnaire was revised. Final questionnaire included 12 close-ended questions with 5-point Likert scale (1= strongly disagree, 5= strongly agree) and 3 open-ended questions, as shown in table 2.

<table>
<thead>
<tr>
<th>Table 2. Dimension and items of the questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question Type</strong></td>
</tr>
<tr>
<td>Close-ended questions</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
3. Social Lounge helped acquaint me with class peers.
5. I felt I was connected with my group to talk and learn together online.
6. I felt I was connected with the class to talk and learn together online.
7. The discussion with foreign peers contributed to my learning of subject content.
8. I had much experience working with students from other country before this cross-cultural online collaborative learning experience.
9. This cross-cultural online collaborative learning contributed to my life experience.
10. This cross-cultural online collaborative learning helped expand my global perspectives.
11. I would like to have similar cross-cultural online discussions in other courses.
12. Overall, my cross-cultural online collaborative learning experience was positive.

At the end of the pilot study, students were requested to complete the questionnaire. 28 valid questionnaires from Chinese students and 34 valid questionnaires from American students were collected. Cronbach’s alpha coefficient was calculated to test the consistency of the items within the valid questionnaires. The internal consistency values (Cronbach’s alpha coefficient) in each dimension ranged from 0.85 to 0.95. Since Cronbach’s alpha coefficient values higher than 0.70 are considered as good, and when the value is close to 1.00, it is considered very good (Fraenkel & Norman, 2003), the internal consistency values calculated in this study can be interpreted as very good. At the same time, a high number of posts (n=938) from the Blackboard were also gathered during the 4 weeks’ collaboration.

In addition, focus group interviews were conducted to properly understand students’ attitudes towards cross-cultural online collaborative learning experience and to elicit their suggestions for improving their experience. Focus group interviews provide multi-faceted instrument that can be used alone or in conjunction with other research methods, allowing the researcher to delve more deeply into the study of a phenomenon and provide enhanced understanding of the research (Vaughn, Schumm, & Sinagub, 1996). Focus group interviews were conducted with 8 students from China and 9 students from the United States, and the scope for these discussions covered participation, affect and cultural identity. The 8 students in China were interviewed as a focus group by one of the authors and the 9 students in the United States were interviewed as a focus group by the teacher in United States. The duration of both interviews was about 1.5 hours each. After interview, the audio of each session was transcribed, and the comments were organized by topic and edited in sequential order until broad themes emerged.

The quantitative data was analyzed using SPSS 21.0 statistical analysis software. The qualitative data, such as the open-ended questions, the posts and the transcripts were analyzed using content analysis method, which is the process of identifying, coding, and categorizing the primary pattern in the data (Patton, 1990). Three researchers carefully reviewed the data and made notes of the important keywords, themes, and categories that emerged from the data. The reviewed transcripts were then analyzed again to compare with previous summaries of key categories and
themes. Finally, the three researchers validated and discussed their coding decisions to reach a consensus for data coding.

Findings

The data obtained in this study was analyzed for three aspects: participation, social and affect aspects.

Participation aspect: Students’ collaboration process was influenced by language and culture

Participation aspect mainly focused on the process of completing a group collaboration activity and mutual inspiration to reach a consensus.

The first question of the questionnaire was about the technological problems using Blackboard learning management system for online learning and communication. The second question focused on the course resources. Analysis of the data showed that there were significant differences between the responses of the Chinese students and the American students for both questions (p<0.05), as shown in table 3. The Chinese students had some technological problems which were not experienced by the American students. On the other hand, the American students thought that the resources were much more helpful to their learning. From the focus group interviews, we found that Chinese students generally had problems in using the learning management system and the course resources in a completely English environment. Language was a barrier for Chinese students, but they expressed a positive attitude to conquer the barrier in future collaboration.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Question No.</th>
<th>China N = 28 Mean(SD)</th>
<th>America N = 34 Mean(SD)</th>
<th>Independent T-Test t</th>
<th>df</th>
<th>Sig(2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Environment</td>
<td>1</td>
<td>3.32 (1.28)</td>
<td>4.41 (0.96)</td>
<td>-3.810</td>
<td>62</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.61 (0.63)</td>
<td>4.36 (0.80)</td>
<td>-4.102</td>
<td>62</td>
<td>.000</td>
</tr>
<tr>
<td>Collaborative Learning</td>
<td>3</td>
<td>3.50 (0.81)</td>
<td>3.70 (1.15)</td>
<td>-.634</td>
<td>60</td>
<td>.528</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3.82 (0.82)</td>
<td>4.08 (0.89)</td>
<td>-1.063</td>
<td>62</td>
<td>.292</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2.32 (0.98)</td>
<td>2.41 (1.21)</td>
<td>-.838</td>
<td>62</td>
<td>.405</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>3.25 (0.93)</td>
<td>3.38 (1.11)</td>
<td>-.527</td>
<td>62</td>
<td>.600</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>3.50 (0.64)</td>
<td>3.50 (1.00)</td>
<td>.065</td>
<td>61</td>
<td>.948</td>
</tr>
<tr>
<td>Attitude to the Learning</td>
<td>8</td>
<td>3.71 (0.85)</td>
<td>3.32 (0.94)</td>
<td>1.782</td>
<td>62</td>
<td>.080</td>
</tr>
<tr>
<td>Experience</td>
<td>9</td>
<td>3.79 (0.79)</td>
<td>3.76 (0.93)</td>
<td>.162</td>
<td>62</td>
<td>.872</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>3.85 (0.91)</td>
<td>3.89 (0.91)</td>
<td>-.162</td>
<td>61</td>
<td>.872</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>3.79 (0.96)</td>
<td>3.68 (1.13)</td>
<td>.442</td>
<td>62</td>
<td>.660</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>4.07 (0.60)</td>
<td>4.24 (0.72)</td>
<td>-1.043</td>
<td>62</td>
<td>.301</td>
</tr>
</tbody>
</table>

From the focus group interviews, we also found that the Chinese students were not familiar with various details of the Chinese ancient culture, such as dragon boat day, the origin of Kuaizi, white snack lady legends, spring festival, etc. In the collaboration process, when Chinese students wanted to share such details of the ancient cultures with the American students, they had to search on the Internet to find both the Chinese version and the English version of these ancient cultural details.

From the content analysis of the forum posts, we found that the Chinese students were not good at asking questions, and they always endorsed the views of the teachers and other students. The views of the American students on others, including peers, teacher and expert were more questioning and critical, and they always asked questions and discussed their views with other students. The reason behind this seems to be the collectivism of Chinese culture and the individualism of American culture, as pointed out by Chen, Hsu, & Caropreso (2006). In the focus group
interviews, we also found that sometimes the Chinese students had the motivation to raise questions and reply to posts, but they felt it difficult to express themselves in English.

The Chinese students attached great importance to the harmonious atmosphere and relationship maintenance in the online collaboration. Even when they disagreed with others, they always exhibited mild and indirect expressions to express their opinions in order to avoid conflicts. The American students, on the other hand, were always focused on the learning task, and when their opinions were different from others, they usually directly expressed what they thought. These two findings could explain the fact that computer-supported collaborative learning (CSCL) in China has primarily remained at a rather superficial stage and the discussions do not go very deep, unlike western countries where CSCL has been found to be very successful (Huang, Zhang, Chen, & Xu, 2007).

Social aspect: Students were interested in each other’s culture

Social aspect mainly focused on mutual cultural consideration, the attitudes towards each other's culture, and culture’s influence on behavior.

“Social lounge” and “cultural orientation” were used as the instruments for social aspects and intended to enable students to get to know each other and understand each other’s culture. The third question of the questionnaire was related to “social lounge”, while the fourth question focused on “cultural orientation”. Analysis of the data showed no significant difference in the attitudes of both the Chinese students and the American students as elicited from the statements “Social Lounge helped acquaint me with class peers” and “Cultural blogs helped foster cultural awareness for cross-cultural online collaborative learning” (p > 0.05). This suggests that the students found these two stages to be somewhat helpful for them in getting to know each other and understand each other’s culture. In the focus interviews, majority of the Chinese students expressed that the social lounge was good for them to get to know the American students, but because they could not see each other and could not communicate synchronously, it was difficult to carry out deeper conversation.

The analysis of the content of the posts during these two stages revealed that the American students were very interested in Chinese culture and they communicated with Chinese students a lot on Chinese traditional culture, asking all sorts of questions regarding various festivals, the traditional legends, the famous great people in old China and so on. At the same time, the Chinese students were very interested in various aspects of American culture, such as Halloween, Christmas, famous movies, etc. But from the focus group interviews, we found that the Chinese students’ unfamiliarity with Chinese traditional culture and their poor written English meant that they could not communicate properly with the American students on cultural issues. As can be seen in table 4, for the fourteenth question in the questionnaire, 28 out of 34 American students expressed that “learning about the Chinese culture” was “most valuable to me about the cross-cultural online collaborative learning experience,” and 22 out of 28 Chinese students expressed that “Culture and Lifestyle” was most valuable to them.

<table>
<thead>
<tr>
<th>Table 4. Analysis of the open-ended questions of the questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items</td>
</tr>
<tr>
<td>The three most significant cultural differences I have observed from the cross-culture collaborative learning</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>The three things I felt most valuable to me about the</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
cross-cultural collaborative learning experience were
Learning about and using technologies to communicate 6 Broaden horizon/Different viewpoints 7

If I had a chance to work with students from a different culture again, I would do the following differently:
Communicate with them more frequently and for longer periods of time 13 Discussions with just individual students not entire classes 10
Have more in-depth conversations 7 More in-depth conversations 9
Have more prior knowledge of each other’s cultures and backgrounds 7 Live chats/instant messaging (Skype) 7

The analysis of the contents of the forum posts also revealed that the Chinese students preferred to post information about themselves and their families at first, and always posted lot of such information, such as birthplace, what their parents did, harmonious relationship of family, personal interest, etc. On the other hand, the American students preferred to introduce only themselves at first, and seldom mentioned their parents. Analysis also showed that the American students always wrote ideas first and then listed evidence, while the Chinese students always wrote evidence first and subsequently provided the conclusions.

Affect aspect: Students’ attitude towards cross-cultural online collaborative learning experience was positive
The affect aspect mainly focused on the attitudes towards the cross-cultural online collaborative learning and satisfaction of team members in the collaboration process.

The results of questions 8-12 and 14-15 revealed that both the Chinese and the American students did not have much collaborative learning experience with students from other cultures prior to this experiment, and most of the students thought that this cross-cultural online collaborative learning experience contributed to their life experience and expanded their global perspective. The same result were also found in the open-ended question “If I had a chance to work with students from a different culture again, I would do the following differently?”, where, as shown in table 4, 13 out of 34 American students said that they would like to communicate and collaborate with the Chinese students for longer time, and 14 of them indicated that they would like to have more in-depth conversations. In the focus group interviews, students from both sides expressed that they really liked this new kind of cross-cultural online collaboration and they wanted to try instant communication tools, such as Skype or video conferencing tools, in order to have more in-depth conversations. The American students also expressed that they should have more prior knowledge of each other’s cultures and backgrounds before carrying out collaborative learning.

Analysis of the chat log revealed the satisfaction level of the team members in the collaborative process. In the first two learning stages, namely “Social Lounge” and “cultural orientation”, there were many posts by each student, and the speed of the replies was also quite high. However, in the third stage of “Technology integration”, posts were mainly those that were required by the teachers, and very few replies were received for each post. This indicated that the students were more satisfied with the first two stages than the third one. The focus group interviews confirmed this where most of the Chinese students expressed that they felt it was not easy to express the professional words in English in the “technology integration” stage. Although the Chinese students wanted to communicate with the American students about their major topics, they could not write many posts and replies due to language barrier.

Discussion
The aim of this research was to find strategies to conduct smooth and effective cross-cultural online collaborative learning. Analysis of the responses received through the questionnaires and interviews, and the analysis of forum posts identified a generally positive feedback from both the Chinese students and the American students on their attitudes towards cross-cultural online collaborative learning. The students felt that this kind of online collaboration could help expand their global perspective to some extent. Moreover, both the Chinese student and the American
students expressed their willingness to have similar cross-cultural online collaborative learning in other courses. At the same time, both the Chinese and the American students realized the importance of having such international exposure and felt that majority of them did not have the chance to communicate face to face with foreigners. Therefore, such cross-cultural online collaborative learning experience would go a long way towards meeting the needs of the students for improving their multi-cultural awareness and cross-cultural collaborative skills.

A number of strategies were identified in this experiment that could promote the implementation of cross-cultural online collaborative learning.

The first strategy identified in the experiment was the topic setting model for cross-cultural online collaborative learning. Collaboration begins with interaction — participants show awareness of each other’s presence and begin to relate as a group (Murphy, 2004). In this study, students had one week to understand each other in “social lounge”, one week to know each other’s culture in “cultural orientation”, and two weeks to exchange their major knowledge on a particular topic. Students were active and engaged in the first two weeks, and they were satisfied with the process of knowing each other and about their culture, which would foster cultural awareness and feeling to relate as a group. Students also showed great interest in each other’s culture. However, in the last two weeks, students were not very engaged, the reason for which was the improperly set topic that did not match with their prior knowledge. Conversations with the teacher in the United States revealed that a proper topic was critical for the collaboration, and setting of the topic should consider the knowledge base and the learning content from both sides, as shown in figure 2. The collaborative learning topic should be the overlap of the three circles which means setting topic should consider the learning content and what the group members from both sides already know. So, the analysis of students’ prior knowledge and analysis of learning content are two main factors that must be considered when setting up the collaborative learning topic. In other words, social interaction and cultural interaction are the basis for cross-cultural online collaborative learning, and the collaborative learning topic should overlap the learning content and the knowledge groups members from both sides already have.

![Figure 2. Cross-cultural collaborative learning topic set Model](image)

The second strategy identified in the experiment was teacher’s task model for implementing cross-cultural online collaborative learning. Teachers design the whole learning environment, control the collaboration process, and evaluate the learning outcomes. It is very important to identify what teacher should do for implementing cross-cultural online learning. Firstly, teachers from both sides should consider group division, which is based on two principles: heterogeneity in a group and homogeneity between groups. Secondly, designing learning environment could consider three sequenced stages including “social lounge”, “cultural orientation”, and “technology
integration.” Thirdly, evaluation methods should contain three domains, namely cultural identity, collaboration process and affect.

The third strategy identified in the experiment was the use of bilingual language resources in the learning environment to overcome language barriers during cross-cultural online collaborative learning. The Chinese students had problems in English learning environments, which added extra cognitive load for them. In such situations, use of bilingual language resources in the learning environment could serve towards reducing the cognitive load, fostering better communication and more effective collaborative learning processes. In the data collected, we found that the Chinese student were not familiar with Chinese ancient cultures, such as dragon boat day, Kuaizi, white snack lady, spring festival, and etc. Therefore, the basic cultural terms could be listed in both languages as resources for students.

The fourth strategy identified in cross-cultural online collaborative learning was the ability to take into consideration the influence of the unfamiliar culture on interaction. Culture has a very important influence on students’ learning style, and students from different cultures behave differently from each other (Kim & Bonk, 2002). Troubles will arise if cultural differences are not taken into account, and as a result, the collaboration may even fail between the students of different cultural backgrounds. For example, the Chinese students were not good at asking questions and always used a mild and indirect way to express their different opinion; while the American students always focused on the learning task and expressed their views directly.

Students’ positive attitudes towards cross-cultural online collaborative learning demonstrated that this kind of collaboration has a future. However, further research on the collaborative process, the culture’s influence and the language barriers will be important to improve such kind of collaborations in terms of smoothness of the process and better effectiveness.

**Conclusions**

This study found that the implementation of cross-cultural online collaborative learning should (1) take the social interaction and cultural interaction as the basis for collaboration and set collaborative topics that overlap the knowledge base of both sides; (2) consider teacher’s task model as a tool for teachers to implement cross-cultural online collaborative learning; (3) provide bilingual language resources; and, (4) pay attention to culture’s influence on the collaborative learning process.

The results of this study should not only be useful for future Chinese-American student collaborations but the generic nature of the results should also be useful for other cultures. Feedback from the participants of this research also indicates that university policy makers should advocate cross-cultural online collaborative learning in order to promote students’ global perspective and multi-cultural understanding.
The students in this study also expressed desire for more synchronous collaborations and indicated that they would want to have online face to face communications. As the high bandwidth internet access is becoming more accessible in educational institutions all over the world, the next step for cross-cultural online collaborative learning should be to integrate synchronous communication tools (such as WebEx, JoinNet and Adobe Connect) into the cross-cultural collaboration.

The findings also indicated that students found it difficult to collaborate when they did not have sufficient background knowledge. Based on the findings and further analysis, a topic setting model is proposed in this paper. However, the model has not yet been verified. As topic setting was found to be a critical factor in the success of collaborative learning, further research is needed to test the model.

This study focused on various strategies for smooth and effective implementation of cross-cultural online collaborative learning by examining the aspects of social, participation, and affect. Research is needed to also understand the effects of cross-cultural online collaborative learning on students’ performances, as that is the ultimate aim of any learning process.

Acknowledgements

This paper is supported by the 2012 national project of 12th five year plan for education science “the implementation strategy for cross-cultural synchronous online classroom (ECA120341)”. The authors also acknowledge the support of NSERC, iCORE, Xerox, and the research related gift funding by Mr. A. Markin.

References


Learning to become a teacher in the 21st century: ICT integration in Initial Teacher Education in Chile

Mario Brun* and J. Enrique Hinostroza
Institute of ICT in Education, University of La Frontera, Chile // mariobrun@cidet.com.ar // enriquehinostroza@iie.cl
*Corresponding author

(Submitted September April 25, 2011; Revised April 23, 2012; Accepted April 12, 2014)

ABSTRACT
This paper presents the most relevant results from a national study about the availability and use of ICT in 46 Initial Teacher Education institutions in Chile, implemented during 2009 as part of the OECD (Organization for Economic Cooperation and Development) international project “ICT in Initial Teacher Training.” Main findings show an overall favorable context for the pedagogical use of ICT in such institutions, in terms of ICT infrastructure, support, policies and teachers’ self-reported ICT related skills. In addition, teachers report a quite frequent use of some ICT resources, giving a high importance to students’ learning of ICT. However, ICT integration is limited to a few specific resources (mainly computers and projectors), mostly applied to perform “traditional” pedagogical activities. Therefore, the expectations about the improvement of teaching and learning in Teacher Education in Chile through the integration of ICT, are not been fulfilled yet. This paper can constitute a significant contribution for developing more innovative and better quality pedagogical practices in this education level.

Keywords
Initial teacher training, Teacher education, ICT in education, Learning, ICT, Chile, ITT

Introduction
This paper presents the results of a study about the availability and use of ICT in the Initial Teacher Training (ITT) in Chile, conducted in the context of the international study “ICT and Initial Teacher Training” developed since late 2008 by the OECD. It stems from the concern shared by several OECD member countries about what they see as a deficit in their ITT systems in providing the required vision, experience and skills for enabling future teachers to integrate ICT into their professional practices in primary and secondary education levels.

Although several studies show that teachers are among the most skilled ICT users, it seems that they are still unable to apply these skills to the way they teach (Pedró, 2009). Some feasible explanations for this paradox highlight the experiences in ITT among the most important factors. However, a claim about the existence of a link between ITT and teaching skills with ICT needs to be supported by empirical evidence: this study aims at generate relevant information on this matter.

In this context, the international study investigates the current use of ICT in ITT in a number of countries from a comparative perspective, analyzing the future prospects and the possibility to formulate recommendations on applicable policies and strategies. Accordingly, the Government of Chile decided to participate in this project to get a deeper understanding about the use of ICT resources in its ITT institutions.

The obtained results provide an overall picture that could contribute to strengthen the processes related to the integration of ICT in this educational level.

National context about ICT in Initial teacher education in Chile
Initial Teacher Education in Chile is conducted by higher education institutions which are officially recognized by the State: Universities and, under certain conditions, Professional Institutes. In early 2009, approximately 65 of these institutions were offering more than 700 programs oriented to all levels of the national system: preschool, primary, special and secondary education.
Chile has a comparatively long history in the introduction and use of ICT in K-12 through the Enlaces project, its “official” nationwide initiative for ICT in education: it has been aiming to integrate ICT as a pedagogical resource for all students and teachers in Chilean public schools (Hinostroza, Hepp & Cox, 2009). However, just recently the Ministry of Education (MoE) has started to pay attention to the use of ICT in ITT institutions: during the last years, the MoE has been developing some actions explicitly aimed at fostering the integration of ICT in ITT, such as the development and promotion of ICT related standards and the integration of ICT-based items in the teachers’ professional certificate implemented on a voluntary basis. Regarding this certificate, the “Inicia” program - conducted since 2008 by the MoE for encouraging the transformation of institutions, curricula and practices in ITT and ensuring the professional quality of graduated teachers at a national level (Meckes, 2010) - proposes a set of pedagogy and content related standards, including the evaluation of new teachers’ “computational skills.”

In addition, the MoE produced the report “ICT Standards for Initial Teacher Training” (Ministry of Education, 2006) and later, the “Functional Map for ICT training and teaching skills” (Ministry of Education, 2007): These documents present an articulated set of teaching competences organized in five functional dimensions, viz. pedagogical, technical, managerial, ethical and legal, as well as professional development.

The presented framework aimed at support the implementation of an ICT curriculum in ITT, opening a space for reflection and debate about the ways in which teacher education institutions are responding to the demands of current society, from a more complex and global perspective related to the appropriation of ICT.

Research background

The information and innovation-based society demands from the education systems continuous improvements to prepare new generations for taking full advantage of the new socio-cultural and economic conditions. Consequently, education systems are recognizing teachers’ key impact on students’ performance (Barber & Mourshed, 2007) and they are increasingly focusing on ITT institutions, transforming them into the new “locus of attention.”

Considering the widespread availability of ICT in schools and society, and being aware of the “difficulties” to easily demonstrate a positive impact of ICT in students’ learning outcomes, researchers in this field are also shifting their attention to the use of ICT in ITT. Although it should not to be assumed that the use of ICT is per se a necessary an sufficient condition for a good pedagogical practice, in cases like Chile - with a major nationwide initiative on ICT in Education implemented for the last 15 years - the preparation of future teachers for using ICT resources in teaching constitutes a real need that cannot be neglected.

Consequently, the process of an effective use of ICT in schools should start in the ITT (Kay, 2006) since this training stage is a key factor for promoting ICT-based skills related to future teachers work at schools (Graham & Thornley, 2000). Nevertheless, the lack of explanation about why many students in ITT feel unprepared to integrate ICT in their future pedagogical practices (Twidle et al., 2006) reveals the complexity of this issue.

Drawing on the results of the review done by Enochsson & Rizza (2009) it seems that ICT aren’t being used in a regular or systematic way in ITT. The main factors that could explain this situation can be summarized as follows:

- **Policies on ICT integration:** Research in K-12 shows that the institutional vision about the role of ICT in teaching and learning is a key factor for its effective pedagogical integration (Anderson & Dexter, 2000; Law, Pelgrum & Plomp, 2008). However, institutional policies are among the least mentioned factors in research works on Teacher Education.

- **Curricular integration of ICT:** Technologies can be integrated as a specific subject or as contents included across most of the disciplines (Schmidt, 1998). It’s generally accepted that the inclusion of ICT just in one or a few subjects is an ineffective way for integrating it in ITT (Kirschner & Davis, 2003; Sardone & Devlin-Scherer, 2008), whereas a cross-curricular approach is more effective (BECTA, 2006).

- **ICT infrastructure and support:** Availability and access to ICT are required conditions to integrate it in teaching and learning (Anderson & Petch-Hogan, 2001; Judge & O’Bannon, 2008), but they aren’t enough to promote its pedagogical use (Selwyn, 2004). Current literature claims that ICT infrastructure is no longer a problem in Higher Education institutions (Enochsson & Rizza, 2009): although there is little information about computers
availability in ITT, a ratio of 14 students per computer in Sweden (Enochsson, 2010) and 16 in Norway (Tømte, Hovdhaugen & Solum, 2010) could be mentioned as references. Additionally, many authors (e.g., Tong & Trinidad, 2005) highlight the importance of an adequate support to teachers.

- **Actors’ competences and visions about ICT use:** The lack of teachers’ ICT-related pedagogical competences has been pointed among the main obstacles for technologies use in teaching (Vrasidas & McIsaac, 2001; Afshari et al., 2009). This demand for training is reflected in the international survey TALIS (OECD, 2009) where the acquisition of “ICT teaching skills” was the second most frequently prioritized item. This prospect about professional development is also backed-up by different studies related to the role of ICT in ITT (Georgina & Olson, 2008; Buckenmayer, 2008) and the “TPACK” (Technological Pedagogical and Content Knowledge) framework (Mishra & Koehler, 2006; Archambault & Crippen, 2009)—which claims for the existence of technological, pedagogical and disciplinary knowledge, not only as separate components, but also as interrelated contents. Furthermore, teachers’ comfort in using ICT, and their attitudes toward the pedagogical use of technologies, are also mentioned as relevant factors (Judge & O’Bannon, 2008; Drent & Meelissen, 2008). Regarding students, they seem to feel comfortable with ICT but they rarely use it for instructional purposes (Twidle et al., 2006); accordingly, many future teachers believe they aren’t being fully prepared for using ICT in teaching (Enochsson & Rizza, 2009), claiming that their training in this matter is patchy and inconsistent (Moeini, 2008).

- **ICT use in teaching and learning activities:** Different studies report a relatively low, basic and repetitive use of ICT in K-12, associated to simple and traditional pedagogical activities, rather than to more complex and/or innovative ones (Trucano, 2005; Law et al., 2008). Many authors consider ICT as an ‘appendix’ to other pedagogical resources (Whitworth & Berson, 2003) or as a support for existing teaching practices (Ottesen, 2006), while others posit an underuse of the ICT resources available at educational institutions (Cuban, 2001; Jimoyiannis & Komis, 2007). In this context, the students’ lack of confidence to integrate ICT in their future teaching seems to be related to the scarcity of pedagogical activities with ICT during their ITT. Perhaps “students don’t know how to use new technology in their classroom instruction when they go to (...) schools, because they lack previous practices of applying ICT into curriculum” (UNESCO, 2002, p. 69).

These studies highlight the importance of future teachers’ understanding, not only about how to use ICT but also about the specific knowledge involved in the pedagogical use of ICT, and the full comprehension of the impact of technology on learning processes. Although more and more teachers and students are achieving increasing levels of ICT-related skills and the availability of technology is constantly growing in educational institutions, the underlying knowledge is still not been fully transferred to teaching practices (Player-Koro, 2007; Pedró & Ananiadou, 2009).

Thereby, the main purpose of this paper is to provide a general description of the current integration of ICT in ITT in Chile, with the aims of respond the following research questions:

- What are the institutional contexts regarding the use of ICT in ITT?
- To what extent and in what ways is ICT been used in teacher education institutions?
- How are students been prepared in ITT institutions to integrate ICT in their future teaching?
- What are the main pedagogical activities performed with ICT?
- What are the most relevant factors considered as barriers and enablers to ICT integration in ITT?
- What recommendations could be issued to policy-makers in this field?

This paper started by presenting the national context in this domain, and the research background. Then, the method is described and main results are shown. Finally, it presents the most relevant conclusions that might have implications for policy-design and further research.

**Method**

To address the research questions, this study used a methodological approach combining both quantitative and qualitative data collection techniques, such as a survey applied to different actors in Initial Teacher Education institutions in Chile using a self-administered paper questionnaires, and several case studies (mainly involving interviews and focus groups) carried out in ITT institutions selected according different criteria.
The questionnaires provided by the international study covered diverse issues (namely: policies, curricular integration, infrastructure and support, frequency of ICT use, teaching and learning activities, enablers and barriers to ICT integration, among others), aimed to get an overview of the use of ICT in ITT. These instruments were firstly translated into Spanish; then adapted to the local context and, finally, some complementary questions were added.

Questionnaires were applied in 46 teacher training institutions participating voluntarily in the study (about 75% of national total), targeting different actors whose corresponding samples were: (a) 46 deans (or equivalent authorities); (b) 495 teachers (considering different programs: primary education, and secondary education in mathematics, national language, social sciences and natural sciences); (c) 164 mentors (supervising students during their teaching practices); (d) 1,675 students (selected from different years); (e) 233 recent graduates; and (f) 50 technical and/or pedagogical responsible (in charge of pedagogical ICT related issues and/or technical aspects, such as maintenance). Fieldwork was conducted from June to September, 2009.

Results have a 95% significance level, allowing to perform inferential processes over the total population. The main quantitative statistical procedures used in this paper correspond to descriptive analysis (i.e., frequencies, and mean differences using t-test and one-way ANOVA), and they were performed with the software SPSS-Statistics v.17.0©.

**Results**

Results are structured in seven sub-sections, according to the dimensions addressed in the theoretical framework: (1) Institutional policies and practices for ICT integration; (2) Infrastructure, ICT resources and support; (3) Actors’ confidence in the use of ICT; (4) Visions about the pedagogical use of ICT; (5) Use of ICT resources in teaching and learning activities; (6) Teaching and learning activities performed; (7) Main barriers and enablers for ICT integration.

**Institutional policies and practices for ICT integration in ITT**

Results show that 63% of deans mentioned the existence of an institutional policy to integrate ICT in teaching, and almost 70% of them reported to have a department specifically oriented to support pedagogical innovations in teaching, including the use of ICT.

They reported that ICT is mostly integrated into the curriculum only in some specific subjects (83% of ITT institutions) instead of across several subjects (cross-disciplinary). Moreover, 56% of the institutions included explicit objectives related to students’ development of ICT-related pedagogical competences in half or less of their programs.

Aligned with this, only 39% of the deans answered that “innovating in teaching using ICT” or “implementing specific actions to integrate ICT” are among the highest institutional priorities regarding teachers’ acquisition of competencies to perform instructional practices (see Figure 1).

![Figure 1. Percentages of deans assigning a high priority to teachers’ acquisition of ICT related competences](image_url)
Consistent with the previous figure, 65% and 68% of the institutions only provided optional courses related to the technical and pedagogical use of ICT, respectively; whereas only less than 20% of such courses were mandatory (Figure 2).

Summarizing, it can be argued that despite the intentions associated to incorporate ICT in teaching and learning, relatively few ITT institutions formally integrate technologies into their curriculum; and for most of them, this is not a high priority (notwithstanding the fact that many institutions do have a department supporting pedagogical innovations). This finding is also consistent with the predominantly optional character of the ICT related courses they offer.

Infrastructure, ICT resources and support

An adequate infrastructure and the availability of technical and pedagogical support are necessary conditions to integrate ICT into educational institutions (Chen & Chang, 2007). Results from this study show that the students per computer ratio in ITT institutions in Chile is 17, whereas all participating institutions provide Internet access (96% with broadband access). Additionally, 59% of the institutions reported to have LMS/VLS systems, supporting 53% of their courses and been used by 44% of teachers at least once a week. Furthermore, almost 75% of teachers and mentors reported to have access to a computer in their institutions.

Figure 3. Percentages of teachers reporting the availability of ICT resources in the classrooms
About the availability of ICT resources in the classrooms, 83% and 82% of teachers reported the existence of computers and audio equipment—respectively—in “some” or “almost all” classrooms; with a lower availability of other resources such as projections systems (64%), interactive whiteboards (52%) and videoconference systems (51%), as Figure 3 shows.

About students’ access to ICT resources, 85% of the students reported to have free access to computers, 50% to projection systems, and only 35% to LMS/VLS systems. Moreover, less than 10% of them reported a free access to interactive whiteboards and digital cameras (Figure 4).

![Figure 4](image-url) Percentages of students reporting to have free access to different ICT resources in their institutions

Regarding the availability of technical and pedagogical support, 96% and 70% of teachers reported—respectively—that these services are available in their institutions. In turn, only 63% and 57% of students mentioned the availability of such support (Table 1). In addition, results show that the overall quality was ranked between “medium” and “high” for both services.

<table>
<thead>
<tr>
<th>Actor</th>
<th>Technical</th>
<th>Pedagogical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers</td>
<td>96%</td>
<td>70%</td>
</tr>
<tr>
<td>Students</td>
<td>63%</td>
<td>57%</td>
</tr>
</tbody>
</table>

In summary, results show that the most available and accessible resources are computers and projection systems; conversely, interactive whiteboards, video-conferencing systems and digital cameras are the least available. Besides, teachers reported more access to ICT infrastructure than students, and a higher availability of technical and pedagogical support. In general, it could be argued that access to computers and projection systems for teachers and students in ITT is widespread; however, availability and access to other resources is relatively scarce (particularly interactive whiteboards).

**Actors’ confidence in the use of ICT**

Broadly speaking, teachers and students reported high levels of comfort in ICT use, either at home or in classes: 78% and 67% of teachers reported to feel “very comfortable” while using ICT at home and in classes, respectively. As seen in Figure 5, students show similar levels of comfort (66% and 62%, respectively) The fact that more teachers reported to feel comfortable with ICT when compared to students, is consistent with the previous results showing teachers’ higher access to ICT.
Regarding teachers’ confidence in performing different activities with ICT, Figure 6 shows that the highest levels were related to: “attach a file and send it via email” ($M = 3.79, SD = 0.559$); “produce a letter with a word processor” ($M = 3.74, SD = 0.559$), and “file a digital document in a computer” ($M = 3.73, SD = 0.726$). Conversely, the activity with the lowest confidence level was “making purchases and payments on the Internet” ($M = 3.13, SD = 1.129$).

In turn, students reported an ICT use of 14 hours per week for academic purposes and another 15 hours for its use at home. Additionally, only 38% of students took workshops or courses about ICT during the two years preceding the study, and just 10% of them have been personally engaged in innovative projects developed by their teachers involving the use of ICT.
Finally, approximately 70% of students reported to feel confident to integrate ICT in their future teaching activities. However, results from the case studies showed that many interviewed students considered that the ICT-related pedagogical skills acquired during their ITT were incomplete and insufficient.

Summarizing, the majority of teachers and students declared to feel confident to take advantage of ICT and to integrate ICT in their teaching practices, however, based on the case studies, there is some evidence pointing at the need to more deeply analyze the type and quality of the ICT-related pedagogical skills that students are learning.

**Visions about the pedagogical use of ICT**

The questionnaire asked teachers and students about the importance of new teachers learning about using ICT for different activities. The corresponding responses were grouped in four indicators:

- **Communication and networking**: This indicator groups the items related to the use of ICT communication tools.
- **Own development and learning**: Item related to ICT use in professional development activities.
- **Organization of teaching work**: This indicator groups the items related to ICT use for administrative and planning activities.
- **Integration in teaching**: This indicator groups the items related to ICT use in teaching and learning activities.

Results show that both teachers and students considered as “quite” or “very” important that new teachers learn how to use ICT to perform those activities (Table 2). It’s noteworthy that differences between and within groups aren’t significant, revealing the high degree of consensus among teachers and students about the importance of learning ICT for teachers’ professional activities.

**Table 2.** Importance attached by teachers and students to students’ learning about the use of ICT for different purposes

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Teacher Items</th>
<th>Cronbach’s alpha</th>
<th>M</th>
<th>SD</th>
<th>Student Items</th>
<th>Cronbach’s alpha</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication and networking</td>
<td>4</td>
<td>.854</td>
<td>3.7</td>
<td>.416</td>
<td>4</td>
<td>.877</td>
<td>3.3</td>
<td>.629</td>
</tr>
<tr>
<td>Own development and learning</td>
<td>1</td>
<td>-</td>
<td>8.8</td>
<td>-</td>
<td>1</td>
<td>3.5</td>
<td>.644</td>
<td></td>
</tr>
<tr>
<td>Organization of teachers’ work</td>
<td>4</td>
<td>.921</td>
<td>3.8</td>
<td>.366</td>
<td>4</td>
<td>.888</td>
<td>3.6</td>
<td>.529</td>
</tr>
<tr>
<td>Integration in teaching</td>
<td>11</td>
<td>.940</td>
<td>3.6</td>
<td>.458</td>
<td>11</td>
<td>.949</td>
<td>3.4</td>
<td>.572</td>
</tr>
</tbody>
</table>

*Note. Means measured in a Likert scale: 1 = not important; 2 = rather important; 3 = quite important; 4 = very important.*

Based on these results, it’s reasonable to argue that teachers and students are aware of the importance of new teachers’ learning about how to use ICT in their professional activities. However, it also emerges from case studies that students perceive they are not learning enough about the use of these tools for their future teaching.

**Use of ICT resources in teaching and learning activities**

Regarding the frequency of ICT use in teaching and learning activities, results show that 76% and 74% of teachers used projection systems and computers, respectively, in half or more of their lessons, whereas most of the remaining hardware resources show percentages below 10%. It’s noteworthy that while 44% of teachers reported the use of digital learning resources in half or more of their lessons, only 13% or less reported the same frequency regarding multimedia or educational software (see Figure 7).

In relation to the relatively low percentage of teachers using some specific ICT resources, it could be argued that it’s a consequence of their lower availability. Nevertheless, when comparing the availability of most of ICT resources (except for computers and projection systems) with their frequencies of use, it can be seen that availability is, in general, quite higher than use (Figure 8). For example, whereas about half of the teachers reported the availability of interactive whiteboards, digital cameras or video-conferencing systems in at least some classrooms, less than 10% of them reported to use them in half or more of their classes.
In addition, teachers were asked about how often they taught their students how to use ICT resources for teaching. Results show that whereas one third of teachers reported to teach their students how to use computers and projection systems for teaching in half or more of their classes, less than 10% reported to teach the use of the other ICT resources. This was somewhat confirmed by students’ interviews, when they claim: “They always require us to make PowerPoint presentations, but in fact this is the only resource we use, the only thing.” This result is interesting and opens questions about the need to fully provide students the conditions for developing the required level of technological-pedagogical-content knowledge (Mishra & Koehler, 2006) to fully take advantage of ICT resources in teaching.

Summarizing, results show that teachers tend to use quite frequently computers, projection systems and some software products during their lessons; however, they do not frequently use other ICT resources. Moreover, they show that relatively few teachers teach their students how to use the available variety of ICT resources for teaching, mainly focusing in teaching the use of computers and projectors, and mostly using presentation software, as students declared.
Teaching and learning activities

The study also asked teachers and students about how often they performed different activities, both overall – regardless the use of ICT – and specifically by using ICT. Particularly, three groups of activities were included: (a) teachers’ educational management related activities; (b) teaching activities; and (c) learning activities.

Regarding teachers’ educational management related activities Figure 9 shows the percentages of teachers reporting the performance of different activities, at least once a month. The majority of them reported to perform the following overall activities: preparation of general classes (92%); search useful learning resources (87%); communication with students (87%) and colleagues (86%); organization of teachers’ work (85%) and design own learning resources (83%). Conversely, relatively few teachers reported: participation in courses or workshops (24%); participation in collaborative projects (37%); and identification of pedagogical situations for the use of ICT (51%).

![Figure 9](image)

Figure 9. Percentages of teachers reporting to perform activities associated to their educational management related work, overall and with ICT (Percentages correspond to a frequency of, at least, once a month)

When considering the percentages of teachers reporting to perform these activities but by using ICT, it emerges that they follow a similar pattern: this finding is quite interesting, since indicates that ICT is not being used in any particular activity more than others, and that teachers’ use doesn’t seem to modify the patterns of implementation corresponding to their practices regardless the use of ICT.

Regarding teaching activities, Figure 10 shows the percentages of teachers reporting to implement different activities at least once a month with and without the use of ICT. In this case, the majority of teachers report they search for information (76%); develop products/reports (63%) and assess learning (66%); meanwhile, relatively few teachers report performing lab activities (18%); extra-institutional activities (19%) and communication with other experts/teachers (31%). Although the percentage of teachers reporting to perform these activities with the use of ICT tends to follow a similar pattern, it should be noted that some activities (e.g., research projects and activities accessed independently by students) show smaller differences, probably implying that, when they are performed, they are mostly done by using ICT.

Finally, Figure 11 presents the percentages of students reporting the implementation of different learning activities, at least once a month. Results show that the majority of students report to: make presentations (75%), work as a
group in class at same pace (70%) and develop learning materials (67%). Conversely, relatively few students report to communicate with external actors (34%) and contribute to the community through learning activities (36%). When comparing the percentage of students reporting the implementation of these activities, in general and with the use of ICT, results reveal the existence of some differences showing that ICT is relatively more frequently used by students in some activities, viz. communicate with external actors, develop learning materials, contribute to the community through learning activities and make presentations. Conversely, ICT is relatively less frequently used in activities such as participate in collaborative projects/activities, work individually in class at own pace, and self-evaluate or co-evaluate with peers. These findings open the discussion about the specific role played by ICT in performing these activities.

**Figure 10.** Percentages of teachers reporting to perform teaching activities, overall and with ICT (Percentages correspond to a frequency of, at least, once a month)

**Figure 11.** Percentages of students reporting to perform learning activities, overall and with ICT (Percentages correspond to a frequency of, at least, once a month)
The comparison of the frequencies of teachers and students activities performed with ICT shows that relatively less teachers report to use ICT for teaching activities than students do for learning activities (30% and 34.5%, respectively); however, the percentage of teachers reporting educational management activities with ICT reach the highest percentage (51.9%).

Summing up, results show that teachers tend to use ICT consistently across their educational management related activities and in most of their teaching practices, without showing any evidence of ICT contributing to the implementation of some specific activities over others. Nonetheless, the use of ICT by students shows some evidence of differentiated roles in certain particular learning activities.

**Main barriers and enablers for ICT integration**

Regarding the main barriers to integrate ICT in ITT, they can be grouped in two categories: (a) personal factors (related to the actors involved: e.g., lack of ICT competencies; visions about ICT; etc.); and (b) institutional factors (aspects clearly related to institutions: insufficient equipment and lack of institutional policies; among others).

As Figure 12 shows, the institutional obstacles most frequently mentioned by teachers were: lack of equipment or digital learning resources (41% and 46% respectively); lack of pedagogical or technical support (44% and 39% respectively); and lack of institutional interest or policies oriented to promote ICT integration (28% and 42%, respectively). Regarding personal barriers, the most frequently mentioned were: lack of time to explore ICT (52%); lack of general or pedagogical ICT-related competences (31% and 29%, respectively); lack of teachers’ confidence or flexibility to try new approaches (24% and 18% respectively) and students’ lack of ICT skills (24%).

**Figure 12. Percentages of teachers considering different factors as barriers to ICT integration**

Results obtained allowed to develop composite indicators to summarize personal and institutional barriers, they show that -according to teachers- the general mean of institutional obstacles is significantly higher than personal barriers ($t = -7.825$, $p = .000$). Conversely, although students show a higher emphasis on personal barriers, the difference is not significant (Table 3).
Table 3. General means of personal and institutional barriers, according to teachers and students

<table>
<thead>
<tr>
<th>Composite indicator</th>
<th>Teachers</th>
<th></th>
<th></th>
<th>Students</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Items</td>
<td>Cronbach’s</td>
<td>M</td>
<td>SD</td>
<td>Cronbach’s</td>
<td>M</td>
</tr>
<tr>
<td>Personal obstacles</td>
<td>7</td>
<td>0.893</td>
<td>2.06</td>
<td>0.683</td>
<td>0.862</td>
<td>2.53</td>
</tr>
<tr>
<td>Institutional obstacles</td>
<td>7</td>
<td>0.911</td>
<td>2.33</td>
<td>0.718</td>
<td>0.925</td>
<td>2.42</td>
</tr>
</tbody>
</table>

Note. Likert scale: 1 = Strongly disagree; 2 = Disagree; 3 = Agree; 4 = Strongly agree (that each group of factors is a barrier to ICT integration).

Regarding the importance attached by teachers and students to different aspects as enablers for the integration of ICT in teaching and learning, the study asked about several factors that can be grouped in three categories: (1) infrastructure and support (ICT resources, technical and pedagogical support); (2) professional development (courses and workshops on ICT for general and pedagogical use); and (3) policies and management (cross-curricular ICT integration, incentives to use ICT, among others). Consistently with the above mentioned barriers, results show that all categories are considered as “quite” or “very” important enablers, by teachers and students (Table 4). Since differences between groups are not statistically significant, it could be said that all dimensions show similar levels of importance for both actors.

Table 4. Importance attached by teachers and students to several aspects related to the promotion of ICT integration

<table>
<thead>
<tr>
<th>Importance attached to the following factors</th>
<th>Teachers</th>
<th></th>
<th></th>
<th>Students</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure and support</td>
<td>3.56</td>
<td>0.463</td>
<td>3.38</td>
<td>0.597</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional development</td>
<td>3.57</td>
<td>0.593</td>
<td>3.49</td>
<td>0.648</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policies/Management</td>
<td>3.49</td>
<td>0.583</td>
<td>3.34</td>
<td>0.605</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. In the scale used, 3 is equivalent to “quite important” and 4 to “very important.”

In summary, results show that, for teachers, institutional barriers to integrate ICT are more important, whereas institutional and personal barriers as similarly relevant for students. In addition, students and teachers assigned similar importance to all groups of enablers for the integration of ICT.

Discussion

Results presented in the previous section will be discussed separately as follows:

- Regarding institutional policies and practices, despite the institutional aims for integrating of ICT in teacher education, there are few institutions that formally put them into practice, revealing that this process is not among their highest priorities, which is consistent with the findings reported by Enochsson & Rizza (2009). Additionally, most of the institutions integrate ICT in the curriculum as specific subjects or courses, rather than as a cross-disciplinary topic to be included in most of the traditional subjects. This approach has shown to be less effective to promote a full and effective ICT integration (Kirschner & Davis, 2003; BECTA, 2006).

- About infrastructure and support, results show that, in general, institutions are well-equipped: For example, students per computer ratio in Chilean institutions is similar to those from European countries (Enochsson, 2010; Tømte et al., 2010). Particularly, there are adequate levels of availability and access to ICT resources, mostly computers and projection systems, as well as regarding technical and pedagogical support.

- Teachers and students show quite high levels of confidence in ICT use, whereas students report an intense use of ICT; consequently, the majority of them declared to be confident in their abilities to integrate ICT in their future teaching activities. However, it was also found that teachers only teach students the use computers and projectors, disregarding other ICT resources such as interactive whiteboards. This finding is important and opens the question about the real understanding of the potential of ICT in teaching and learning as a key factor on the lack of the required competencies to teach with technology. This is also supported by the findings of the TALIS study (OECD, 2009), which show that one of the professional development priorities most mentioned by teachers is...
related to the integration of ICT into their teaching (which is also concordant with the study of Jimoyiannis & Komis, 2007).

- The actors’ visions about ICT reveal a general awareness of the importance given to new teachers’ learning about how to use ICT for pedagogical purposes (Littrell, Zaguny & Zaguny, 2005); however, as mentioned before, these visions has not been fully put in practice yet (Jamieson-Proctor, Burnett, Finger & Watson, 2006).

- Regarding pedagogical activities, results show that teachers tend to use ICT consistently across their educational management-related activities and in most of their teaching activities, without revealing any evidence of ICT contributing to perform some activities more than others. Nevertheless, ICT use in learning activities shows some evidence of differentiated roles in specific activities such as “communicate with external actors,” “develop learning materials,” “contribute to community through learning activities” and “make presentations.” Most of these activities are more related to a “student-centered” pedagogy, coincidently with the findings reported by Inan, Lowther, Ross & Strahl (2010) which claim that ICT use helps to promote this pedagogical approach.

- Moreover, results show that teachers and students are not frequently using ICT in activities associated to students’ assessment (consistently with the results reported by Law et al., 2008). However, the highest frequencies of ICT use by teachers and students correspond to the most basic and conventional activities, such as “prepare general classes” and “organize teachers’ work” (educational management-related activities); “search information” and “develop products” (teaching activities); “make presentations” and “work as a group at the same pace” (learning activities). This is also consistent with previous findings (Trucano, 2005; Law et al., 2008) and particularly with the claim that ICT is just another pedagogical resource to support preexistent teaching activities (Ottesen, 2006) rather than to revolutionize them (Karasavvidis, 2009).

- Concerning the main barriers and enablers to ICT integration in ITT, results are relatively homogenous and show little differences among the diverse actors and dimensions. However, according to teachers, most obstacles are institutional: this is consistent with the preliminary findings reported by Pedró & Ananiadou (2009).

Broadly speaking, it could be said that the foundations for a widespread integration of ICT in the ITT seem to be present in Chile, but this hasn’t been enough to fully harness the potential of ICT for supporting teaching and learning activities, which is aligned to the findings reported by other authors (see for example Player-Koro, 2007). According to Fluck (as cited in Pearson, 2003), “pre-service teachers are expected to graduate (…) with the same competencies that working teachers have gained”; nevertheless, results show that new teachers are not learning yet how to exploit the potential of ICT in their future professional activities.

Conclusions

Based on the presented results, it can be said that institutional contexts are -generally speaking- auspicious to the integration of ICT in ITT: the existence of institutional policies on this matter; the overall infrastructure (e.g., the students per computer ratio); the availability and access to ICT resources, and the availability and quality of technical and pedagogical support have reached adequate levels, showing that these aspects shouldn’t be a cause of particular concern in this ITT. In addition, teachers use ICT in teaching quite frequently and, to some extent, they also teach their students how to use ICT in teaching, although these latter practices are only restricted to some ICT resources. However, the use of technologies still seems to remain bounded to a set of basic teaching and learning activities, whereas the more advanced and complex pedagogical activities are significantly less frequent. Therefore is not unexpected that although teachers show high self-reported levels of comfort and confidence for using ICT, they ask for more professional development opportunities to integrate ICT into their teaching activities, recognizing implicitly their lack of competencies in this field and confirming the importance to teach future teachers how to integrate ICT in teaching and learning.

Consequently, it could be considered that the expectations about the pedagogical integration of ICT in Chilean institutions of ITT are not been fulfilled yet, since the majority of students are not being taught how to use in teaching the full range of available ICT resources, and nor are experiencing situations where ICT is used as an innovative pedagogical tool for acquiring the specific knowledge involved in some teaching practices (e.g., the ability to solve problems by using ICT) and the full comprehension of its impact on learning processes.
In this scenario, both professional development opportunities and pedagogical support for teachers focused on the pedagogical integration of ICT should be fostered in order to take more advantage of the current favorable context and to promote more innovative pedagogical practices which provided students relevant experiences involving the use of ICT in teaching. Additionally, more research in this area should be encouraged, to deepen the understanding of the contribution of ICT in specific pedagogical activities, thereby enlarging the knowledge base in this field. Results might also contribute to shed light on policy-related issues oriented to foster a more complete, updated, innovative and high-quality Initial Teacher Education in Chile.

Acknowledgements

Our gratitude to the Chilean Ministry of Education, through its Center for Education and Technology (CET) and to the Centre for Research on Educational Policy and Practice (Grant CIE01-CONICYT), for supporting this research. We are also grateful to the authorities and other actors from the institutions participating in this research, for their collaboration and readiness.

References


Ministry of Education. (2006). ICT standards for the initial teacher training. Santiago, Chile: Ministry of Education of Chile


The Game Embedded CALL System to Facilitate English Vocabulary Acquisition and Pronunciation

Shelley Shwu-Ching Young and Yi-Hsuan Wang*
Institute of Learning Sciences, National Tsing Hua University, Taiwan // Department of Information Technology and Management, Tzu Chi College of Technology, Taiwan // scy@mx.nthu.edu.tw // annywang12345@hotmail.com
*Corresponding author

(Submitted April 23, 2013; Revised June 30, 2013; Accepted August 15, 2013)

ABSTRACT

The aim of this study is to make a new attempt to explore the potential of integrating game strategies with automatic speech recognition technologies to provide learners with individual opportunities for English pronunciation learning. The study developed the Game Embedded CALL (GeCALL) system with two activities for on-line speaking practice. For the drill practice, learners get immediate scores with evaluation feedback on their utterances; in the game-based practice, they have to correctly choose and pronounce the English vocabulary from the candidate answers to pass the four-level barriers. To enable the study to confirm the factors that facilitate students’ English learning, a total of 52 learners from Taiwan participated in this experiment, divided into an experimental group (E.G.) and a control group (C.G.). The learners in the E.G learned English with both the drill and game-based practice, while the students in the C.G. learned only with the drill practice. The empirical evaluation reveals that those who learned with both drill and game-based activities performed better in terms of their English pronunciation than the learners who only participated in the drill practice, but the learners in the C.G. performed better in the delayed vocabulary retention test. Besides, the study indicates that the learners, especially the low-achievement ones, showed great involvement and were active in practicing speaking in the game-based scenario. The findings are in accordance with the previous research affirming that gameplay is the driving force that promotes learners’ educative engagement, and that learners of different levels of learning achievement are active in learning together while practicing their oral speaking in the stress-free environment. In sum, the GeCALL system is an educative aid that could reduce learners’ language speaking anxiety and to provide flexible chances for individual students to do self-speaking practice within the given limited teaching time. The findings of this study support the value of applying game strategies with ASR techniques for language learning purposes.

Keywords

CALL system design, Pronunciation learning, Game strategy

Introduction

English has become an international language and many Asian countries have established English learning policies to promote students’ English ability. In Taiwan, the Ministry of Education has set up curriculum guidelines to enhance learners’ oral communication and listening skills at the primary school stage. However, English instructors have indicated that coping with students’ heterogeneity in classes with limited teaching time and arousing learners' speaking motivation are the main barriers when teaching (Su, 2006). In order to deal with teachers’ difficulties, study have indicated that providing students with a one-on-one computer assisted learning system is a helpful method for encouraging learners of various language proficiency levels to learn in an enjoyable way (Dolati & Mikail, 2011). The concept of Computer Assisted Language Learning (CALL) to achieve learner-centered learning can be traced back to the early 1960s, and with the arrival of the World Wide Web, CALL systems were no longer limited to personal editions but enabled asynchronous and synchronous communication between instructors and students. Over the last decades, a variety of educational applications in the CALL research have been implemented, from web-based applications (Golonka et al., 2012; Hsu & Ou-Yang, 2013) to ubiquitous mobile-based learning environments (Wu et al., 2012), and from vocabulary acquisition to English writing skills. Within the topic of adoption of technology in language learning, there is strong support for the claim that integrating automatic speech recognition techniques (ASR) into computer-based learning systems had a measurable impact on pronunciation learning (Chiu, Liou, & Yeh, 2007; Chen, 2011). For example, Neri et al., (2008) used a story-based CALL system, PARLING, to encourage young beginner learners of Italian to memorize the words in a story. The PARLING system uses the word game to encourage children to pronounce the words in the story, and adopted an ASR component to analyze the learners’ utterances. SPELL is another CALL system that created an interactive virtual world learning environment for secondary school students in French to have English conversations with virtual characters (Morton & Jack, 2010).
The above studies demonstrated that integrating an automatic speech recognition (ASR) component in a learning system benefits language learning, and especially facilitates an improvement in pronunciation. However, most of the CALL studies conducted comparison research regarding learners’ performance of practicing speaking with and without an ASR-based environment (Neri et al., 2008; Morton & Jack, 2010), and little research explores the learning factors or interaction design in the ASR-based environment which may enhance the effectiveness of speaking. Many design variables such as self-learning approaches or pedagogy need and are worth investigating (Kim & Gilman, 2008; Neri et al., 2008). Neri et al. (2008), for instance, integrated game-like speaking practices into their system design; however, their results focus on analyzing whether the ASR system can help young learners improve their pronunciation comparable to that through traditional teacher-led scenarios, while the results related to students’ perceptions of participating in the game-like speaking activities in the ASR-based learning environment or the effect of combining game strategy with ASR techniques for promoting the learning of speaking were less investigated.

On the other hand, some studies indicated that adopting educational games into language instruction has a positive impact on promoting students’ learning motivation and performance (Tsai et al., 2012; Chiu, Kao, & Reynolds, 2012). For example, researchers have demonstrated that using scaffolding games helps students acquire targeted knowledge better; and the adoption of game strategies in instruction decreases frustration resulting from excessive numbers of learning retries (Sun, Wang, & Chan, 2011; Tsai et al., 2012). Moreover, the use of web-based games also facilitates language learning such as grammar or writing skills (Tsai et al., 2012). Despite the application of computer-based learning environments and game strategies in learning having separately received considerable attention for promoting learners’ learning ability, research on using game scenarios for assisting English pronunciation training has been less studied.

Research purpose and questions

To make up for this deficiency, this study integrates game strategies into developing an English speaking learning system to assist students’ acquisition of English vocabulary as well as pronunciation. The researchers designed a four-level speaking game to help students learn English progressively through the game process. To enable the study to confirm the factors that facilitate students’ English learning, the developed system supported learners with two learning scenarios: providing learners with speaking drill practice with and without game activities. In the drill practice, learners can practice speaking with the system repeatedly with a score and evaluation feedback; in the game activity, learners have to pronounce the English vocabulary from the candidate answers according to the given picture to pass each barrier and complete the game. This study aimed to explore whether or not applying game-like learning interactions for speaking practices in the ASR-based language learning system would bring various learning effects for learners, and also to observe if students would have diverse learning interactions through participating in different speaking activities. Thus, the following questions are investigated to understand the effectiveness of the developed GeCALL system:
- Which learning scenarios of the system would help the learners achieve better learning performance in acquiring English vocabulary and pronunciation?
- How did the learners perceive practicing pronunciation with the system?
- Were there different learning interactions or behaviors between the learners in the two learning scenarios?

Literature review

Related CALL studies

A variety of educational applications in the CALL research have been conducted, and the applications of CALL bring language learning a wealth of new possibilities (Chen, 2011; Golonka et al., 2012; Hsu & Ou-Yang, 2013). For example, Johnson and Heffernan (2006) conducted a reading CALL project in Japan in which high-beginner to intermediate students were provided with opportunities of exposure to target vocabulary through a series of online story readings. Ma and Kelly (2006) applied a web-based CALL system, WUFUN, to help Chinese students overcome their learning difficulties in reading English vocabulary (Ma & Kelly, 2006). Harbusch et al. (2008) developed a Sentence Fairy CALL system for elementary students to enhance their English writing in virtual writing conferences through the use of the Natural Language Processing technique. Lee et al. (2009) built a web-based
CALL system for adults to learn academic essay writing. The system used Latent Semantic Analysis to give learners immediate writing feedback including content sub-themes and organization. Fidaoui et al. (2010) proposed that fourth graders had more confidence and self-esteem in English writing through using a CALL system in their writing classes. The abovementioned CALL studies focus on enhancing learners’ reading and writing abilities.

More recently, with the mature development of technology, more advanced techniques such as the ASR technology has been applied to CALL applications for immediate learning. For example, Chen and Chang (2011) integrated mobile devices, personal digital assistants, as learning tools to support listening comprehension activities for college students in Taiwan. Research conducted by Tsai and Young (2010) and Chwo (2012) all integrated the speech recognition technology or text-to-speech function to enhance students’ multi-language, speaking and listening ability. The above studies all focused on adopting CALL applications to facilitate students’ listening and speaking ability, and all revealed positive support for the claims that integrating learning technology into language learning systems offers learners extended opportunities and benefits students’ language performance. We can conclude from the above studies that CALL programs offer learners self-directed learning environments and allow students to train at a self-paced speed.

**The potential of game-based learning environments**

The research on game-based computer assisted learning has increased rapidly since 2006 (Hwang & Wu, 2012) and has become a popular trend in the field of learning technology for promoting learners’ educative engagement and motivation (Hong et al., 2009; Tsai, Yu, & Hsiao, 2012). Prensky (2001) mentions that teaching with game strategies could address some of the pitfalls found in traditional education, and most researchers believe that educational games could be learning aids to be used as reinforcement to support traditional learning (Tsai, Yu, & Hsiao, 2012) because such games can lower anxiety and make learning acquisition more likely. Furthermore, games have the power to get learners to learn enthusiastically, and the repetition of gameplay is the driving force that motivates learners to search for target knowledge through the chance of learning by playing (Coyne, 2003). Although many studies claim that educational games are beneficial for learning, there are still some researchers who deny the effectiveness of educational games because of the problems of students’ distraction by gameplay and not paying attention to the targeted knowledge (Ke, 2008; Papastergiou, 2009). Therefore, finding ways to help students acquire learning content efficiently through playing educational games is a critical issue.

**The principle of game-based activities design**

To address the concern of how to appropriately design game-based learning activities, researchers of Behaviorism consider learning to be produced by stimulation and reinforcement, and the game-based learning environment is useful in autonomous learning. That is to say, learners in games have to know their learning goals and what actions help them to achieve those goals (Wu et al., 2012). During the processes of gameplay, frustration control should be designed to prevent learners from becoming stuck in the games (Sun, Wang, & Chan, 2011). On the other hand, the Constructivist perspective proposes that providing scaffolding processes in game-based design may be a possible way of directing learners to pay attention to learning contents during gameplay (Tsai et al., 2012). The scaffolding strategy could also be used for assisting learners to accomplish learning tasks at their own pace (Wu et al., 2012). Besides, the design of the game has to meet the target learners’ prior knowledge level. If the provided games can achieve the balance between learners’ skills and challenge, the possibility of experiencing learning through gaming is higher (Kiili, 2005). Meanwhile, multimedia learning materials in educational games are another factor impacting the effectiveness of game-based learning. There may be a risk of overloading learners’ working memory due to inappropriate ways of presenting too many multimedia elements (Kiili, 2004). To appropriately balance the attractive factors of games and educational content in order to optimize the best learning effectiveness while designing educational games is a challenging task.

**Game-based language learning study**

Several studies have applied game strategies to language acquisition for enhancing English grammar or vocabulary acquisition (Chiu, Kao, & Reynolds, 2012; Tsai et al., 2012). For many language learners, the traditional way of
learning a new language is to "look and remember" repeatedly without internalization, and sometimes students even write down the translation of new words but cannot pronounce them. Research has shown that the cramming method of language learning is not an effective way and that after a short period of time, many learners may slowly forget the learned vocabulary (Decarrico, 2001). However, language learning could be facilitated by educational games and much research has indicated that game-based teaching strategies have positive effects on language acquisition (Barendregt & Bekker, 2011). For example, Decarrico (2001) applied a web-based game to facilitate learners' English vocabulary acquisition and found that the learners had the opportunities to be in control of the lessons during games and were able to move at their own pace to complete the learning. Besides, both advanced students and less advanced ones worked together to finish the learning task. On the contrary, learners who learned without using games as aids were passive and depended on teachers totally for the input of new materials without involvement for self-development. Meanwhile, another study pointed out that the use of games had positive impacts on supporting teachers to engage learners of various language proficiency levels in a more enjoyable way of learning English compared to traditional teaching methods such as mechanical drilling (Chiu, Kao, & Reynolds, 2012).

In summary, the advantages of adopting game strategies for language learning shed light on the potential for 1) learning and retaining knowledge more easily, 2) keeping learners’ interested and getting involved actively, and 3) repeated practice in a game-based manner. However, most of the reviewed studies adopt game strategies for language reading or writing skills; using game-based scenarios for speaking training has been less studied. Besides, in the application of CALL, there is a lack of a systematic approach to investigating the key factors in ASR-based CALL systems that may enhance the effectiveness of speaking learning, and further research is suggested to explore the effectiveness of using games for teaching speaking (Dolati & Mikail, 2011). Hence, this study aims to develop an interactive speaking system from the perspective of learning-theory foundations invoked to underpin educational computer game design. Meanwhile, the study examines whether the developed system adopting the game-based activities could facilitate students’ vocabulary acquisition and pronunciation and also to identify the potentiality of using games for pronunciation learning.

**The Game embedded CALL (GeCALL) system**

The goal of developing the GeCALL system is to provide learners with self-learning opportunities to achieve vocabulary acquisition and pronunciation ability. The researchers adopted a self-developed ASR application for evaluating speakers’ English pronunciation based on the factors of tone, speed, volume and timbre (Chen, Lo, & Jang, 2004).

**Learning materials and activities design**

The learning materials in the GeCALL system were selected from the required elementary school level vocabulary from school textbooks to balance the learners’ language skills and game challenge so as to promote the possibility of learning through gaming (Kiili, 2005). The elementary English teachers further differentiated the vocabulary into four categories based on the frequency of their appearance in the textbooks, with Category One containing the most common vocabulary while Category Four includes the least common vocabulary. Then the researchers designed the learning activities according to the vocabulary in the four categories. Two types of activities were designed in the system, drill practice and game-based practice. The drill practice includes four parts with each containing the vocabulary from one category. For example, Part 1 contains the vocabulary from Category One and Part 4 contains the vocabulary from Category Four (Figure 1-1).

The researchers designed the game-based practice as a four-level challenging barrier game for scaffolding learners while acquiring the vocabulary and accomplishing learning tasks at their own pace. The scaffolding strategy is adopted to design the question of each level, and the vocabulary from each category is selected according to different proportion as learning questions to assist the students in learning the vocabulary of each category gradually (Wu et al., 2012). For example, Level 1 contains the vocabulary from Category One; in Level 2, 40% of the vocabulary is from Category One and 60% is from Category Two; in Level 3, 20% is from Category One, 20% is from Category Two and 60% is from Category Three. If the students get stuck in Level 2, then they have to practice the vocabulary of Category One and Two again to pass the Level 2 barrier in the game. The structure of the game-based practice is presented as Figure 1-2.
Figure 1-1. The structure of the drill practice: 4 parts

Figure 1-2. The structure of the game-based practice: 4-level barrier game

Rules of the two activities: drill practice and game-based practice

In the drill practice, students can listen to sounds produced by a native speaker and then record themselves repeating the sounds. Their pronunciation accuracy is displayed with a score and evaluation feedback on the screen, and then they can have repeated practice with the system (Figure 2).
In the game-based practice, the learners have to select and pronounce the English vocabulary from three candidate answers on the screen according to the given picture (Figure 4-a). If they choose the correct answer and pronounce it correctly, the system will give them positive feedback and then they can go on to the next question (Figure 4-b). When learners complete one level of barrier game, they will get one color of virtual medal as reward. After they collect the four colors of medals, they complete the barrier game. Meanwhile, the system also offers a frustration control tool that gives the learners textual hints to prevent them from becoming stuck (Sun, Wang, & Chan, 2011) (Figure 4-c). The flow diagram of the game-based practice is shown as Figure 3. Through the gaming process, the learners are asked to pronounce the vocabulary, and the motivation to practice speaking comes from the desire to pass the different-level barriers in the game, thus empowering and enriching their self-speaking opportunities.

![Figure 3. Flow diagram](image)

![Figure 4. Sample screens of the game-based practice](image)

**Methodology**

The study used comparative test data to report on the performance of learning English in the GeCALL system with two learning scenarios, and both qualitative and quantitative approaches were employed. The learners in the experimental group (E.G.) learned English with both the vocabulary drill practice and the game-based practice, while the students in the control group (C.G.) learned only with the vocabulary drill practice. In order to explore how the
GeCALL system facilitated the learners’ English learning and what conditions contributed to their English vocabulary acquisition and pronunciation learning, the researchers conducted a paper-based test and speaking test before and after the experiment as the learning pre-test and post-test. In the speaking test, the learners had to pronounce five words learned in the GeCALL system, while the paper-based test consisted of ten multiple-choice questions that required the learners to select the corresponding English vocabulary according to Chinese items. The total scores of the paper-based test and speaking test were 100 each. Moreover, the questionnaire developed by the researchers of the present study was administrated to the learners. The items included two parts, learning feedback and system operational feedback. The perceived learning scales were modified based on the Motivation-and Attitudes-toward-Learning-English-Scale-for-Children (MALESC) questionnaire used to measure the English motivation and attitudes of elementary students (Carreira, 2006). The reliability coefficient of the questionnaires was 0.92 as measured by Cronbach’s α. To increase the validity, the wording of the survey items were reviewed by one learning technology expert and one English teacher. Meanwhile, in order to understand the participants’ experiences of participating in the speaking activities, a total of 15 students were chosen for interviews. The semi-structured interview included the questions: (1) How did you feel when participating in the speaking activities in the system? and (2) Did you enjoy the speaking activities? Finally, a delayed vocabulary retention test was administered to understand whether the learners in the two groups had remembered the learned words because the purpose of the delayed test was to measure learning outcomes not only before and immediately after the experiment, but also several days later (Draper et al., 1996). A one-week delay period was used in this study (Iwashita, 1999; Mohammad et al., 2011).

Participants

A total of 52 fourth graders (aged from 7 to 9) from an elementary school in Taiwan participated in this study and were divided into two groups. The E.G. consisted of 27 learners practicing English with the system using both drill practice and game-based practice, while the C.G. consisted of 25 students only given drill practice. Learners in both groups were further divided into three subgroups based on their English grades during the previous semester for further qualitative analysis. The high-achievement learners were those students whose grades were in the top one third of the class. The grades of the low-achievement learners were in the bottom third of the class, and the rest of the students were categorized as medium-achievement learners.

Data collection

The duration of the data collection was 8 weeks. In the first class, the English instructor gave the students a detailed demonstration of the system operation and its functions. Then the learners had to take the pre-test. During the experiment period, the learners in the E.C. and the C.G. learned English with the system in the drill practice, while only the learners in the E.G. participated in the game-based speaking practice. Besides, observation forms detailing the students’ learning performance and processes in the classes were recorded by the trained observers. The observational scales were documented in the observation forms. The content of the evaluation criteria included: Did the students show motivation and involvement during the practice session with the system? How did the learners interact with the system and their classmates while practicing speaking? After 8 weeks of practice, the learners were asked to take the post-test. Moreover, the delayed vocabulary retention test was administered to the students one week later.

Data analysis and results

Data from the pre-test, post-test, delayed test, questionnaires and interviews were analyzed according to the research questions. Descriptive statistics were calculated to describe the means and standard deviations, and analysis of covariance (ANCOVA) was adopted to compare the final learning results (post-tests) of the two groups with pre-test scores as covariates to eliminate the effect of any existing pretest difference on the results. The homogeneity of variances was assessed by the Levene’s test, and the results confirmed that the data met the equality of variance assumption.

245
Learning performance of the two groups

The pre-test, post-test and delayed test data were analyzed to answer the first research question. Table 1-a shows the descriptive statistics, including the means and standard deviations of the pre-test and post-test scores of the two groups. Both groups showed improvement on the post-test, but, noticeably, the ANCOVA results show a significant difference only in the pronunciation post-test scores of the C.G. and E.G. (Table 1-b; p = 0.03*). However, it is noticed that the learners in the E.G. lagged behind the learners in the C.G. a little in the delayed vocabulary retention test. This indicates that the learners with game-based practice achieved better speaking improvement than those with only drill practice during the experiment, but the learners with only drill practice memorized the vocabulary better than the learners with game-based practice from a long-term perspective.

<table>
<thead>
<tr>
<th>Table 1-a. Descriptive statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.G.</td>
</tr>
<tr>
<td>Pre-test</td>
</tr>
<tr>
<td>M   S.D.</td>
</tr>
<tr>
<td>Pronunciation Test</td>
</tr>
<tr>
<td>Paper-based Test</td>
</tr>
<tr>
<td>C.G.</td>
</tr>
<tr>
<td>Pre-test</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1-b. ANCOVA table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest_ Pronunciation Test</td>
</tr>
<tr>
<td>SS</td>
</tr>
<tr>
<td>618.77</td>
</tr>
<tr>
<td>Between(Group)</td>
</tr>
<tr>
<td>Within(Error)</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Corrected total</td>
</tr>
</tbody>
</table>

Evaluation of learners’ learning reflection

The questionnaires regarding the learners’ reflections on using the GeCALL system were analyzed in order to answer the second research question. In all, the general evaluation by the learners was positive and they reflected that the GeCALL system helped them to memorize and pronounce the English words (Table 2, Q4 & Q5). Besides, the participants claimed that using the system promoted their opportunities of English speaking and most of them expressed the expectation to practice speaking with the aid of the GeCALL system for further learning (Table 2, Q1 & Q3). Furthermore, learners from the E.G. gave affirmative feedback on participating in the speaking game. They acquired the English vocabulary by completing the different levels of the speaking activity and showed strong motivation to practice speaking in order to experience achievement in the speaking barrier game (Table 3, Q1–Q4.).

<table>
<thead>
<tr>
<th>Table 2. Learners’ feedback on the system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Using the system enhances the opportunities of English speaking.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>I like to practice English speaking and I am not afraid of making mistakes.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>I would like to use the system for further learning.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>I can pronounce the words clearly with the aid of the system.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>I can memorize the words with the aid of the system.</td>
</tr>
</tbody>
</table>

246
as observed that by using the speaking game, the students had
behavior
Learning
interactions
Learning


to complete the challenging barrier game.

Moving around the classroom, the environment was controllable because they
more interaction with each other and
in front of students with better performance. But it
mentioning that originally some students were shy and afraid of being mocked by their classmates when they spoke
through the
activities. The low
levels to complete the challenging barrier game.

Meanwhile, qualitative feedback from the interviews was also collected. The learners pointed out that speaking
English with the computers was easier than practicing with a real person. In addition, most of the students agreed that
the levels of the game were moderate, while a few low-achievement learners revealed that the game was challenging
but they were not frightened to keep practicing speaking because it was just a game. Their comments included: “Yes,
I did play the game and I think it was a little hard. But I am not afraid of it because I was just playing a game” (EL9)
and “It was hard and I had to try lots of times to complete the game. I was nervous but I think it was still quite interesting” (EL10). In contrast, the researchers also found that very few (only two) of the high achievement learners showed any anxiety about playing the game because they were afraid of making mistakes in the speaking practice which might affect their score. They commented that, “The game was not hard and I had learned some of the words before. But I had to be careful so that I wouldn’t make any mistakes in the game” (EH22). For a better understanding, a summary of the observational results of the two groups is provided in Table 4.

### Table 3. The E.G. learners’ feedback on the system

<table>
<thead>
<tr>
<th>Items</th>
<th>Avg</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participating in the speaking game enhanced my speaking motivation.</td>
<td>4.47</td>
<td>0.83</td>
</tr>
<tr>
<td>Learning speaking through speaking games is quite interesting</td>
<td>4.55</td>
<td>0.76</td>
</tr>
<tr>
<td>I learned lots of new vocabulary through playing the speaking barrier game.</td>
<td>4.82</td>
<td>0.57</td>
</tr>
<tr>
<td>I would like to practice speaking repeatedly for experiencing achievement in the speaking activities.</td>
<td>4.52</td>
<td>0.73</td>
</tr>
</tbody>
</table>

### Evaluation of learners’ learning interaction and behavior when using the system

The observational data and interviews regarding the learners’ learning processes and interaction while using the GeCALL system for speaking practice were analyzed in order to answer the third research question. The classroom observation indicated that the learners in the two groups all engaged in participating in on-line speaking learning. The learners in the C.G. had little learning interaction except comparing speaking scores with each during the drill practice, while, some of the high-achievement learners in the E.G. shared their experience of completing the game-based activity with the low-achievement ones, and some of the high-achievement learners in E.G. even accompanied the others especially the low-achievement learners to practice speaking in the game-based activity.

Noticably, learners of various learning achievement levels in the E.G. had different learning behaviors during the activities. The low-achievement learners in the E.G. tended to firstly participate in the game-based speaking activity and went to the drill practice only if they got stuck in the speaking game (Figure 6a); however, the higher learning achievement learners in the E.G. usually practiced the vocabulary and read the results of the speaking evaluation through the drill practice first before challenging themselves in the game-based activity (Figure 6b). It is worth mentioning that originally some students were shy and afraid of being mocked by their classmates when they spoke in front of students with better performance. But it was observed that by using the speaking game, the students had more interaction with each other and practiced the target language together comfortably. Although the students were moving around the classroom, the environment was controllable because they needed to pay attention to the learning contents to complete the challenging barrier game.

### Table 4. Summary of the observational results

<table>
<thead>
<tr>
<th>Learning interactions</th>
<th>C.G.: Drill practice</th>
<th>E.G.: Drill practice and Game-based practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learners only compared speaking scores with each other</td>
<td>• Learners had little learning interaction except for comparing the speaking scores</td>
<td>• Learners of different learning achievement had good learning interaction:</td>
</tr>
<tr>
<td>Learners had little learning interaction except for comparing the speaking scores</td>
<td>• Learners of different learning achievement had good learning interaction:</td>
<td>✓ High-achievement learners shared the learning experience with the low-achievement learners</td>
</tr>
<tr>
<td></td>
<td></td>
<td>✓ High-achievement learners accompanied the others to practice speaking during the game-based activity</td>
</tr>
<tr>
<td>Learning behavior</td>
<td>Learners were moving around the classroom</td>
<td>Learners of different learning achievement had different learning paths:</td>
</tr>
<tr>
<td>Learners were pleased with practicing speaking with the system</td>
<td>• Learners of different learning achievement had different learning paths:</td>
<td>✓ Low-achievement learners firstly participated in the game-based speaking activity and then</td>
</tr>
<tr>
<td>Learners all sat in their own seats and were well</td>
<td>• Learners were moving around the classroom</td>
<td></td>
</tr>
</tbody>
</table>

247
behaved

- Learners were engaged in the activities

• went to the drill practice
  Higher-achievement learners practiced the drill activities first before challenging themselves in the game-based activity
  - A few low-achievement learners revealed that the game was challenging
  - A few high-achievement learners showed anxiety about playing the games

Learning feedback

- I am more willing to practice speaking with the activities. (CH1)
- Practicing with the computer is better. It is difficult to speak with classmates. (CH9)
- I see the classmates get good scores and I like to practice more. (CL12)
- The game is easy. I am not afraid. (EH1)
- Yes, I did play the game and I think it was a little hard. But I am not afraid of it because I was just playing a game. (EL9)
- Practicing speaking with games makes me want to learn English. (EL12)
- It was hard. I had to try lots of times to complete the game. I was nervous but I think it was still quite interesting. (EL10)
- The game was not hard and I had learned some of the words before. But I had to be careful so that I wouldn’t make any mistakes in the game. (EH22)

Discussion

Different impacts of the game-based practice and drill practice on English learning

The evaluation results from the pre-test and post-test indicated that the use of the GeCALL system enhanced the learners’ vocabulary acquisition as well as their pronunciation ability, especially for the learners in the E.G. who achieved significant improvement in their speaking performance from the results of ANCOVA analysis. Those who learned with the game-based activity compared their speaking performance with each other during the speaking activity and “the desire to complete the barrier game” became a trigger that motivated them to speak out. The process of completing the game helped the learners acquire the target vocabulary level-by-level, and it was observed that the learners went back to the drill practice to practice pronunciation again if they got stuck in one particular level. This finding echoes the research affirming that games have the power to get learners to learn enthusiastically, and the repeated gameplay was the driving force that motivated the learners to internalize the target knowledge gradually by playing (Coyne, 2003). On the other hand, the learners in the C.G. who learned only with the drill practice performed better in the delayed vocabulary retention test one week later. The reason can be explained and inferred from the qualitative observation that the learners in the E.G. put more focus on speaking the words rather than memorizing them during the barrier game; however, those students who learned with drill practice read the English vocabulary and the content description carefully, but sometimes they did not or forgot to pronounce the words. Accordingly, those who learned with the game-based activity performed better in terms of their pronunciation performance, but their vocabulary retention regressed compared with that of the learners who just used the drill practice. We conclude from the results that the game-based speaking activity enhanced the students’ speaking motivation as well as their speaking performance; yet, the adoption of traditional teaching instruction, the drill-and-practice method, may facilitate learners’ vocabulary retention more than that of the game-based activity.

Various achievements resulting from the game-based speaking scenario

According to the qualitative data, learners in the English class were all engaged in speaking practice with the GeCALL system. A previous study had indicated that sometimes noisy and disorganized learning happens while adopting game activities for teaching (Evans, 1979). However, even though we found the same situation, the atmosphere in the class was full of joy while the students were moving around the classroom during the game-based activity. One observation worth noting is the different learning behaviors of the learners in the C.G. and E.G. during the practice sessions. In the C.G., the learners were willing to practice speaking with the system and most of them
worked on their own in the drill practice. They liked to compare their speaking scores with each other, but little learning interaction happened while they were practicing speaking.

On the contrary, learners in the E.G. who participated in the game-based speaking activities tended to help each other to complete the game. The desire to try and the willingness to make mistakes among the learners were the factors which led to the students in the E.G. having learning interaction with classmates without stress. This observation was particularly obvious for the low-achievement learners who showed great involvement and were active in practicing speaking. Similar results were found in a previous study (Decarrico, 2001) in which the research mentioned that adopting game-based activities motivated advanced and less advanced students to work together to finish the learning tasks. The flexible and comfortable learning atmosphere during games made the students feel more relaxed and promoted their educative engagement so as to facilitate language acquisition (Barendregt & Bekker, 2011; Dolati & Mikaili, 2011; Chiu, Kao, & Reynolds, 2012). In addition, another noticeable observation is that (very) few advanced learners were afraid of making mistakes in the games due to their high self-expectations, and they showed little pressure while participating in the activities.

Conclusions

This study is a new attempt to explore the potential of integrating game strategies with state-of-the-art ASR technologies to provide learners with individual opportunities of immediate English pronunciation learning. The research concludes that, firstly, the learners’ pronunciation performance improved significantly through practicing with the game process in comparison with the students who used the drill practice only, while the learners with only drill practice performed better in a delayed vocabulary retention test. Secondly, the low-achievement learners tended to participate in the barrier game first, only trying the drill practice if they got stuck pronouncing words. However, the high-achievement learners progressed the other way. This implies that the GeCALL system with both drill and game-based practices fulfilled learners’ diverse needs. Thirdly, the results of the study indicate that the learners were not afraid of speaking English in front of peers even though they might make mistakes; meanwhile, the high-achievement learners were willing to share their speaking experiences with the low-achievement ones. The study affirms that the learners of various achievement levels had friendly learning interaction with each other while practicing speaking with the GeCALL system, and the lively interaction and great involvement during the game scenarios helped the learners overcome their anxiety about speaking and hence has positive potential for speaking acquisition. Despite the drill practice in the GeCALL system seeming to have better affirmation for vocabulary retention that benefits learners in passing traditional paper-based exams which are a feature of Taiwan’s education system, the game-based activity in the GeCALL system obviously provides learners with a stress-free environment in which they are stimulated to actively practice their oral speaking, an important communicative skill required in the authentic world, comfortably.

Currently, most learners in Taiwan attend English classes in which they have to share their teacher’s attention with each other, and this may reduce the amount of time that each student spends producing language output. However, the adoption of the GeCALL system provides students with an immediate one-on-one on-line speaking scenario, which could be an educative aid that school teachers could use to provide flexible chances for individual learners to do self-speaking practice within the given limited teaching time. Meanwhile, it is suggested that English instructors apply the GeCALL system in their formal English courses or for after-school learning, for example, once a week, as a remedial strategy for speaking practice. The learning contents could be modified according to the curriculum, and English instructors could assign the learners with diverse levels of proficiency to different kinds of speaking activities. For the high-achievement learners, they learn through advanced drill practices to further memorize the spelling of the vocabulary and internalize the knowledge; for the low-achievement students, the game-based activities can arouse their speaking motivation and, as long as the learners have interest in learning English, they could start engaging in the language learning scenario. The role of the English teachers could be as facilitators who could move around the classroom to monitor the students’ learning processes and to provide help when needed. It could reduce teachers’ teaching burden in coping with students’ heterogeneity in one class; meanwhile, it can increase learners’ speaking opportunities compared with learning in a traditional practice environment. Besides, for informal learning, teachers could hold on-line speaking competitions in the GeCALL system to encourage learners to keep practicing speaking after school. The game-like and relaxed learning scenarios of GeCALL may be a trigger that fosters learners, both at the advanced and low-achievement levels, to join together practicing speaking in a pleasurable learning environment. The findings of this study support the value of applying game research with ASR techniques for language learning. Other related studies on integrating game speaking training with ASR applications will be further explored in the near future.
References


Chwo, S.-M.-G. (2012). Enhancing listening proficiency via multi-modality technology - An initial finding from technology university non-English major EFL learners in central Taiwan. In J. Colpaert et al. (Eds.), *Proceedings of the fifteenth international CALL conference* (pp. 212-216). Taichung, Taiwan: Providence University.


An Integration Architecture of Virtual Campuses with External e-Learning Tools

Antonio Navarro1*, Juan Cigarrán2, Francisco Huertas3, Miguel Rodríguez-Artacho2 and Alberto Cogolludo2

1Departamento de Ingeniería del Software e Inteligencia Artificial, Facultad de Informática Universidad Complutense de Madrid, C/ Profesor José García Santesmases s/n, 28040 Madrid, Spain // 2Departamento de Lenguajes y Sistemas Informáticos, Facultad de Informática Universidad Nacional de Educación a Distancia, C/ Juan del Rosal 16, 28040 Madrid, Spain // 3Center for Open Middleware, Universidad Politécnica de Madrid, CTB building, floor1, Campus de Montegancedo, 28223 Pozuelo de Alarcón, Spain // anavarro@fii.ucm.es // juanci@lsi.uned.es // francisco.huertas@centeropenmiddleware.com // miguel@lsi.uned.es // acogolludo@lsi.uned.es

*Corresponding author

(Submitted July 29, 2013; Revised October 18, 2013; Accepted December 20, 2013)

ABSTRACT

Technology enhanced learning relies on a variety of software architectures and platforms to provide different kinds of management service and enhanced instructional interaction. As e-learning support has become more complex, there is a need for virtual campuses that combine learning management systems with the services demanded by educational institutions and users. However, nowadays, the functions of virtual campuses need to be enhanced with external e-learning applications. This paper deals with the issue of integrating these applications in virtual campuses using a software architecture based on the merging of mashup and multitier patterns. Our solution is based on an asymmetric architecture that provides seamless interconnection of the external tools with virtual campuses. We illustrate it with a real example of the integration of a research prototype that searches for and retrieves learning objects with a virtual campus designed independently of the underlying e-learning platform.

Keywords

Software architecture, Software patterns, LMS, Moodle, Sakai

Introduction

Looking back over the past decade we have witnessed the rapid evolution of technology-enhanced learning, leading to the appearance of a heterogeneous set of e-learning tools, including Learning Management Systems (LMSs), virtual campuses, Managed Learning Environments (MLE), virtual object editors, Massive Open Online Courses (MOOCs) and serious games. This is only a small subset of the e-learning tools that can be found in educational institutions.

However, a constant in educational institutions is the need to manage courses, students and teachers, a need usually addressed by virtual campuses (CCP, 2012). According to Yábar et al., (2007), virtual campuses play several roles: they provide support for face-to-face teaching, encourage teaching innovation and the authoring of learning material, foster communication between the various different users, facilitate learner monitoring, self-learning and self-assessment, and provide blended teaching experiences with different degrees of virtuality.

These campuses are software applications that can be conceived from different points of view. According to Van Dusen (1997), virtual campuses are a metaphor for the electronic teaching, learning and research environment created by the convergence of several relatively new technologies including, the internet, World Wide Web, computer-mediated communication, video conferencing, multimedia, groupware, video-on-demand, desktop publishing, intelligent tutoring systems, and virtual reality. In more recent studies (Allison & DeBlois, 2008; Epper & Garn, 2004; Green, 2005; PLS, 2004) virtual campuses are understood, in a broader sense, to be the integration of Information and Communication Technologies (ICT) in terms of both educational and organizational university settings. This matches our view of virtual campuses as modular software systems used by higher education institutions to give support to their teaching and learning activities (Navarro et al., 2012a).
As a first step, virtual campuses were built using a simple LMS for dealing with core e-learning activities. Educational institutions quickly realized that an LMS was not enough to deal with the management activities (i.e., registration in the LMS course of a group of students enrolled in an in-person course) involved in a virtual campus. So, virtual campuses evolved to become complex applications. The virtual campus deployed at the Universidad Complutense de Madrid (UCM) is a complex software system made up of four different applications (Navarro et al., 2012a): (i) two LMSs (Moodle 1.9 & Sakai 2.4), responsible for core e-learning facilities (e.g., content publication); (ii) a J2EE (Weaver, 2004) data load application, responsible for transferring course and user data from the university management system to the virtual campus administration database; (iii) a J2EE administration application, responsible for administrative tasks (e.g., changing a password); and (iv) an integration page that provides a unified view of the courses virtualized in different LMSs.

However, once virtual campuses were running, users demanded three different types of feature (Navarro et al., 2010): (i) the use of different LMSs in the same virtual campus; (ii) independence of the underlying LMS; and (iii) the inclusion of additional e-learning features not contemplated in LMSs. In other papers we have focused on solving issues (i) and (ii). This paper focuses on the third issue.

We have found in the literature two main approaches for dealing with the integration of software applications: (i) those based on service integration using a multitier architecture (Alur et al., 2003; Fowler, 2002) such as Event-Driven Architecture, EDA (Etzion & Niblett, 2010), Service Oriented Architecture, SOA (Erl, 2005) and classic integration patterns (Juric et al., 2001); and (ii) those based on mashup integration (Hanson, 2009; Ogrinz, 2009).

The first type focuses on service integration and does not usually consider cases where one application needs to use the services of another through its user interface. These integration patterns thus focus on isolating service invocation, monitoring transactions and managing interchanged data, omitting user interface integration. The second type, based on mashups, basically pick up information gathered from different service providers and display it in a user-friendly way. Interaction between the different service providers can exist, but this is limited. In addition, the portal (i.e., the result of the mashup) is a new application that probably did not exist previously (Adams & Gerken, 2008; Ogrinz, 2009).

When integrating virtual campuses with external e-learning tools we have considered both types of integration: service-based and user-interface based. Thus, the external tools have to invoke virtual campus services, for example, to upload content for a specific course. In addition, because external e-learning tools usually have a complex user interface, the virtual campus has to render the external tool user interface in the context of its own user interface, enabling the integrated invocation of the external tool services. For example, virtual campus users may need to visualize specific content, previously generated and uploaded in the virtual campus, using the external tool user interface.

The integration of external e-learning tools in the context of virtual campuses has the following integration requirements: (i) two standalone applications must be integrated: the virtual campus and the external e-learning tool; (ii) the virtual campus contains key information about users and courses; (iii) the external tool is able to find or create contents that have to be deployed (or uploaded) to specific courses for particular users of the virtual campus; (iv) because of the preceding requirement, the external tool has to be invoked for a user who is logged on in the virtual campus, thus providing a context for the content to be uploaded; and (v) the external tool has a complex user interface that makes its integration in the virtual campus user interface necessary.

Taking these requirements into account, we can see that these applications play two different roles, which can be generalized for other contexts: on the one hand, the virtual campus is the host application, whose user interface has to embed the external tool user interface, providing a context for the uploading of content provided by the tool. On the other hand, the external tool is the guest application, which has to be able to respond to virtual campus requests via its integrated user interface (the terms “host” and “guest” are ad-hoc terms defined in this paper).

This raises an important research question: how to provide integration architecture for independent software applications that share services through an embedded user interface, based on the use of design patterns. The use of design patterns guarantees maintainability and other positive characteristics (enumerated in the related work section) in software architectures. The research question arose in the context of the AACV Project, where we had to integrate a virtual campus that we have created with an external e-learning tool.
Reviewing the literature about software integration, we have found that solutions for service integration do not pay attention to user interface integration and those focusing on mashup definitions do not pay attention to complex service interactions between both types of application. This paper provides a generic software architecture for the integration of virtual campuses with external e-learning tools, combining mashup and multitier architecture patterns. The proposed architecture is contextualized in terms of e-learning applications, but could be reused in other applications with similar requirements. Section 2 presents related work. Section 3 describes the proposed integration architecture. Section 4 provides an example of this architecture, describing the integration between a virtual campus and an external e-learning tool, going more deeply into details than the proof of concept architecture described in (Navarro et al., 2012b). Section 5 provides educators’ with a simple description of the overall approach described in this paper. Finally, Section 6 describes conclusions and future work.

Related work

Currently almost any entity of medium size makes extensive use of ICT. The bigger the entity, the more likely it is to need to use different software systems to manage different aspects of its business model: payroll, sales management, supplier management, business intelligence, etc. Although some years ago the main problem was the development of these systems, in recent years the main problem has become the integration of the software systems deployed in an organization (Alur et al., 2003; Fowler, 2002). Starting with classic integration patterns (Juric et al., 2001) several software architectures have been proposed to ease the task of integrating different systems that can run as stand-alone. Two of these architectures are Service-Oriented Architecture (SOA) (Erl, 2005), and Event-Driven Architecture (EDA) (Etzion & Niblett, 2010). SOA architecture is based on the concept of web service: A software function synchronously invoked over a network, irrespective of the platform on which this service is implemented. EDA architecture is based on the concept of an event: a stimulus able to initiate an asynchronous service invocation. Although SOA and EDA architectures have utility per se, they show their full potential in the context of multitier architecture (Alur et al., 2003). This architecture divides the software system into five tiers:

- **User tier**: The system actors that interact with the software application.
- **Presentation tier**: Responsible for displaying the application’s functions.
- **Business tier**: Responsible for the implementation of the business rules.
- **Integration tier**: Responsible for interaction with the resource tier.
- **Resource tier**: The resources used for the software applications, such as databases or other systems.

Multitier architecture is a de facto standard in the development of enterprise applications (Fowler 2002), because it can easily accommodate SOA or EDA solutions using web service broker and service activator patterns (Alur et al., 2003).

Mashup literature provides another source of integration patterns (Hanson, 2009; Ogrinz, 2009). A mashup is a new application genus that describes the combination of two or more sources into an integrated site, the output usually being created by one of the sites participating in the mashup (Ogrinz, 2009). Mashups have evolved into complex mixtures of applications, and mashup patterns have been defined to deal with this complexity. These patterns can be classified as follows (Ogrinz, 2009):

- **Harvest patterns**: A general class of solutions that lend themselves to obtaining data from sources outside the reach of traditional tools.
- **Enhance category**: This demonstrates how systems can be extended and improved.
- **Assemble patterns**: These show how fresh solutions can be minted by combining data and presentation from multiple sources.
- **Manage patterns**: These use mashups as a vehicle for helping an organization leverage its existing assets more effectively.
- **Testing patterns**: These test mashup applications.

Both approaches are used in software design due to the advantages of the architectures that include them (Alur et al., 2003; Fowler, 2002; Erl, 2005; Gamma et al., 2005; Hanson, 2009), such as better integration and reusability, encapsulation, distribution, partitioning, scalability, performance, reliability, manageability, consistency and flexibility, packaging, and configurability. Other advantages are: support for multiple clients, independent development and faster development times.
Although traditional web applications and mashups are web applications that are built in the server and use the browser for displaying data, mashups have additional characteristics (Adams & Gerken, 2008): they aggregate disparate data, typically leverage public data, provide a common user interface for data, produce something new, produce something interesting, and add value by associating data.

Clearly, the problem of integrating a virtual campus and an external tool does not fall in the mashup category, because a virtual campus does not leverage public data and both the virtual campus and the external tool existed before they were integrated. However, their integration can take advantage of some user interface integration concepts belonging to the mashup pattern domain.

The architecture described in this paper implements mashup patterns in terms of multitier architecture patterns, taking advantages of both types of patterns to promote maintainability and other positive characteristics provided by pattern-based architectures. We have not found solutions of this type in the reviewed literature.

Finally, there are some software architectures for e-learning systems that we should mention: OKI Open Service Interface Definitions (OKI OSID) (OKI, 2011), IMS Learning Information Services (IMS LIS) (IMS, 2010a), IMS Basic Learning Tools Interoperability (IMS BLTI) (IMS 2010b) and IMS General Web Services (IMS GWS) (IMS, 2006). These standards are related to the integration of LMSs with other tools and are intended for the interaction between the virtual campus and the underlying LMS. However, in new generation virtual campuses, LMS core e-learning services tend to be wrapped by a virtual campus e-learning layer (Huertas & Navarro, 2013). Thus, the external tool usually interacts with this wrapping, which isolates LMSs from external tools. The use of these software architectures is not, therefore, of value for the type of integration described in this paper.

Integration architecture

This section describes the integration architecture proposed to provide a pattern-based solution to the problem of integrating host and guest applications. Although the detailed requirements have been listed in the introduction, the basic requirement is that the host application’s user interface has to embed the guest application’s user interface, providing a context for the uploading of content provided by the guest application. Therefore, the guest application has to be able to respond to the host via its integrated user interface.

In order to provide a homogeneous characterization of both applications, and taking into account the latest development architectures, let us suppose that both applications use a multitier architecture. Taking this restriction into account, Figure 1 depicts the architecture of the unrelated virtual campus and external tool using a UML 2 component diagram (Arlow & Neustadt, 2005).

The integration layer of the Virtual Campus architecture is made up of an LMS and a virtual campus database exposed via their corresponding APIs. On this layer a set of management services and core e-learning services are built. The management services make it possible, for example, for a group of students enrolled in a class using the university online system to be automatically registered in an LMS course. The e-learning core functions are intended
to wrap the LMS-dependent functions into an LMS-independent system, thus isolating the virtual campus from the underlying LMS (Huertas & Navarro, 2013). If this component is not present in the architecture, it can be replaced by a web service broker (Alur et al., 2003) that mediates between the LMS and its clients. Finally, users of the virtual campus access their services through the user interface layer, built on the business service layer.

The architecture proposed for the external tool applies multitier architecture with integration, business logic and presentation layers, isolated via their corresponding interfaces.

In order to integrate the user interface of the guest application into the host application, the virtual campus uses the HTML element Object (W3C, 1999), which, assisted by a view helper (Alur et al., 2003), embeds the user interface of the guest application within its user interface. In this way, the content aggregation pattern of the mashup approach (Ogrinz, 2009) and the presentation layer mashup (Hanson, 2009) are implemented in the context of a multitier architecture (using front and application controllers (Alur et al., 2003)). Content aggregation is the mashup pattern most directly related to the integration of e-learning applications in the Ogrinz (2009) pattern catalogue.

Using this architecture, the presentation tier of the host application interacts with the guest application using the URLs that give access to its user interface directly, thus creating a RESTful-type interaction between the host application and the user interface of the guest application (note that in this figure, the interface with the REST stereotype only has HTTP methods available: GET, POST, PUT and DELETE (Hansen, 2007)). This is a very simple, functional solution. However, it has an important drawback: although it can be used from Apple Safari, Google Chrome, and Mozilla Firefox, it does not work with MS Internet Explorer. The user interface does not work properly in Microsoft’s browser when a complex object is included in a web page, which makes the solution proposed in this paper unviable. However, we have not found this problem in the rest of the browsers previously mentioned.

For its part, the guest application does not need to interact with the host application using host user interface. The host application services (Alur et al., 2003) are therefore exposed directly, or using a web service broker (Alur et al., 2003), as web services to the external tool services.

This is, therefore, an asymmetric architecture: the host application interacts with the guest application using its user interface as RESTful web services (depicted in blue in Figure 1), while the guest application interacts directly with the host application services using SOAP or RESTful web services (depicted in green in Figure 1) (Hansen, 2007). Note that the asymmetry does not come from the nature of the web service implementation (SOAP vs. RESTful), but has to do with how the host application interacts with the guest application, via its user interface, while the guest application interacts with the host application via its business service layer.

This architecture is so generic that it could be applied to any two applications built using a multitier architecture that had to be integrated in line with the requirements listed in the introduction. Using this architecture, two main types of use case appear: those that are invoked by the user in the host application user interface to trigger an action in the guest application, and those that are invoked by the user in the guest application user interface (integrated in the context of the host application) to trigger an action in the host application.

In both types the same actions have to be performed:
- Information encoding.
- Application invocation and information sending.
- Information reception and decoding.
- Information processing.

However, there are differences due to the asymmetric architecture. In the case of invocation from host to guest the requests are made to the guest using its user interface. This forces a RESTful interaction between host and guest, as well as the adaptation of the code of the guest application to transmit the information belonging to the host application that has been received and has to be transferred to the business layer (e.g., a virtual campus user id).

In the case of invocation from guest to host, the requests are made to the guest using its business layer. Therefore, this interaction can be made using SOAP or RESTful web services, and it may be necessary to provide new services.
in the business layer of the host application for dealing with the new request from the guest application (e.g., the uploading of a new type of content).

If both applications do not belong to the institution, the proposed approach could not be applied, because the initial requirements call for changes in both parts, at least in their user interfaces:
- The host application user interface has to embed the guest application user interface.
- Both applications need new widgets (e.g., buttons or links) for invoking the new functionality that appears as the result of the integration.

Example of integration

Integration of VCAA virtual campus and CREASE tool

This section provides a specific example of the integration architecture proposed: the integration of the VCAA virtual campus with the CREASE tool (the virtual campus with integrated tool is accessible at http://solaris.fdi.ucm.es/CVMB2/ with user demo and password demoKey). The main goal of the VCAA project was to develop a virtual campus independent of the underlying LMS responsible for implementing the core e-learning processes (Navarro et al, 2010). Thus, users interact with a generic virtual campus that hides the underlying LMSs. Figure 2 illustrates the VCAA architecture. According to this figure there are two main components: one responsible for core e-learning facilities and the other for management issues. These business processes are isolated from LMSs using the VCAA canonical interfaces, which are defined as SOAP web services. These interfaces are designed for the management of users, courses, resources, announcements, forums and platforms, and implemented for Blackboard Learn 9.1, Moodle 2.0 and Sakai 2.7.

![Figure 2. VCAA software architecture](attachment:image.png)

This architecture has two main advantages: (i) platforms can be changed, and users are unaware of these changes; and (ii) whenever implementation for the VCAA Canonical Services is provided, platforms can be changed without the need to rewrite a single line of code of the virtual campus itself.

At present, this virtual campus has been deployed and is undergoing testing by three control groups of undergraduate and graduate courses, with a total of 91 users.

Our external tool is a search tool called CREASE (http://creaselsi.uned.es), developed to cluster and organize unstructured didactic resources using Formal Concept Analysis (FCA), a mathematical theory that can be used for
the building of lattices of resources (Cole & Eklund, 1996; Ganter & Wille, 1999). The CREASE tool enables users to perform web searches for didactic resources, organizing them as lattices of resources depending on their descriptors. Thus, the developing framework of the CREASE search and classification tool is used to author learning content. To undertake this task, we first characterize instructional uses (i.e., goals) by means of a set of terms accurately defined by instructional designers. Secondly, we perform a number of search engine queries related to a domain keyword and enriched with the previous terms. Finally, the resources retrieved are analyzed and organized using FCA, in order to generate a concept lattice which could be explored by the user to select resources relevant to the initial pedagogical needs (Mayorga et al., 2012). As a working hypothesis, we assume that the resources retrieved by each query will be bound to a goal if there is sufficient overlap between the enriching terms in the queries and those characterizing the goal. Figure 3 shows how the components of the CREASE tool work together through the orchestrator component to perform these tasks.

Architectures depicted in Figure 2 and 3 fit the architectures of the host and guest applications in Figure 1. According to this integration architecture, the goal is to obtain an integrated application such as the one depicted in Figure 4. In this figure the VCAA virtual campus user interface embeds the CREASE tool user interface. The CREASE tool searches for resources in the web, organizes them in lattices, and uploads these lattices in the VCAA virtual campus where the interaction originally started.

Thus, users from the virtual campus can launch the CREASE tool, search for resources in the web, and upload lattices of resources to specific virtual campus courses. Later, these lattices can be visualized in the context of the course using the CREASE tool. For example, Figure 4 illustrates how the CREASE tool shows the final concept lattice for the original query “Algebra” in the context of the VCAA virtual campus user interface. The diagram depicted is based on Hasse diagrams (Wille, 1982; Wille, 1992). The blue concept node represents the formal concept currently selected and the green ones represent those formal concepts which are closely related. The red formal concept is the top of the concept lattice and, finally, the green concept above the top formal concept represents a formal concept not comparable or not related to the formal concept that the user is currently exploring.

Using these two applications, this section provides a detailed design for the generic architecture of Figure 1, considering an implementation for the two main types of use which are of interest: those that are invoked by the user in the host application user interface to trigger an action in the guest application (use case uc1), and those that are invoked by the user in the guest application user interface (integrated in the context of the host application) to trigger an action in the host application (use case uc2).

Examples of these two main types are as follows:

- **uc1**: We consider the use case *show lattice*. This is a use case with its origin in the virtual campus user interface. Thus, a lattice previously uploaded to a course is displayed using the integrated CREASE tool user interface, as shown in Figure 4.

- **uc2**: We consider the use case *upload lattice*. This is a use case with its origin in the integrated CREASE tool. A search lattice built by the CREASE tool is thus uploaded to the virtual campus course, where the interaction with the CREASE tool originally started.
Figure 4. CREASE tool (pointed blue) integrated within the VCAA Virtual Campus (pointed green)

Uc1 show lattice

Figure 5 depicts the detailed class diagram for the presentation tier of the host application sketched in Figure 1 using a **UML 2 class diagram** (Arlow & Neustadt, 2005) enhanced with **UML WAE notation** (Conallen, 1999). Basically, according to UML WAE notation, client pages represent contents directly retrieved by the web server (e.g., HTML pages), server pages represent computational resources run by an extension to the web server (e.g., Java servlets run by Tomcat), client pages can be built by server pages and navigated associations represent hyperlinks.

The key issue is to be able to include the CREASE tool in the context of the Virtual Campus user interface. In order to solve this issue, the HTML element Object is used. This element allows an HTML page to be loaded inside another HTML page, enabling a nested interaction. The virtual campus user interface uses a view helper to build its user interface, including an HTML Object element with the URL of the CREASE tool. This URL includes encoded information about the virtual campus session (course id, user id, etc.) as well as information about the lattice.

The interaction between object instances of these classes for implementing use case uc1 is depicted in Figure 6 using a **UML 2 sequence diagram** (Arlow & Neustadt, 2005). In this figure the view helper builds the URL that invokes the CREASE tool user interface in the context of an Object element nested in the VCAA virtual campus user interface. The information is encoded using a dedicated class and the Object element with the target URL is printed in the HTML page generated.
Figure 5. Detailed design of the presentation tier for the use case uc1: show lattice

Figure 6. Virtual campus user interface including the nested Object element that loads the CREASE tool responsible for showing the lattice in the VCAA virtual campus user interface
Main is a GWT entry point for the implementation of use case uc1

For the management of this invocation, submitted via an HTTP GET request, the CREASE tool user interface has to decode the information received, and to display the lattice as it would be displayed by the application running as a standalone application. The CREASE tool is a J2EE application whose user interface has been built using Google GWT (Tacy et al., 2013). Therefore it is necessary to define a GWT EntryPoint for the invocation of the component responsible for the implementation of use case uc1. Figure 7 depicts this entry point (called Main.)

As Figure 8 shows, when this entry point is invoked from the VCAA Virtual Campus, it receives the information sent (virtual campus session id plus the lattice), decodes it, stores the virtual campus session id in the HTTP session, shows the lattice, delegating to the application service defined in Figure 7.

**Figure 7.** Main is a GWT entry point for the implementation of use case uc1

**Figure 8.** Interaction at the GWT entry point for the implementation of use case uc1 in the CREASE tool
Uc2 upload lattice

For the implementation of use case uc2 (upload lattice) the CREASE tool simply needs to invoke the web services provided by the VCAA Virtual Campus using the WSDL web services provided by VCAA Virtual Campus. Figure 9 depicts the CREASE application service responsible for implementing this use case.

![Figure 9. Detailed design of the business logic for the implementation of use case uc2](image.png)

When the application needs to upload a file or a lattice to the VCAA Virtual Campus, the CREASEIntegrationASImp class extracts the necessary information from the session (virtual campus id as well as the lattice of the file that has to be uploaded), encodes the information, and invokes the implementation of the WSDL VCAA web services using the JAX-WS proxies (Hansen, 2007). Figure 10 shows this interaction.

![Figure 10. Interaction in the CREASE tool for lattice uploading in the VCAA Virtual Campus](image.png)
Discussion

This paper focuses on the software component of application integration. However learners and educators are the main beneficiaries of the solution. For example, in the context of a large virtual campus (such as the UCM virtual campus) particular users from specific Faculties or departments may need specialized learning object repositories to handle the unique metadata from each domain (Navarro et al., 2013). It would be desirable to have a generic integration architecture that is able to combine those learning object repositories and virtual campuses. Learning objects could thus be built using the specific tools from each domain and these objects could then be integrated in the virtual campus.

Therefore, educators should be aware that they would be able to develop tools for their specific learning needs and that these tools could subsequently be integrated in virtual campuses, making it easier to provide domain-specific learning objects in generic virtual campuses.

Examples of these tools are almost endless, provided that they can be integrated in the architecture described in this paper: repositories of learning objects, tools that promote collaborative and constructive learning or simulators. Tools that can be integrated in virtual campuses can be conceived as satellite tools (Fernández-Valmayor et al., 2011). From an educator’s point of view this paper provides a software architecture that can integrate software applications into a single user interface.

In simple terms we could divide the software application into two parts: the part responsible for interacting with the user through a user interface (the view) and the part responsible for processing the input provided by the user (the model, which in this paper has been split into business logic –processes- and the integration layer -data access-, in line with multi-tier architectures). In traditional integration architectures a software application (let’s call it application1, with view1 and model1), interacts with another software application (application2, with view2 and model2) using both models in a coordinated manner to obtain application1', with view1' and model1+model2.

With the approach presented in this paper a new software application is created that not only integrates models but also views, which could be expressed as application12, with view view1+view2 and model model1+model2.

This integration has an important advantage: the complex user interface (view) needed to access an application’s business logic (which is part of the model) can now be reused in a coordinated manner with the user interface of another application. Traditional integration architectures mainly focus on model integration and not on views integration.

There is, however, one important drawback to this solution: programmers must have access to both applications in order to integrate them because the user interface of the first application, the host application, has to include the user interface of the second application. Conversely, the guest application has to include facilities for interacting with the host application. This cannot be resolved without access to the source code of both applications.

Conclusions and future work

Nowadays virtual campuses have become complex software applications that enlarge LMS functions with services needed by educational institutions or by users.

In these campuses, users demand the enhancement of the learning possibilities provided by classic LMSs with the use of external e-learning tools, such as those designed for the definition, search and management of complex learning objects. However, in most cases, virtual campuses and external e-learning tools are developed independently.

This paper describes a pattern-based integration architecture for virtual campuses with external e-learning tools. According to this architecture, virtual campus users directly interact with the tool thanks to the inclusion of the tool in the virtual campus user interface using the HTML Object element, and some REST-inspired techniques for communication between both applications. The external tool interacts with the virtual campus using a set of services published as web services (SOAP or RESTful).
This solution is therefore based on an asymmetric architecture: the virtual campus exposes its services as web services, whereas the external tool is directly integrated in the context of the user interface of the virtual campus. This asymmetry is necessary because there is no sense in exposing the external tool services as web services and then rebuilding the external user interface in the context of the virtual campus, as the direct inclusion of the user interface of the external tool is a simpler and more straightforward solution. In addition, the presence of the virtual campus services as web services facilitates the integration of the virtual campus in the context of a large educational institution, where integration needs are very significant. The proposed architecture is contextualized in terms of e-learning applications, considering virtual campuses as host applications and external tools as guest applications, but it could be reused in other applications with similar requirements.

Regarding the integrated interaction between host and guest applications, there are differences due to the asymmetric architecture. In the case of invocation from host to guest the requests are made to the guest using its user interface. This forces a RESTful interaction between host and guest, as well as the adaptation of the code of the guest application to transmit the information belonging to the host application that has now been received and has to be transferred to the business layer (e.g., a virtual campus user id).

In the case of invocation from guest to host, the requests are made to the guest using its business layer. Therefore, this interaction can be made using SOAP or RESTful web services, and it may be necessary to provide new services in the business layer of the host application to deal with the new request of the guest application (e.g., the uploading of a new type of content).

The architecture is based on the integration of mashup and multitier architecture. The presence of these patterns guarantees design maintainability as well as other positive properties. In the proposed solution the mashup pattern’s content aggregation and presentation layer mashup are implemented using a view helper multitier pattern that builds the URL used by the embedded HTML Object of the host application user interface. Other multitier patterns used in this integration architecture are front and application controller, application service and web service broker. Obviously the rest of the twenty one multitier patterns (Alur et al., 2003) can still be used in the architecture.

This architecture has been validated by integrating an FCA-based search and classification tool in a virtual campus. Thus, users of the virtual campus can use this tool in the context of a virtual campus course, reusing the content extracted by the tool.

Currently, the VCAA Virtual Campus is being tested by two university courses with 91 users, with promising results. Future work includes the development of more functionality in the VCAA Virtual Campus, and its integration with other external e-learning tools. In addition, we are working on a solution to enable the integrated use of both tools in Microsoft Internet Explorer and validation of the approach with a broader range and number of users.

Acknowledgements

Ministerio de Ciencia e Innovación (projects AACV TIN2009-14317-C03-01 and CREASE TIN-2009-14317-C03-03), Universidad Complutense de Madrid (group 921340), and e-Madrid project (S2009/TIC-1650, Investigación y Desarrollo de tecnologías para el e-learning en la Comunidad de Madrid) have supported this work.

References


Navarro, A., Rodríguez-Artacho, M., Huertas, F., Cigarrán, J., & Buendía, F. SOA integration of a tool for retrieving open learning resources into a modular virtual campus. In M. B. Nunes & M. McPherson (Eds.), IADIS e-learning 2012 (pp. 262-269). Lisbon, Portugal: IADIS.


The Effects of Student Question-Generation with Online Prompts on Learning

Fu-Yun Yu1* and Kuan-Jung Pan2

1Institute of Education, National Cheng Kung University, No. 1, University Road, Tainan City 70101, Taiwan // 2National Nanke International Experimental High School, No. 6, Lane 12, Dashun 6th Rd., Tainan City 74146, Taiwan // fuyun.ncku@gmail.com // shake1230@gmail.com

*Corresponding author

(Submitted June 27, 2013; Revised October 28, 2013; Accepted November 25, 2013)

ABSTRACT

The focus of this study was to investigate the effects of student-question generation with online prompts on student academic achievement, question-generation performance, learning satisfaction and learning anxiety. This study adopted a quasi-experimental research design. Two classes of eighth grade students (N = 64) from one middle school participated in weekly 45-minute online question-generation learning sessions for 6 consecutive weeks. Data analyzed using ANCOVA indicated statistically significant differences between the two groups in regard to academic performance and question-generation performance with students assigned to the online prompts group performing significantly better than their counterparts assigned to the without prompts group. However, there were no statistically significant differences between the two groups in the areas of learning satisfaction and anxiety. Empirical and pedagogical significance of this study together with its implications for instructional implementation, system development, and future studies are provided.

Keywords
Attention capacity, Learning effects, Online prompts, Scaffolding, Student question-generation

Introduction

Student question-generation (hereinafter named SQG), grounded in constructivism, information-processing theory, metacognitive theory and self-determination, is an instructional and learning strategy. While definitions vary in regard to specificity, SQG is applied in most subjects as a process by which students reformulate given questions (from textbooks or teachers) or generate new questions around areas of the study material or in response to their previous classroom activities or experiences (Yu, Wu & Hung, 2013). During SQG, students are directed to be creative and motivated to construct, organize, connect and interact with information and personal experience so that higher-order cognitive competencies are targeted (Yu, Liu & Chan, 2005) rather than the lower level of Bloom taxonomy (i.e. memorization). Evidence from existing studies for the past few decades has supported the use of SQG for the enhancement of comprehension of learned content (Brown & Walter, 2005; Drake & Barlow, 2008) and the promotion of motivation (Chin, Brown & Bruce, 2002), group communication (Yu & Liu, 2005) and higher-order cognitive skills (Brown & Walter, 2005; Dori & Herscovitz, 1999; Yu & Liu, 2008).

Even though empirical studies generally have supported the positive effects of SQG, obstacles affecting its widespread diffusion and swift adoption in the classroom have been noted. Specifically, it has been found that an extensive proportion of students did not have experience in SQG during their formal schooling years (Moses, Bjork & Goldenberg, 1993; Vreman-de Olde & de Jong, 2004), and that students have had concerns about their capability and performance as related to this task (Yu & Liu, 2005). Additionally, a majority of students have considered SQG to be either difficult or very difficult (Yu, 2009). As such, issues pertaining to how to include theoretically sound designs to support SQG activities and build an empirical basis for its effectiveness should be a topic of great relevance.

With the distinct features of computers and network technologies, more than a dozen online learning systems with SQG as the focus have been developed during the last decade (Yu & Liu, 2009). While recently developed systems have come to recognize the need for incorporating scaffolds with a theoretical basis to support the task, the efficacy of these built-in support mechanisms has yet to be substantiated (Yu, Tsai & Wu, 2013).

The need for building up solid empirical evidence is compelling in light of recent findings indicating that some of these scaffolding designs are not as potent as they are purported to be. For instance, the effects of online access to
peer-generated questions during SQG activities were examined in one study (Yu & Yang, in press). Contrary to its strong theoretical underpinning in observational learning theory and scaffolding, significant differences in academic performance and weekly question-generation performance were not found between groups with and without online access to peer-generated questions. Another example of a counter-theoretical argument was found in Yu, Wu and Hung’s study (2013). In light of its sound theoretical grounds in motivational theory and social construction of knowledge, the potential of cooperative learning intended to ease the concerns and anxiety associated with SQG was examined. To the researchers’ surprise, student engagement in SQG while working with a partner cooperatively was not found to lead to lessened learning anxiety, nor did it promote student perceptions of the value of the activity for enhancing self-competence, as compared to the individual SQG mode.

Research purpose and questions of this study

Along this line of research, empirical investigation into “if and how theoretically-derived online support will hold up” served as the focus of this study. In particular, the effects of online prompts in the form of key terms along SQG on learning were examined. To provide a comprehensive view of the issue under current examination, both cognitive and affective effects were investigated. Specifically, in addition to academic achievement and SQG performance, taking into consideration that intervention that leads to increased performance but induces negative affective responses from students is not optimal, and that affective components, including satisfaction toward past learning experiences and learning anxiety, have been suggested to have a decisive impact on successful learning (Krashen, 1988), data on these affective areas were collected. In sum, three research hypotheses are proposed:

- There will be significant differences in student academic performance between the group receiving online prompts and the group receiving no online prompts along SQG.
- There will be significant differences in SQG performance between groups receiving versus not receiving online prompts along SQG.
- There will be significant differences in students’ past learning experiences and learning anxiety between the two different groups.

In the following sections, the theoretical foundations of the provision of online prompts for learning (i.e., scaffolding, and limited attention capacity and selective attention) are briefly described below before proceeding to the methodology section of this study.

Scaffolding

Scaffolding is support offered alongside a task to help learners attain pedagogical goals they might not otherwise be able to accomplish (van de Pol, Volman & Beishuizen, 2010). It is aimed at bridging the cognitive gap between a learner’s current ability level and potential developmental level.

Scaffolding has often been suggested as an effective instructional method (Cole, 2006; Pawan, 2008). In practice, various types of scaffolding, each with a distinct purpose, have been designed. For instance, supportive or procedural scaffolding in the form of suggestions and examples alongside a task emphasizes its function as providing guidance, coaching, and modeling (Hannafin, Land & Oliver, 1999; Jackson, Krajcik & Soloway, 1998). Reflective or metacognitive scaffolding, on the other hand, is intended to prompt learners to reflect back on their thinking processes and to promote better execution and management of self-regulatory strategies and self-reflection during the process (Ge & Land, 2004; Hannafin et al., 1999; Jackson et al., 1998). Elaboration scaffolding is geared toward engaging learners in articulating thoughts, constructing explanations, and making justifications (Ge & Land, 2004). Intrinsic scaffolding is aimed at reducing the complexity of a task (Jackson et al., 1998). Strategic scaffolding provides learners with alternative techniques for approaching tasks (Jackson et al., 1998). Despite their differences in terms of functions and foci, overall, scaffolding acts to decrease working memory load while dealing with novel information (Kirschner, Sweller, & Clark, 2006; Myhill & Warren, 2005; Van Merriënboer, Kirschner, & Kester, 2003), to minimize frustration on the part of learners (Rummel & Kramer, 2010; Wood et al., 1976), and to clarify misunderstandings or fill in incomplete information pertaining to targeted operations or concepts (Sharma & Hannafin, 2007).
Scaffolding is a frequently studied concept (van de Pol et al., 2010). Empirically, scaffolding has been found to have a facilitative effect in regard to supporting students' metacognitive and cognitive activities (Azevedo, Winters & Moos, 2004; Belland, Walker, Olsen & Leary, 2012; Berthold, Nuckles & Renkl, 2007; Ge, Chen & Davis, 2005; Hmelo, Holton & Kolodner, 2000; Wolf, Brush & Saye, 2003). Based on a meta-analysis of 180 research studies, Swanson (1999) concluded that scaffolding is one of the nine most effective instructional interventions. In addition, based on a review of the past decade of research (from 1998-2009) on scaffolding in teacher-student interaction, the conclusion that scaffolding is effective was reached (van de Pol et al., 2010). More recently, the effects of scaffolding on student cognitive outcomes were demonstrated by a meta-analysis of STEM education at the K-12, college, graduate, and adult levels, and the results showed that scaffolding positively impacts student learning, producing an average effect size of .53 (Belland et al., 2012).

While researchers have been successful in devising different forms and means of scaffolding (Tharp & Gallimore, 1988; Wood et al., 1976; van de Pol et al., 2010) to support various learning activities and contexts, its effects for SQG have yet to be fully realized and understood. Hence, it serves as the focus of this study.

**Limited attention capacity and selective attention**

In most learning scenarios, a great deal amount of information is presented continuously, which needs to be effectively and efficiently processed. Because people have a definite attention capacity at any given time, only a restricted portion of the thousands of stimuli reaching our senses can be attended to (Ruz & Lupiáñez, 2002).

According to Treisman's attenuation theory (1964), the attention to some incoming information from different sensory channels and sources will be selectively processed to various extents so as to prevent the information-processing system from becoming overloaded (McLeod, 2008). With this limited capacity constraint, issues regarding “how to better direct learner attention to promote further processing of relevant information” should be topics of great importance.

Previous research shows that learners can use information about an upcoming target or event to enhance task performance (Moher & Egeth, 2009; Simons, 2000). To enable materials deemed important and relevant being attended to and processed as intended, several strategies have been proposed and effects attested to. Specifically, a great deal of attention was given to the influence of directing learner focus of attention by inducing pre-instructional strategies, such as objectives, overview/outlines of keywords, and advance organizer on subsequent learning in the sixties and seventies (Hartley & Davies, 1976). Their function to alert students to aspects of the learning material that might be overlooked while reducing time and attention allocated on irrelevancies has been acknowledged, and their significant effects on aiding learning have generally been supported (Hartley & Davies, 1976).

In light of the predominately discernible effects of pre-instructional strategies used to preface the teaching to come, in this study, the effects of providing students with online prompts to preface SQG to come is examined in this study. It is well-known that goal-directed attention, one of the two determinant factors of selective attention (Chun, Golomb, & Turk-Browne, 2011; Ruz & Lupiáñez, 2002; Yantis, 2000), is affected by a subject’s expectancies and requires conscious awareness (Ruz & Lupiáñez, 2002). The authors hypothesized that a clear set of key terms could act as informative cues for SQG, and could therefore help modulate and orient learner goal-directed attention to important areas for further processing to yield positive learning effects.

**Methods**

**Study context and participants**

This study was conducted in a middle school in Tainan, Taiwan. While attaining high performance on exams is still imperative to enter top-tier universities, in response to the many impending challenges nation- and world-wide, the Ministry of Education in Taiwan has undertaken a series of educational reforms, where core competencies are identified. Among these, problem-solving, creativity, ICT skills, and independent thinking as part of the ten identified core competencies for 1st to 9th grade students (Ministry of Education, 2012) resonated well with online SQG. As such, the online SQG activity implemented in this study was well received. The activity was introduced to support...
student learning of civics and citizenship, part of social studies curriculum, at the secondary school level. Two eighth grade classes (N = 64) taught by the same instructor were randomly selected to participate in this study for 6 consecutive weeks.

The integrated online learning activity was held in social studies weekly study sessions in the school’s computer lab to accompany the weekly 45-minute instructional session allocated for civics and citizenship education. As multiple-choice is the only question type used in the school-wide and high school entrance examination for social studies, it was chosen as the question type for the online SQG activities.

**Instructional content**

Two chapters were covered during the study. The first chapter is about “civil laws and life”, with four lessons that cover topics including the governing of civil laws, principles regarding the exercising of civil laws, civil law liability, and the capacity of civil laws. The second chapter is on “criminal law and administrative regulations”, with four lessons. Topics covered include crime and punishment, elements of crime, basic knowledge of criminal law, and administrative regulations and responsibilities.

**Experimental design and treatment groups**

A quasi-experimental research method was adopted for this study. For the purpose of this study, two treatment conditions were devised: the online prompts (experimental) group and the without prompts (contrast) group. Except for the online prompts that were accessible only to the experimental group, all instructional elements and procedures were kept the same for both groups.

For the experimental group, prompts were included in the system and made accessible to students as part of the SQG function for the duration of this study (see Figure 2). Prompts were included in an attempt to simplify the SQG task for the student (Tharp & Gallimore, 1988) and for better attention management (McLeod, 2008; Ruz & Lupiáñez, 2002). Essentially, the online prompts were in the form of a set of key terms related to the study material and were derived from the “the answer is” approach, which was first proposed by Stoyanova & Ellerton (1996) for SQG. By encouraging the construction of questions matched with ‘the answer’ (i.e., targeted key terms), not only the language, concept, and structure with regard to the ‘target’ may have a better chance of being recognized and understood (Stoyanova, 2003), but also different aspects of the target as observed in daily life can be linked and appreciated (Silver, 1994). Simply taking ‘civil law’ as the target, the following question exemplifies how SQG enables students to construct contextual questions and interconnect different concepts related to the target: Chi-Hong and Shu-Hwa (1) are getting married, so (2) they took out a loan of 2 million dollars from a bank, (3) bought an apartment in the An-Ping district in Tainan, and (4) paid property taxes on time. Which of the above behaviors are governed by "civil law?" (a) 123 (b) 124 (c) 234 (d) 134 (answer key: a). As for the contrast group, no prompts were provided while the students completed the SQG task throughout the study.

In sum, the main difference between the two groups lies in the fact that students in the experimental group were advised to generate questions around key terms in reference to a set of online prompts plus the study material, while students in the contrast group were instructed to generate questions around key terms only with reference to the study material.

**Online learning system**

An online learning system with a focus on SQG, called QuARKS, was adopted (Yu, 2009). Like all similar online systems on the market, QuARKS enables multimedia files to be included as parts of the question; text with different fonts, sizes, and styles can be used, and questions can be constructed to be saved, retrieved, revised, and deleted with ease by users. Nevertheless, QuARKS is different from other systems in at least one way. To the authors’ knowledge, it is the only system that supports customizability in terms of specific function(s) available to students and provides context-sensitive scaffolding at any pre-specified time or in real time.
For this study, both SQG and drill-and-practice (D&P) functions were activated for both groups, but the online prompts function was only made accessible to the experimental group along with SQG. With QuARKS’s customizability, different sets of online prompts (in the form of a set of key terms related to the current study content) can be incorporated, updated, and made available in a timely fashion by individual instructors. In the following sections, the activated functions and associated online learning space are briefly explained.

SQG function. After logging into QuARKS and specifying the course, unit, and class (step 1), students simply click on the SQG function button (step 2) and the multiple-choice question icon (step 3) to be directed to the multiple-choice SQG space. For multiple-choice SQG, students simply compose the question-stem (step 1), provide four alternative options (step 2), identify the correct answer key (step 3), and list the key term(s) to be tested (step 4) in the respective fields. After finishing, they click on the “Submit” or “Temporarily save” button (step 5). For the experimental group, students may click on the “Prompts” button placed on top of the “question” field to be directed to a set of online prompts on the current learned content for reference for SQG (see Figure 2).

For D&P, students simply click on the D&P function (step 1) and specify the number of items to answer (step 2, Figure 3) to proceed. At the conclusion of each D&P activity, the correct answer key to each question is shown, and the average accuracy rate is calculated (Figure 4). If time permits, students can choose to engage in the D&P activity again by re-entering the number of questions to answer. The system will then randomly re-select the specified number of questions from the online database and re-sequence the order of the questions and options within the multiple-choice questions.
Experimental procedures

The experimental procedure is delineated in Figure 5. A pilot study involving one eighth-grade class in the participating school (N = 29) was conducted to ensure that the planned procedures and time allocation for various activities were appropriate, and that the instruction for experimental implementation was clear prior to the actual study.

This study took place right after the school-wide first-term exam. Participating classes were randomly assigned to the two treatment groups. Data on participants’ academic performance on the exam in regard to civics and citizenship was collected.
At the first class session, adopting the suggestions of Rosenshine, Meister and Chapman (1996), elements supporting SQG including the following were included and explained first: (a) the set of criteria that the teacher would use to assess SQG performance (see measurement section for details), and (b) models of appropriate questions. Emphasis was placed on various possible ways of generating questions targeting any key term(s) in such a way that significant details (e.g., important features, conditions, characteristics) could be addressed. For instance, for ‘abandoned inheritance,’ two examples were provided as models. Example #1: When her father died, Mei-Ling decided not to inherit either the estate or debts. Which of the following inheritance types is Mei-Ling practicing? (a) Unconditional inheritance (b) Limited inheritance (c) Abandoned inheritance (d) None of the above (answer key: c). Example #2: When turning 20, Mai-Li could inherit from her father 5 million dollars in estate and 7 million in debts. If Mai-Li wants to abandon inheritance, when should she file the claim to the court? (a) within 2 months upon the known fact (b) within 3 months upon the known fact (c) within 2 months of the death of her father (d) within 3 months of the death of her father (answer key: b).

In addition, rules of thumb for generating multiple-choice question items and the operational procedures for the adopted system were demonstrated before having students practice generating questions on one paragraph of the study material. Then, students assigned to both treatment groups were instructed to generate as many multiple-choice questions as possible on key terms learned in civics and citizenship class in the current week in the corresponding spaces in the QuARKS interface during the remaining time allotted. Students were encouraged to refer to the study materials and supplementary handouts, including the SQG criteria given during the SQG activities.

Each week, starting from the second SQG activity, whole-class feedback on SQG was provided first by highlighting the strengths and weaknesses of three purposefully selected pieces of student-generated questions (10 minutes). This was followed by a 10-minute D&P activity on peer-generated questions selected by the instructor before the SQG activity for the remainder of the class time (25 minutes). At the end of the sixth SQG session, students were advised to complete a posttest and a questionnaire to collect data on outcome measures.

**Measurements**

**Academic performance on civics and citizenship**

A posttest developed by the participating teacher was used to assess student academic performance on the covered content. It included 50 multiple-choice question items consisting of both teacher-generated questions and student-generated questions selected from the two respective participating classes. To eliminate the practice effect on student-generated questions, only teacher-generated questions (27 items) were included for data analysis. Based on item analysis, the quality of the posttest was satisfactory, with test difficulty averaged at .73 and discrimination index averaged at .43.

**SQG performance**

To assess SQG performance, the set of criteria developed by Yu and Wu (2013) was referred to and adapted, which entails the following six dimensions:

- **Fluency (0-3):** question stem and options free from typos (1); precise meaning with one correct answer key (1); complete: question-stem, four options, answer key and annotation (1)
- **Complexity (0-2):** question involves one key term (0), two key terms (1), more than two terms (2)
- **Elaboration (0-3):** interconnectedness between the currently study topic/unit and prior topics/units (1); interconnectedness among topics within the currently studied lesson (1); plausibility of alternatives (1)
- **Originality (0-2):** link to personal life experience or current news (1); scenario-based with contrived characters and story (1)
- **Cognitive level (0-2):** use of language taken directly from the learned materials (0); use of one’s own words to define or describe learned content (1); link across topics/lessons/chapters and inference is needed (2)
- **Importance (0-1):** important concepts of the study material (1)

All questions that students generated during SQG during the first and last activities were analyzed, scored, and summed up according to the defined criteria. For inter-rater consistency, one civics and citizenship education teacher
from the participating school was recruited and trained to use the criteria. The Pearson correlation efficient between the two raters was .92, based on SQG performance assessment of 20% randomly drawn student-generated questions. Definitions and examples of each of the six dimensions of SQG performance were explained to students during the training session. With an explicit scoring scheme for SQG performance assessment coupled with weekly whole-class feedback to students, constructing questions that tap into higher-level cognitive skills (e.g., comprehension, application and analysis), as opposed to memorization, were made clear to participating students.

**Satisfaction toward the past learning experience and learning anxiety**

Hung’s (2000) Learning Satisfaction Scale (nine items, Cronbach’s $\alpha = .91$) and Learning Anxiety Scale (nine items, Cronbach’s $\alpha = .85$) were adopted. On the ‘Learning Satisfaction Scale,’ students were asked to assess their satisfaction toward the activity they were engaged in. Sample items on this scale included: It was enjoyable to be able to participate in this activity; I like to learn via this kind of learning arrangement; I am satisfied with my performance in this activity.

With regard to learning anxiety, students were asked to gauge their emotional states during the activity, such as their level of tension, restlessness, uneasiness, concern, and insecurity. Sample items were: I felt uneasy and confused in this kind of learning arrangement; I would become upset if I knew I would have to participate in similar activities again in the future; I felt pressure learning in this way; it terrified me to learn in this way.

Each statement on the scales was rated on a six-part discrete scale with corresponding verbal descriptions ranging from very inconsistent, through inconsistent, somewhat inconsistent, somewhat consistent, consistent, to very consistent (each response received a weight of 1, 2, 3, 4, 5, or 6, respectively). Both positive and negative statements were included in the scales to counteract possible response-set tendencies. As such, scoring on the statements was adjusted so that negative and positive responses could be summed and analysed, with higher scores reflecting more satisfied feelings and higher anxiety toward the exposed learning experience. Internal consistency reliability values calculated after the actual study were .91 and .85 for ‘Learning Satisfaction Scale,’ and ‘Learning Anxiety Scale,’ respectively.

**Data analysis**

An analysis of covariance technique (ANCOVA) was used to test whether there were significant differences in academic and SQG performance between the two treatment groups. Scores from the first-term exam on civics and citizenship and the 1st SQG performance were used as the covariates. Finally, student satisfaction and learning anxiety were compared using independent-group t-tests.

**Results**

**Academic achievement on civics and citizenship**

Table 1 summarizes the descriptive statistics for both groups’ academic achievement on civics and citizenship. The assumption of the homogeneity of the regression was satisfied, $F(1,62) = 0.05, p = .82 > .05$, before proceeding to ANCOVA. Results of ANCOVA revealed significant differences between the two treatment groups, $F(1,63) = 7.86, p = .01 < .05, \eta^2 = 0.48$.

**SQG performance**

Student SQG performance for the first and last sessions is listed in Table 1. The assumption of the homogeneity of the regression was satisfied, $F(1,62) = 0.931, p = .338 > .05$, before proceeding to ANCOVA. Results of ANCOVA revealed significant differences between the two treatment groups, $F(1,63) = 12.683, p = .01 < .05, \eta^2 = 0.168$. 
Table 1. Descriptive and F-test statistics for academic and SQG performance

<table>
<thead>
<tr>
<th>Treatment groups (n)</th>
<th>Observed variables</th>
<th>With prompts (n = 32)</th>
<th>Without prompts (n = 34)</th>
<th>F</th>
<th>p</th>
<th>η^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>Pre^1 M (SD)</td>
<td>9.78 (3.09)</td>
<td>10.38 (3.11)</td>
<td>7.86^*</td>
<td>.01</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Post^1 M (SD)</td>
<td>10.38 (3.16)</td>
<td>9.20 (3.21)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adjusted M</td>
<td>10.60</td>
<td>8.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQG</td>
<td>1^st, M (SD)</td>
<td>15.72 (8.80)</td>
<td>11.38 (9.63)</td>
<td>12.683^*</td>
<td>.000</td>
<td>0.168</td>
</tr>
<tr>
<td></td>
<td>6^th, M (SD)</td>
<td>22.94 (14.94)</td>
<td>10.29 (9.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adjusted M</td>
<td>21.86</td>
<td>11.3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. ^1^ 1st school-wide exam on civics and citizenship; maximum possible score: 16. ^2^ Teacher-designed test on civics and citizenship; maximum possible score: 27.

**p < .01.

Satisfaction toward past learning experience and learning anxiety

The descriptive statistics of both groups’ responses to the Learning Satisfaction Scale and Learning Anxiety Scale are presented in Table 2. Data analysis indicated that student satisfaction toward past learning experience was not statistically significant, t(66) = 0.88, p = .57 > .05, nor was it for learning anxiety, t(66)= 0.50, p = .34 > .05.

Table 2. Descriptive and t-test statistics for learning satisfaction and anxiety

<table>
<thead>
<tr>
<th>Observed variables</th>
<th>With prompts</th>
<th>Without prompts</th>
<th>t(66)</th>
<th>p</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m, M (SD)</td>
<td>m, M (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction</td>
<td>4.81, 43.31 (7.52)</td>
<td>4.64, 41.79 (6.47)</td>
<td>0.88</td>
<td>.57</td>
<td>-1.94 4.98</td>
</tr>
<tr>
<td>Learning anxiety</td>
<td>4.73, 42.56 (8.06)</td>
<td>4.73, 41.62 (7.11)</td>
<td>0.50</td>
<td>.34</td>
<td>-2.81 4.69</td>
</tr>
</tbody>
</table>

Note. ^1^ m: mean/item; maximum possible score: 6 ^2^ M: Mean/scale; maximum possible score: 54.

Discussion and conclusions

By allowing students to generate questions around the study material, many self-regulatory processes (e.g., rehearsal, organization, elaboration, planning, monitoring, evaluation, revision, reflecting, and so on) are mobilized by the many sub-tasks that SQG entails (Yu & Liu, 2008). Associated sub-tasks may include: identifying important concepts as well as significant details around important concepts, building linkages or distinguishing among related concepts, writing up questions of adequate quality, solving the generated questions, and so on.

As is evident, any task composed of complex sub-tasks, each involving the activation and use of complex processes, as is the case with SQG, will benefit from some type of supportive structuring to help the learner proceed in an effective way. This kind of support is compelling in such learning situations in view of the fact that learners with finite attention and cognitive capacity are frequently challenged to process a vast amount of information. To attain better learning effects, the use of instructional measures intended to enable learners to deploy limited available attention and cognitive capacity to relevant content during learning to bring out expectant learning outcomes while directing attention away from less relevant and sometimes excessive, redundant or even unnecessary signals (attenuating attentional capture phenomena) is important.

In this study, the effects of online prompts to preface SQG were examined. The results indicated that the provision of online prompts helped enhance student academic and SQG performance better, as compared to the without prompts group. Nevertheless, both groups were not distinguished from each other in terms of learning satisfaction and anxiety. The underlying mechanism by which online prompts in this study worked for the benefit of student cognitive growth seemed to manifest at several levels. This can be understood with reference to scaffolding and limited attention capacity and selective attention, as follows:

First, as suggested by van de Pol et al. (2010), Kirschner et al. (2006), Myhill and Warren (2005) and van Merriënboer et al. (2003), the provided prompts, acting as scaffolding, helped reduce the learner’s cognitive load; thus, limited capacity can be devoted to other important aspects of the learning task at hand. Second, they helped
simplify the task for the student by reduction in the degrees of freedom (Tharp & Gallimore, 1988). Third, they functioned as direction maintenance (Wood et al., 1976), or informational cues (Moher & Egeth, 2009; Simons, 2000) for better attention management (McLeod, 2008; Ruz & Lupiáñez, 2002), which aids in keeping the learning on target when a student is pursuing a particular objective.

Briefly explained, locating key terms in the study material will be one of the first tasks, if not the first task, to be attended to when engaged in SQG activities. Plausibly, each student might come up with a different set of key terms out of any specified study content. With key terms highlighted upfront for students, not only can time and effort spent on identifying key terms be saved (i.e., cognitive capacity reduction, task simplification, reduction in the degrees of freedom), students' attention and available cognitive capacity can be focused on understanding, building linkages, and constructing questions around essential content (i.e., direction maintenance, attention management).

As hypothesized, being prompted with a set of key terms along with SQG activities should help modulate and orient learner goal-directed attention. Students restricted by finite attention and cognitive capacity thus will be better able to be directed to key areas of the study material for further processing, and in turn, this will positively influence learning. As was found, students in the online prompts group performed significantly better than those in the without prompts group in terms of academic and SQG performance.

Significance of this study

This work has some empirical and pedagogical significance. First, this study empirically attested to the supportive effects of online prompts based on “the answer is” approach for learning. While evidence substantiating the positive effects of SQG prompts was available decades ago (Rosenshine et al., 1996), its pedagogical values have been mostly examined through comparisons to other instructional arrangements (e.g., summarization, note-taking and review, as done by King, 1992), rather than through a “without prompts” contrast group. To the best of the authors’ knowledge, this is the first study empirically substantiating the superiority of online prompts in the form of a set of key terms to preface SQG.

Second, “the answer is” approach, since proposed by Stoyanova and Ellerton (1996), has been applied exclusively in math. Its applicability beyond its originally designed domain area where mastering vocabulary and terminologies is one essential learning objective (such as language learning, sciences, and social sciences) is legitimate. As demonstrated, its use in civics and citizenship education was explored, and its effects for promoting academic and SQG performance were substantiated. This study demonstrates its uses beyond the currently applied context and helps broaden its pedagogical use.

Implications of this study

Suggestions for instruction, system designs, and future studies are provided on the basis of the current findings and previous studies supporting the pedagogical benefits of SQG strategy. First, instructors interested in SQG are advised to include online prompts in the form of a list of key terms along with SQG. As found, the online prompts devised in this study acted effectively as scaffolding and informative cues to preface SQG and could lead to escalated cognitive effects.

Second, while more than a dozen online learning systems with SQG as the focus have been developed during the past decade, few systems have recognized the need for building up support for complex tasks. Online prompts in the form of a set of key terms were confirmed in this study to be a beneficial design to preface SQG. Yet, to allow context-sensible scaffolds to be incorporated, systems with scaffolding along with SQG should adopt a dynamic rather than a fixed structure. This way, the supports provided can accommodate the current study material and the learner’s cognitive developmental state, as the instructor sees fit, to reflect the notion of scaffolding (Pea, 2004).

Finally, several types of prompts have been proposed to assist the learning and use of a SQG strategy. This study examined the effect of one specific type of prompt (i.e., the answer is) for civics and citizenship education. Because each prompt type has its intended use context and is different in terms of level of concreteness, specifics of focus, and demands on cognitive capacity (Rosenshine et al., 1996), the generalizability of the effects found in this study to
other contexts (e.g., different subject matter, age groups, prompt types) will need to be exercised with caution. Investigations along this line should be fruitful avenues for future research.

References


Yu, F.-Y., & Yang, Y.-T. (in press). To see or not to see: Effects of online access to peer-generated questions on performance. *Educational Technology and Society*.


The Difference in the Online Medical Information Searching Behaviors of Hospital Patients and Their Relatives versus the General Public

Hung-Yuan Wang¹, Jyh-Chong Liang* and Chin-Chung Tsai²

¹Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taipei, Taiwan // ²Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan // D9622304@mail.ntust.edu.tw // aljc@mail.ntust.edu.tw // cctsai@mail.ntust.edu.tw

*Corresponding author

(Submitted May 19, 2013; Revised August 24, 2013; Accepted September 3, 2013)

ABSTRACT

The purpose of this study is two-fold: to explore the differences in online medical information searching behaviors, including evaluative standards and search strategies, of the general public (general group) and those of hospital patients and their relatives (hospital group); and to compare the predictive relationship between the evaluative standards and search strategies of the two groups. A Medical Information Searching-behavior Survey (MISS) was administered. A total of 247 people in the hospital group were surveyed while they were in hospital, and 293 volunteers in the general group were surveyed. The results reveal that the hospital group showed higher tendencies to verify online medical information with mixed evaluative standards and to use more sophisticated search strategies than the general public group after comparing the descriptive results. From the results of regression analysis, the evaluative standards of the hospital group play a less important role in their search strategies. In contrast, the significant relationships between the evaluative standards and search strategies are relatively complex in the general group. Even though some of their evaluative standards are significant factors for predicting search strategies, other factors should be considered in future studies to fruitfully explain their online medical information searching behaviors.

Keywords

Online medical information, Evaluative standard, Search strategy

Introduction

With the explosion of online information in recent years, the Internet has become a preferred source of medical information in our society (Lemire, Pare, Sicotte, & Harvey, 2008; Kammerer, Braten, Gerjets, & Stromso, 2013). Related studies have demonstrated that many people around the world have become accustomed to searching for medical and health information on the Internet (McMullan, 2006; Oh, Kreps, Jun, Chong, & Ramsey, 2012; Renahy, Parizot, & Chauvin, 2010). The growth in online information provides Web users with more opportunities to manage their own health problems and to make decisions about medical issues (Morahan-Martin, 2004; Renahy et al., 2010). That is, in contemporary information society, widely accessible online medical information provides more possibilities to facilitate Web users’ health management (Eysenbach, Powell, Kuss, & Sa, 2002), and thus to expand the health care system (Haux, 2002).

However, the plentiful information on the Internet can not only make Web users feel overwhelmed (Loeber & Cristea, 2003), but it also suffers from low credibility (Metzger, 2007), including inadequate medical information (Hong, 2006; Morahan-Martin, 2004). For example, Web users usually get a great deal of medical information when they are searching with a search engine, but they rarely visit those search results beyond the first page (Morahan-Martin, 2004). Although the Internet is a favorite channel of seeking health information, Oh et al. (2012) reported a higher percentage of trust in traditional media (i.e., newspapers or magazines). The quality of numerous sources of online medical information is one of the issues that has received considerable attention (Hanif, Read, Goodacre, Chaudhry, & Gibbs, 2009; Lemire et al., 2008), as the questionable quality of the information has resulted in potential dangers related to its unsuitable use (Benigeri & Puyte, 2003; Benotsch, Kalichman, & Weinhardt, 2004). Additionally, in some developed nations, the growth of aging population implies the societal need of credible online medical information for health management (Seçkin, 2010).

To respond to this crucial problem, standards and filtering tools have been developed by previous studies (Hanif et al., 2009). On the one hand, many institutions have established standards for evaluating the quality of online medical information, such as the JAMA benchmarks from the Journal of the American Medical Association and the HON

*Corresponding author

280

ISSN 1456-4522 (online) and 1176-3647 (print). © International Forum of Educational Technology & Society (IFETS). The authors and the forum jointly retain the copyright of the articles. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear the full citation on the first page. Copyrights for components of this work owned by others than IFETS must be honoured. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from the editors at kinhuhk@ieee.org.
(Health On the Net) codes from Health on the net foundations. These standards provide many criteria (e.g., authority, accuracy, attributions of source, and so on) for web users to assess, control and assure the quality of the information they find. On the other hand, some quality controls or filtering tools to help them to access trustworthy online medical information have also been established. For example, Sladek, Tieman, Fazekas, Abernethy, & Currow (2006) built a search filter to select appropriate information from the medical literature.

However, Hanif et al. (2009) indicated that the quality of the websites can still vary even if they possess a benchmark, and the filtering system might give users a false sense of security. Metzger et al. (2007) have reviewed students’ cognitive models for evaluating online information and highlighted the importance of related issues, and recommended that the study of users’ perceptions or beliefs when evaluating the credibility of online information should be explored further. After reviewing 129 related studies, Anker, Reinhart, & Feeley (2011) reported that the previous studies were mainly conducted to measure users’ perceptions of source quality or credibility. Studies investigating how Web users assess the credibility of online medical information are relatively few in number. As online medical information begins to play an active role in people’s management of their personal health, exploring Web users’ personal perceptions or beliefs regarding evaluating online medical information and the way they select online medical information is essential (Rains, 2007), and is thus one of the research purposes in this study.

A framework of medical information searching behaviors

In order to explore the standards Web users adopt to filter or select online information, Tsai (2004) conducted an interview study and then proposed a theoretical framework, called “Information commitment.” The framework is an attempt to illustrate Web users’ online searching behaviors, including their evaluative standards and information search strategies on the Internet.

![Diagram of the framework of online searching behaviors](image)

**Figure 1. The framework of online searching behaviors (Categorized as advanced orientations)**

As shown in Figure 1, the framework consists of two components: “Evaluative standard” and “Search strategy.” In the first component, two aspects and four orientations are identified. The first aspect is related to “Standards for Accuracy,” including “Multiple Sources” and “Authority” orientations. Web users would verify the accuracy of web information by either triangulating information from multiple sources or checking the authority of the website. The second aspect covers “Standards for Usefulness,” including “Content” and “Technical” orientations. Web users would consider the relevancy of the content to determine its usefulness on the one hand, or its ease of retrieval on the other. The second component is “Search Strategy,” which refers to the approach that Web users employ to seek information online. It investigates their preferences toward information selecting and processing when fulfilling the purpose of search. In this regard, two orientations are also identified. Web users will use “Elaboration” to integrate collected information, or “Match” to compare existing Web information with intended searches. Among these six
orientations, “Multiple Sources,” “Content” and “Elaboration” are categorized as advanced, and are often adopted by experts, while the remaining three are categorized as less advanced, and are often used by novices (Tsai, 2004).

Wu and Tsai (2005, 2007) then undertook subsequent studies to validate the framework. An instrument for exploring college students’ evaluative standards and search strategies of online information was then developed. After exploratory and confirmatory factor analyses, the results showed that the framework is coherent with Tsai’s study (2004) and the instrument is an adequate tool with sufficient reliability and validity. Accordingly, the framework and the instrument might be suitable for exploring Web users’ online medical information searching behaviors, including their evaluative standards and search strategies.

The possible disparity of online medical information searching behaviors

In a previous study, Liang and Tsai (2009) argued that medical students showed different online information searching behaviors from university students in general, and gender difference existed in the students’ usage of certain searching strategies. Moreover, researchers have indicated that disparities exist in access to and needs of online health information among individuals (Anker et al., 2011; Renahy et al., 2010). Web users with more perceived health risks, such as patients or their relatives, have shown tendencies to search for medical information online (Rice, 2006; Yun & Park, 2010). For example, relatives and patients showed higher motives for online medical information seeking than general populations in Rice’s report. Anker et al. (2011) concluded that Web users’ personal health conditions play an important role in their searching behaviors. Most patients believe that getting enough medical information will be helpful and empowering when discussing their condition with their physicians and making crucial decisions about their health (Rains, 2007). Based on the above literatures, Web users’ online searching behaviors display individual differences, and the disparity of online medical information searching behaviors may exist in different groups. The stronger motives held by relatives and patients might make certain differences from general public. However, researchers have rarely conducted comparative studies to illustrate the differences and to examine the role of relatives’ and patients’ immediate health concerns in their online medical information searching behaviors.

Thus, the present study hypothesized that, people facing medical problems, such as patients and their relatives, are assumed to have a greater need for abundant and accurate medical information, and consequently use different evaluative standards and search strategies from those used by the general public. To compare the different populations’ (i.e., hospital patients and their relatives versus the general public) online medical information searching behaviors is one of the research purposes in this study.

The relationships between Web users’ evaluative standards and their online search strategies

Tsai (2004) has asserted that Web users’ evaluative standards play an important role in their search strategies. Wu and Tsai (2005) then identified the implicit component (i.e., evaluative standards) and the explicit component (i.e., search strategies) from the Information commitment framework, and indicated that the implicit component influences the explicit component. In their study, students who used multiple sources and relevance of web content to evaluate the accuracy and usefulness of online information tended to employ the elaborative information search strategy, while those who used the authority and technical elements of online information tended to use the “Match” search strategy. The results imply that Web users’ different evaluative standards could lead to different search strategies when searching for medical information on the Internet.

As previously mentioned, patients and relatives showed higher motives and more pressing needs in searching online medical information, and may exhibit different searching behaviors from general public. Accordingly, it is hypothesized that the predictive relationships between standards and search strategies may be also different between these two populations. However, the interplay between the evaluative standards and search strategies was mainly explored in single target. The relationships across different groups should be explored, but there is a lack of related research.

In sum, this study aimed to investigate the possible disparities of online medical information searching behaviors existed in different groups, and then compare their predictive relationship between the evaluative standards and
search strategies. For fulfilling these research purposes, an instrument for measuring online medical information searching behaviors was validated as well. The research questions are specified as below:

- Is there any difference in online medical information searching behaviors between hospital patients and their relatives and the general public? If yes, what are they?
- How do people’s evaluative standards predict their search strategies in these two groups? Is there any difference in the predictive relationship between these two groups?

**Methods**

**Participants**

The participants of this study were 540, ranging in age from 30 to 69 years old ($M = 45.1$ years), who self-reported that they had experience of searching online for medical information in Taiwan. The average was 3.67 when asked their frequency of searching for medical information on the Internet on a 7 point Likert-scale (from always [7] to never [1]). Over half of the participants (62%) spent less than 20 hours online per week.

These participants consisted of two groups: hospital group and general public group. The hospital group included 247 patients and their relatives (135 males and 112 females) who aged from 30 to 69. The general public included 293 (140 males and 153 females) whose age was from 30 to 62. In terms of frequency of searching online medical information, their average was 3.71 and 3.64 (7 point Likert-scale) in the hospital group and general public groups, respectively. The participants in both groups showed medium frequencies. All of them were volunteers recruited in the hospital or public places.

**Instrument**

In order to investigate the participants’ evaluative standards and search strategies, a questionnaire called Medical Information Searching-behavior Survey (MISS) was employed in this study. The MISS, which was mainly modified from Wu and Tsai’s (2005) questionnaire, comprised three aspects: (1) standards for accuracy of online medical information, (2) standards for usefulness of online medical information, and (3) search strategies for online medical information. The modified items were examined by three experts who are all university professors and experienced researchers, and their reliability and validity were also examined through exploratory factor analysis (EFA). As previously mentioned, each aspect contains two orientations. The six orientations that include 36 items construct the main structure of the MISS (items are presented in Table 1):

- **Multiple sources as accuracy scale** (Multiple sources) *with 5 items*: measuring the extent to which web users will validate the correctness of unknown online medical information by various sources, such as related websites, prior knowledge, peers or other printed materials.
- **Authority as accuracy scale** (Authority) *with 7 items*: assessing the extent to which web users will examine the accuracy of unknown online medical information by the “authority” of the websites or sources.
- **Content as usefulness scale** (Content) *with 5 items*: measuring the extent to which web users will assess the usefulness of the online medical information by the relevancy of its content.
- **Technical issues as usefulness scale** (Technical) *with 7 items*: assessing the extent to which web users will judge the usefulness of the online medical information by the ease of retrieval, the ease of searching or the ease of obtaining information. Therefore, their standard for evaluating online medical information is more closely related to some technical issues.
- **Elaboration as search strategy scale** (Elaboration) *with 6 items*: measuring the extent to which web users will have purposeful (metacognitive) thinking or integrate online medical information from several websites to find the best fit that fulfills their purpose.
- **Match as search strategy scale** (Match) *with 6 items*: investigating the extent to which web users will tend to start searching from a single search engine, or find only a few websites that provide the most fruitful and fitting information when they search for online medical information. Their strategy is oriented towards matching the purposes of their search.
The MISS in the present study employed a six point Likert-scale in which statements were presented with bipolar strongly disagree/strongly agree anchors (from strongly disagree = 1 to strongly agree = 6). As previously mentioned, three of the six scales (“Multiple sources,” “Content,” and “Elaboration”) are categorized as sophisticated information searching behaviors, while the remaining three (“Authority,” “Technical,” and “Match”) are categorized as less sophisticated. Therefore, the people who score higher in the former three scales could be recognized as possessing more sophisticated online medical information searching behaviors.

Data collection and analysis

In the present study, a convenient sampling method was adopted to collect data. According to the research purposes, the participants consisted of hospital group and general public group. On the one hand, the researchers went to hospital and asked patients or their relatives to answer the MISS. The total number of the hospital group was 405. On the other hand, the researchers invited people who showed interests in searching online medical information to answer the MISS in parks and train stations. The total number of the general public group was 497. Uncompleted data were excluded. The valid sample included 540 adults in total.

In order to validate the MISS, EFA with principle component analysis was conducted. The items with factor loadings less than 0.5 were excluded. Then, for comparing the difference in the online medical information searching behaviors of the two groups, the average scores of the six subscales were calculated and compared by a t-test. Moreover, in order to answer the second research question, multiple regressions were conducted to illustrate how evaluative standards predict the search strategies in each group. Comparing the regression models of the two groups will be helpful to understand the differences in the predictive relationships between the evaluative standards and the search strategies of the two groups.

Results

The validation of the MISS

To clarify the structure of MISS, EFA was used. The EFA results revealed that a total of six factors were extracted with eigenvalues exceeding 1.0: “Multiple sources,” “Authority,” “Content,” “Technical,” “Elaboration” and “Match.” Six items with factor loadings of less than 0.5 were removed to increase its total explained variance and reliability coefficients. In sum, these factors which consisted of 30 items accounted for 61.11% of variance. The factors and responding factor loadings of the items are presented in Table 1. Moreover, the reliability (alpha) coefficients for these factors are 0.74, 0.80, 0.87, 0.79, 0.86, and 0.86 respectively, and the overall alpha is 0.86. Therefore, the MISS is suggested to be a sufficiently reliable tool for assessing online medical information searching behaviors.

Table 1. Factor loadings, Cronbach’s α values and descriptive results for the six scales of the MISS (n = 540)

<table>
<thead>
<tr>
<th>Item</th>
<th>Measure</th>
<th>Factor loadings</th>
<th>Alpha value</th>
<th>Means (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple sources 1</td>
<td>I will compare with relevant medical knowledge I have learned to judge whether the online medical information is correct.</td>
<td>.64</td>
<td>.74</td>
<td>4.96 (0.53)</td>
</tr>
<tr>
<td>Multiple sources 2</td>
<td>I will discuss with teachers or peers, and then judge whether the online medical information is correct.</td>
<td>.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sources 3</td>
<td>I will explore relevant content from books (or print materials), and then evaluate whether the online medical information is correct.</td>
<td>.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sources 4</td>
<td>I will try to find more websites to validate whether the online medical information is correct.</td>
<td>.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authority 1</td>
<td>I will believe in its accuracy if the online medical information is posted in well-known websites.</td>
<td>.61</td>
<td>.80</td>
<td>4.67 (0.69)</td>
</tr>
<tr>
<td>Authority 2</td>
<td>I will believe in its accuracy if the online medical information appears in government websites.</td>
<td>.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authority 3</td>
<td>I will believe in its accuracy if the online medical information is posted in professional (official) websites.</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authority 4</td>
<td>I will believe in its accuracy if the online medical information appears in some websites recommended by experts when I view or navigate the information on the Internet.</td>
<td>.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authority 5</td>
<td>I will believe in its accuracy if the online medical information is presented professionally.</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content 1</td>
<td>If its content fits my searching goal, I will consider the online medical information as useful to me.</td>
<td>.72</td>
<td>.87</td>
<td>4.99 (0.59)</td>
</tr>
<tr>
<td>Content 2</td>
<td>If it can provide more related links, the online medical information is useful to me.</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content 3</td>
<td>If it can help me search further for relevant information, I will think the online medical information is useful to me.</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content 4</td>
<td>If the online medical information is closer to my searching purpose, I will believe more in its usefulness.</td>
<td>.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content 5</td>
<td>If it is highly related to my intended searching content, the online medical information is useful to me.</td>
<td>.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical 1</td>
<td>If the website provides enough online medical information, I will stop searching further and believe in its usefulness.</td>
<td>.73</td>
<td>.79</td>
<td>4.60 (0.80)</td>
</tr>
<tr>
<td>Technical 2</td>
<td>If it is presented by animation, I will think the online medical information is useful to me.</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical 3</td>
<td>If it does not take much time to be retrieved, the online medical information is useful to me.</td>
<td>.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical 4</td>
<td>If it does not require a password or registration, I will think the online medical information is useful to me.</td>
<td>.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaboration 1</td>
<td>I am used to searching for online medical information from different websites (or pages).</td>
<td>.68</td>
<td>.86</td>
<td>4.75 (0.66)</td>
</tr>
<tr>
<td>Elaboration 2</td>
<td>I am used to summarizing a variety of online medical information.</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaboration 3</td>
<td>I can use some acquired online medical information for advanced search to find the best-fit information.</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaboration 4</td>
<td>I will continue to judge what kind of online medical information I need during the searching process.</td>
<td>.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaboration 5</td>
<td>I can integrate the online medical information obtained from a variety of websites.</td>
<td>.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaboration 6</td>
<td>I can compare different online medical information from related websites (or pages).</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match 1</td>
<td>Sometimes I will forget the purpose of my search during the searching process.</td>
<td>.68</td>
<td>.86</td>
<td>3.91 (0.96)</td>
</tr>
<tr>
<td>Match 2</td>
<td>Sometimes I do not know how to start the searching of online medical information that I want.</td>
<td>.76</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I usually enter the first few websites suggested by the search engine. .80

I usually use a search engine to find one website which is most fitted to my search purpose. .79

When I find the first relevant website, I will not search for others. .80

I am eager to find a single website that contains the most fruitful information. .74

Note. Overall alpha = .86; Total variance explained is 61.11%.

Scores on the six scales

For investigating online medical information searching behaviors, Table 1 also shows the mean scores and standard deviations of the six scales of the MISS. The participants scored highest on “Content” (an average of 4.99), followed by “Multiple sources” (an average of 4.96), “Elaboration” (an average of 4.75), “Authority” (an average of 4.67), “Technical” (an average of 4.60), and “Match” (an average of 3.91). The average scores for the first four scales (i.e., evaluative standards) are all higher than four points, and the mean scores for the “Multiple sources” and “Content” scales are close to five points. This reveals that the participants in this study might opt to employ all these evaluative standards, a so-called “mixed standards” tendency. Moreover, using “Multiple sources” for accuracy and “Content” for usefulness, which are considered as more sophisticated evaluative standards, is much preferred. Regarding search strategies, on the one hand, the participants showed a neutral preference for employing “Match” as a search strategy (M = 3.91). On the other hand, using “Elaboration” as a search strategy was preferred (M = 4.75).

Group differences in information searching behaviors

To compare the differences in the MISS of the two groups (people in hospital and the general public), a series of t-tests were used. Table 2 shows the results of the comparison of the MISS scales identified by the t-tests. The results indicated that the hospital group showed higher tendencies to use “Multiple sources” (t = 2.53, p < .05), “Authority” (t = 4.44, p < .001), “Content” (t = 4.15, p < .001), and “Technical” (t = 2.86, p < .01) as evaluative standards, and “Elaboration” (t = 2.47, p < .05) as their search strategy to seek online medical information, but less orientation toward using “Match” (t = -3.90, p < .001) than the general public group.

The hospital group showed a higher tendency than the general public group on all standards for judging the accuracy and the usefulness of online medical information. In other words, hospital patients and their relatives appear to show stronger preferences for filtering the online medical information with mixed standards. Moreover, they also preferred to use sophisticated search strategies (i.e., Elaboration), and to show less orientation toward using the match strategy.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Groups</th>
<th>Mean</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple sources</td>
<td>Hospital</td>
<td>5.03</td>
<td>2.53*</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>4.91</td>
<td></td>
</tr>
<tr>
<td>Authority</td>
<td>Hospital</td>
<td>4.80</td>
<td>4.44***</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>4.55</td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td>Hospital</td>
<td>5.10</td>
<td>4.15***</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>4.90</td>
<td></td>
</tr>
<tr>
<td>Technical</td>
<td>Hospital</td>
<td>4.71</td>
<td>2.86**</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>4.51</td>
<td></td>
</tr>
<tr>
<td>Elaboration</td>
<td>Hospital</td>
<td>4.82</td>
<td>2.47*</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>4.69</td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>Hospital</td>
<td>3.73</td>
<td>-3.90***</td>
</tr>
<tr>
<td></td>
<td>General</td>
<td>4.05</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05. * * * p < .01. ** * * * p < .001.
The predicting relationships between groups

In order to examine the predictive power of evaluative standards for online medical information in their search strategies within each group, multiple regression analyses were used.

Table 3. Multiple regression estimates for predicting search strategies

<table>
<thead>
<tr>
<th>Predictors</th>
<th>General public group</th>
<th>Hospital group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Elaboration</td>
<td>Match</td>
</tr>
<tr>
<td>Multiple sources</td>
<td>0.27***</td>
<td>0.07</td>
</tr>
<tr>
<td>Authority</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>Content</td>
<td>0.17*</td>
<td>0.07</td>
</tr>
<tr>
<td>Technical</td>
<td>0.12</td>
<td>0.05</td>
</tr>
<tr>
<td>R²</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>Adjust R²</td>
<td>0.21</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 3 shows the results of the multiple regression models of the two groups. In the general public group, “Multiple sources” ($\beta = 0.27, p < 0.001$) and “Content” ($\beta = 0.17, p < 0.05$) could significantly and positively predict “Elaboration,” and explained 22% of usage of the “Elaboration” search strategy. “Authority” ($\beta = 0.28, p < 0.001$) and “Technical” ($\beta = 0.34, p < 0.001$) were significant and positive predictors for “Match,” while “Multiple sources” ($\beta = -0.18, p < 0.01$) was also a significant but negative predictor for the “Match” search strategy. These three factors accounted for 19% of variance. Among the general public group, those who tended to assess the accuracy and usefulness of online medical information by “Multiple Sources” and “Content” might be more oriented to employ the “Elaboration” search strategy. Those who tended to assess online medical information according to “Authority” and “Technical,” and those who were less oriented to “Multiple sources,” might be more likely to employ the “Match” strategy.

In the hospital group, “Multiple sources” ($\beta = 0.26, p < 0.001$) and “Content” ($\beta = 0.30, p < 0.001$) made a significantly positive prediction of the “Elaboration” search strategy (23% of variance explained), while “Multiple sources” ($\beta = -0.16, p < 0.05$) was the only significant but negative predictor for the “Match” search strategy with a relatively low explained variance (5% of variance explained). That is, those participants in the hospital group who were inclined to use “Multiple Sources” and “Content” were more likely to employ “Elaboration” as a search strategy. On the other hand, those who showed fewer tendencies to judge the accuracy from multiple websites were likely to use the “Match” strategy.

These results show that these two models share some commonalities. By and large, in both groups, “Multiple sources” and “Content” are significant positive predictors of “Elaboration,” and “Multiple sources” is a negative predictor of “Match.” However, divergences were also found between these two models. The “Authority” and “Technical” evaluative standards made positive prediction of the “Match” search strategy only within the general public group.

In the hospital group, the significant predictive relationships between their evaluative standards and search strategies are comparatively fewer. For example, only “Multiple sources” and “Content” have predictive roles in their “Elaboration” search strategy. Though “Multiple sources” is a negative predictor for “Match,” the variance explained by this variable is quite low (5%). However, the relations between “Authority,” “Technical” and “Match” were not revealed in the hospital group. Therefore, a finding proposed here is that evaluative standards might play different roles in the search strategies of these two groups. Compared to the relatively rich relationships between these two components in the general public group, the evaluative standards play a more limited role in the search strategies of the hospital group.

Discussion and conclusion

In this study, the structure of the MISS is validated by EFA, and is consistent with the framework proposed by Tsai (2004). Therefore, the MISS is a sufficiently reliable tool to measure Web users’ online medical information.
searching behaviors. It would provide more insights to uncover their perceptions of standards toward evaluating online medical information and preferences of selecting such information.

According to the scores of each scale in the MISS, the participants in this study tend to adopt mixed evaluative standards and sophisticated search strategy. In terms of evaluative standards, this finding is similar to the result obtained from Liang and Tsai’s (2009) study. Web users have to verify the information with numerous criteria simultaneously in order to access accurate online medical information. Hong (2006) indicated that depth and expertise are both contributing factors influencing Web users’ credibility judgments. That is, a website which provides in-depth information (i.e., “Content” is this study) or appears to have experts (i.e., “Authority” in this study) is considered as more credible.

Due to the specific need for accurate medical information, the evaluative standards which are commonly held by novices (i.e., authority and technical) might act as basic criteria on the one hand. For example, Web user can filter some improper information and gain preliminary understandings of specific medical terminology from government or experts’ websites. On the other hand, for comprehensively figure out the trustworthiness of acquired online medical information, advanced evaluative standards (i.e., multiple sources and content) are needed. Tsai (2004) has indicated that the advanced evaluative standards would lead to effective information searching, and be helpful for obtaining adequate online information. Thus, it is supposed that Web users’ might use mixed standards as “basic” and “advanced” criteria for assessing online medical information might be one of the possible tendencies.

Moreover, in this study, the participants even showed higher orientations toward employing not only advanced evaluative standards, but also sophisticated search strategy. “Elaboration” was also considered as a much preferred search strategy by university students and medical students in previous studies (Liang, & Tsai, 2009; Wu, & Tsai, 2007). Due to the abundant amount of unreliable information online, it is possible to find more credible medical information with these advanced evaluative standards and search strategies.

Regarding the first research question, significant differences in the online medical information searching behaviors of the hospital group and the general public group were found in this study. First, the hospital group showed a higher orientation towards employing “mixed standards” to judge online medical information. People in hospital (patients or their relatives) have often encountered a more critical medical problem than members of the general public, and showed higher tendency to search online (Rice, 2006). They have a need to seek the most credible medical information as references to deal with their medical problems.

Second, the hospital group also tended to use more “Elaboration” and less “Match” as search strategies than the general public group. In order to deeply and carefully understand the diagnoses and the information about their diseases, they may summarize and compare the online medical information they found. Elaborating from abundant information will help them to understand the diseases they are facing and to feel empowered when they have to make decisions about their health management. Searching for medical information on the Internet, which is a database including an ever-expanding amount of information, patients and their relatives have to use more advanced search strategies to find credible and useful information.

In terms of the relationships between evaluative standards and search strategies, that is, the second research question, the findings from this study indicate that the predictive model of online medical information searching behaviors of the general public is compatible with the model of Wu and Tsai’s (2005) study. For the general public, their evaluative standards of online medical information could be viewed as important predictors for their search strategies. In other words, the predictive model of the general public was similar to the original theoretical framework proposed by Tsai in 2004.

Nevertheless, the predictive relationship between evaluative standards and search strategies of people in hospital is comparatively limited. The search strategies they use cannot be well predicted by their evaluative standards as they can be for their counterparts in the general public. This finding implies that the decision of which search strategies to use might be influenced by other factors. Relevant studies have indicated that many factors (e.g., gender, age, and medical literacy) influence people’s search strategies (Anker et al., 2011; Pharo & Jarvelin, 2004). As such, the relationship between evaluative standards and search strategies is vague among the people in hospital, and more factors should be considered in future studies.
Furthermore, as the Internet is becoming a popular source of medical information in recent years, people’s low information literacy will impede their access and proper use of online medical information (Benigeri & Pluye, 2003). People in this information society should acquire adequate online medical information literacy to effectively judge the credibility of online medical information. In this regard, the framework of the MISS might provide some appropriate guidelines for instructional design of teaching online medical information literacy. Those evaluative standards and search strategies (i.e., “Authority,” “Technical,” and “Match”) which are commonly held by novices could be introduced first as a basic evaluative standard and search strategy. Then, people should learn how to assess the information with advanced evaluative standards and search strategies (i.e., “Multiple sources,” “Content” and “Elaboration”). To improve people’s online medical literacy, such a course designed with the framework might be useful, but the actual effects should be investigated with more studies.

Additionally, using the MISS could support the development of quality control tools. First, medical information providers should make some additional efforts to help people to verify the credibility of online medical information as a form of guidance. Some ideas from the framework might be applicable, such as providing multiple sources automatically. For example, Schwarz and Morris (2011) have developed an experimental searching engine module that could provide some statistical visualization to “augment” the search results. For each search result, the module will generate a visualization including credibility indicators by checking a database. Accordingly, the medical information providers could offer more relevant credibility information as references for users to judge the information, just as the sophisticated searching standards: “Multiple sources” and “Content.”

Second, designers could use the MISS to identify users’ preferred evaluating standards and searching behaviors. Then, the adaptive quality control tool could provide users with many more references according to their preferred evaluating standards to judge the credibility of online medical information. On the other hand, more sophisticated standards provided by the external quality control tool might stimulate users to rethink their evaluating standards and searching skills, and then to reach the maturity of medical information literacy.

Regarding to the limitation of this study, one is the small number of participants. To clearly demonstrate users’ medical information searching behaviors and validate the MISS, more participants should be surveyed. Moreover, interviews or other qualitative methods should be employed to illustrate participants’ online medical information searching behaviors. It will be helpful for improving the validity and reliability of this study.

Acknowledgements

Funding of this research work was supported by the National Science Council, Taiwan, under grant numbers NSC 99-2511-S-011-005-MY3, NSC 99-2511-S-011 -008-MY3 and NSC 101-2628-S-011-001-MY3.

References


Digital Game-Based Learning Supports Student Motivation, Cognitive Success, and Performance Outcomes

Jeng-Chung Woo
Department of Arts and Plastic Design, National Taipei University of Education, Taiwan // wwwwc2000@gmail.com

(Submitted May 23, 2013; Revised August 23, 2013; Accepted October 21, 2013)

ABSTRACT
Traditional multimedia learning is primarily based on the cognitive load concept of information processing theory. Recent digital game-based learning (DGBL) studies have focused on exploring content support for learning motivation and related game characteristics. Motivation, volition, and performance (MVP) theory indicates that cognitive load and learning motivation simultaneously influence performance. To provide efficient management of learning effectiveness by understanding the latent relationship among cognitive load, motivation, and performance, this study investigated 63 university students using an online game entitled “Operating a Small Factory in Computer-Aided Manufacturing” for 8 weeks. We collected data based using an instructional materials motivation survey, a cognitive load scale, and performance (skill and cognitive) scales. The data analysis results showed that motivation and cognitive load exhibited a significant canonical correlation with performance. This preliminary finding suggests that when designing DGBL, designers should increase motivation and germane cognitive load to enhance learning effectiveness. In addition, the attention component must be compromised. However, relevance, satisfaction, and confidence do not conflict. This study proposes an application of game characteristics to the attention, relevance, satisfaction, and confidence model, and integrated multimedia effects in using DGBL design methods.

Keywords
DGBL; Motivation; Cognitive load; Performance; Computer-aided manufacturing

Introduction
Relevant studies have indicated that digital game-based learning (DGBL) possesses significant potential for increasing students’ learning motivation (Papastergiou, 2009; Huang, 2011) and enhancing their learning effectiveness (Yang, 2012; Wang & Chen, 2010). Higher education was the most frequently referenced research category (from 2001 to 2010) and has shown the highest growth in the number of DGBL studies (Hwang & Wu, 2012).

The attention, relevance, confidence, and satisfaction (ARCS) motivation model (Keller, 1987) has been widely applied to instructional designs to improve learning motivation (Liu & Chu, 2010; Karoulis & Demetriadis, 2005) and evaluate the performance of DGBL learners following motivation stimulation (Kebritchi, Hirumi, & Bai, 2010; Liu & Chu, 2010). Several studies suggest that games can enhance learning motivation because the game characteristics attract learners. Consequently, game characteristics have been systematically analyzed and summarized (Malone & Lepper, 1987; Garris, Ahlers, & Driskell, 2002; Prensky, 2007; Huang & Johnson, 2008). The relevant game software adopted by DGBL studies can be divided into the following two categories: (a) applications of existing commercial software (Kebritchi et al., 2010; Huang, 2011; Yang, 2012); and (b) designs and developments of new software (Papastergiou, 2009; Liu & Chu, 2010). Although both types stimulate learning motivation, new software designs and developments based on specific instructional objectives can better satisfy course demands. Nevertheless, how DGBL design methods can be employed to apply various game characteristics that match motivational strategy procedures remains unclear.

Although using DGBL can improve learning motivation, enhancing motivation in a multimedia learning environment, where stimulating external materials (Mayer, 2001) and interaction in a complex intrinsic game-based learning environment (GBLE) exist, requires considerable cognitive investment from learners to process environmental and social stimulation. This may lead to cognitive overload problems among learners. Improper learning process management can lead to early interruption or termination because of limited learner motivation and possible cognitive process overloads (Keller, 2008).
Various cognitive loads in learners’ working memory, including intrinsic cognitive load, extraneous cognitive load, and germane cognitive load, are task-based mental loads induced by task characteristics, learners’ performance, and mental effort invested (Paas, Tuovinen, Tabbers, & van Gerven, 2003). Because multimedia can process and display information using various methods, employing well-structured multimedia-based instructional designs can reduce learners’ extraneous cognitive loads (Khalil, Paas, Johnson, & Payer, 2005). Furthermore, teaching material designs should enhance learners’ learning motivation and germane cognitive load. Based on the dual-channel assumption, people use separate systems to process visual/pictorial and auditory/verbal representations of information. By combining limited capacity and active processing assumptions, Mayer (2001) developed a cognitive theory for multimedia learning; however, the limitation of this theory is its exclusion of motivational factors. The limitation of the ARCS macro model proposed by Keller (1987) is that it cannot explain how information processing elements are integrated with the learning process or how these elements interact with motivation. Deimann and Keller (2006) indicated that self-oriented multimedia-based learning environments use non-linear (i.e., nodes and hotlinks) and random information access characteristics (i.e., learners determine the sequence of information access), which empowers learners to control the learning process and results completely. Therefore, the cognitive theory of multimedia learning proposed by Mayer (2001) and learners’ volition control have been included to explain the learning motivation process in multimedia-based learning environments. Keller’s MVP model provides more comprehensive explanatory latent variables for the relationships among motivation, learning, and performance. The model also shows the influence that learning motivation and cognitive load during the learning process have on performance (Keller, 2008). However, the theoretical framework has not been practically developed or applied to DGBL environments, nor thoroughly investigated using empirical research.

**Literature review**

**ARCS model**

The ARCS model is a motivational model initially proposed for diagnosing motivational problems and providing strategic suggestions. Comprising the components of attention, relevance, confidence, and satisfaction, the ARCS model has been widely applied to instructional design processes, thereby connecting learning motivation with effectiveness (Keller, 1987). Experimental studies have confirmed that instructional designs of the ARCS model can significantly enhance learning motivation and effectiveness (Kebritchi et al., 2010; Liu, & Chu, 2010).

**Game characteristics**

From a pedagogical perspective, students’ attraction to computer games has been considered to address student learning interests effectively (Thomas, Thomas, Mark, & Elizabeth, 2011). Moreover, game characteristics are what attract learners. Several researchers indicated different game characteristics (shown in table 1).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>categories</td>
<td>challenge, curiosity, fantasy, control</td>
<td>fun, play, rules, goals, interactive, adaptive, outcomes and feedback, win states, conflict/competition/challenge/opposition, problem solving, interaction, representation and story</td>
<td>challenges, competition, rules, goals, fantasy and changed reality, story or representation, engagement and curiosity, role-playing, control, multimodal presentation, tasks</td>
<td>control, rules/goals, challenges, sensory stimuli, mystery, fantasy</td>
</tr>
</tbody>
</table>

Activities should stimulate a portion of learners’ sensory control; in other words, learning effectiveness or achievement is a feeling determined by learners’ actions. Furthermore, learners can be prompted to participate in role-play activities or fantasy scenarios, and can experience imaginary situations and develop intrinsic motivation.
(Malone & Lepper, 1987). Sensory stimuli should be appealing or novel visual and auditory stimuli (Garris et al., 2002). Games typically employ multimodal presentations to effectively increase interest and the instructional effects, and also integrate auditory, visual, and textual presentations to enrich players’ experience (Huang & Johnson, 2008). Therefore, sensory stimuli are similar to multimodal presentations, which include representations or visual factors. Curiosity is a product of perceived discrepancies or knowledge inconsistencies (Garris et al., 2002). Incongruous information, complexity, novelty, surprise, and violations of expectations strengthen mystery (Berlyne, 1960). Therefore, mystery is similar to curiosity. In a multimedia environment with abundant stimuli, curiosity and achievement are positively correlated (Keller, 2008). Prensky (2007) indicated that adaptive characteristics enable learners to generate “flow.” The learner is fully absorbed in the immediate activity of the game and does not rely on metacognitively induced strategies of self-regulation to remain on task (Keller, 2008). Players typically stop playing games because overly easy tasks bore them and tasks that are excessively difficult discourage them. Thus, appropriate challenges are essential for maintaining a “flow” state (Prensky, 2007) and fostering a sense of winning and challenge. Accordingly, adopting a suitable strategy can enable learners to exhibit superior performances (Wang & Chen, 2010).

In summary, game characteristics include fantasy, curiosity, control, role-playing, fun, play, rules, goals, interactive, adaptive, win states, conflict/competition/challenge/opposition, problem solving, interaction, multimodal presentation and story, task, and outcomes and feedback. The video "Dumb Ways to Die," which has a story and is fun, induces curiosity, and involves fantasy, has attracted more than 60 million views (http://www.youtube.com/watch?v=IJNR2EpS0jw&feature=youtu.be). A curious character pokes a bear with a stick. Another character is electrically shocked, becoming a fantasy skeleton. This reminds viewers to be safe around trains (Metro Trains Melbourne, 2012).

**Digital game-based learning**

Students use games to explore and ultimately construct concepts and relationships in authentic contexts. The concept of learning-by-doing comprises core constructivist principles that underlie game-based learning (Yang, 2012). Lenhert et al. (2008) found that 97% of Americans between the ages of 12 and 17 play digital games. Yang (2012) used commercially available games to assist students with understanding economic life and global issues. Kebrich et al. (2010) examined the effects that a computer game had on students’ mathematics achievements and motivation. Liu and Chu (2010) conducted a study investigating how ubiquitous games influence English learning achievements and motivation. Papastergiou (2009) assessed the learning effectiveness and motivational appeal of a computer game for computer science learning. The results show that DGBL can provide effective and motivating learning environments regardless of students’ gender. Huang (2011) used the Trade Ruler game to introduce economic theory and found that it enhanced students’ learning motivation. Therefore, DGBL can enhance learning motivation (Papastergiou, 2009; Kebritchi et al., 2010; Huang, 2011) and improve the learning effectiveness of students (Kebritchi et al., 2010; Liu & Chu, 2010; Yang, 2012). People acquire new knowledge and complex skills from game play (Federation of American Scientists, 2006). The games adopted by DGBL studies can be divided into two categories: (a) applications of existing commercial software (Kebritchi et al., 2010; Huang, 2011; Yang, 2012); and (b) designs and developments of new software (Papastergiou, 2009; Liu & Chu, 2010). Although both stimulate learning motivation, new software designs and developments based on specific instructional objectives can better satisfy course demands. Nevertheless, how DGBL design methods can be employed to apply various game characteristics appropriate for motivational strategy procedures remains unclear.

**Reducing cognitive load in multimedia learning**

The ideal teaching material format should integrate information from various sources to limit unnecessary mental integration by learners and reduce learners’ extraneous cognitive load. Mayer (2001) indicated that in multimedia learning, active processing involves five cognitive procedures: selecting words, selecting images, organizing words, organizing images, and integrating. Cognitive requirements are divided into three types: essential processing, incidental processing, and representational holding. Essential processing refers to the cognitive processes necessary to comprehend teaching material, for which substantial cognitive capacity is used to select, organize, and integrate words and images. Incidental processing refers to cognitive processes that are not required to comprehend the presented materials but are primed by the learning task design. Finally, representational holding refers to the
cognitive processes aimed at retaining a mental representation in working memory over time to support subsequent learning. When the total intended cognitive processing load exceeds the learner’s cognitive capacity, cognitive overload occurs. The solution is to reduce the cognitive load by reallocating essential processing, reducing incidental processing, and/or decreasing representational holding. Mayer (2001) applied the results of multimedia research to propose various solutions (Fig. 1).

Figure 1. Cognitive load reduction methods for multimedia instruction (Mayer, 2001)

The MVP model for motivation, volition, and performance

Keller (2008) extended the ARCS model into an MVP model (Fig. 2), which represents the external input, psychological environment, and outputs. The external input of environmental conditions influences learners’ psychological environments and processes, resulting in outputs of effort, learning and performance, and consequences influenced by contingency management and intrinsic reinforcement strategies. The psychological environment and relevant processes include motivation and volitional processing, motivation and information processing interfaces, information and psychomotor processing, and outcomes processing. Volition is included in motivation and volitional processing. Information processing elements and their interactions with motivation and volition can be explained using information and psychomotor processing and the motivation and information processing interface, respectively.

Figure 2. The MVP model developed by Keller (2008)

Kuhl (1985) defined volition as a mediating factor that “energizes the maintenance and enactment of intended actions.” Gollwitzer (1996) investigated implementation intention and determined that it comprises the following three phases: (a) motivation (pre-decisional phase), (b) the beginning of volition (pre-actional phase), and (c) the implementation of volition (actional phase). Finally, the post-actional phase involves evaluating whether further goal pursuit is necessary and worthwhile, and then this evaluation provides feedback to motivation. In summary, volition can be considered a mediating factor of motivation, and the post-action phase is similar to the cognitive evaluation
proposed in the ARCS model by Keller (1987). In addition, Astleitner and Wiesner (2004) indicated that in a multimedia environment, the instructional strategies proposed using the ARCS method can support self-regulated (Zimmerman, 1998) learners in various phases. However, multimedia elements can enhance the level of reality in the learning environment.

The limitation of the ARCS model developed by Keller (1987) is that it cannot describe how information processing elements are integrated into the learning process and how motivation interacts with information processing elements. Information processing theory and its applications (Mayer, 2001) also fail to consider motivational factors. Based on studies conducted by Keller (1987) and Mayer (2001), Astleitner and Wiesner (2004) proposed a model that integrates information processing and motivation, including motivational processing and elements of mental resource management. This model facilitates investigation of the interaction between cognitive and motivational elements in the learning process. The concept underlying information processing theory is cognitive load (Paas et al., 2003), which refers to the amount of information in working memory that a person is capable of processing. Cognitive load is comparatively more challenging to manage during multimedia-based instruction because various stimuli, such as hot links, Internet excursions, and problems such as “seductive details” and “lost in hyperspace,” may occur in GBLEs (Deimann & Keller, 2006).

Researchers investigating cognitive loads must confirm the motivational influences of instructional conditions and verify strategies for maintaining student focus when using learning materials (Paas, Tuovinen, van Merrienboer, & Darabi, 2005). The MVP model in Fig. 2 shows that motivation and cognitive load are variables that influence learning and performance.

Research purposes

Based on the preceding discussion, this study aimed to address the following questions:

- What DGBL design method can be employed to apply various game characteristics that match the motivational strategy procedures?
- Do empirical relationships exist between motivation, cognition, and performance in DGBL, as suggested in Keller’s MVP theory (Keller, 2008)?

Methods

Online DGBL (operating a small factory using computer-aided manufacturing (CAM; OSF-CAM)

Computer-aided design (CAD)/CAM has become the primary industrial production type and a leading course for mechanical and industrial design departments at universities. Dankwort, Weidlich, Guenther, and Blaurock (2004) indicated that students should acquire the ability to use CAD and CAM abilities when at university. Employing such techniques and technology facilitates the manufacturing of models designed by students of design education.

OSF-CAM design methods build on the game characteristics and ARCS model (Keller, 1987) to enhance student learning motivation using the systemized instructional strategies shown in Table 2 and the game characteristics employed in various design strategies, as shown in Table 3. Regarding the design procedure, three corresponding ARCS subcategories were first adopted, then matching game characteristics were analyzed (Table 2) and applied to DGBL design (Table 3). Finally, multimedia research effects (Mayer, 2001) were integrated. This process can reduce extraneous cognitive loads, enable learners to establish schema through assimilation and accommodation, and increase germane cognitive loads.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Subcategories</th>
<th>Game characteristics</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>Perceptual arousal</td>
<td>Multimodal presentation and story, fantasy, and fun</td>
<td>Digital games are designed with effects such as rich pictorial and auditory variations and using multimedia characteristics to induce players’ curiosity. In addition, the use of story contributes to various types of knowledge learning.</td>
</tr>
</tbody>
</table>
Inquiry arousal

Games possess exploratory and experiential factors that correspond with constructive learning, assisting learners in resolving problems. Role-playing prompts participation (Huang & Johnson, 2008). Attempting to solve problems in the game and maintaining the level of conflict, competition, challenge, or opposition can balance the game.

Variability

Including varying plots and tasks can increase the fun and challenge of the gaming process and incite learners’ curiosity through learning contents to sustain their attention, thereby stimulating their motivation for continuous participation.

Familiarity

This learning system adopts the following two strategies to compensate for the insufficiency of familiarity: (a) introduction of product model examples from the completed CAD course, and (b) designing objects in the game using 3D virtual machines for realistic simulation.

Goal orientation

In the game tasks, including tasks that accumulate toward achieving game goals, learners are often required to undertake a series of tasks to complete the final goal of the game (Huang & Johnson, 2008).

Motive matching

Goals give us motivation. Playing with others is fun and helps a player become involved in a community. Losing and winning have strong emotional and ego-gratification influences (Prensky, 2007).

Learning requirements

Specific game rules and limitations show students criteria for achieving certain goals (Garris et al., 2002).

Success opportunities

Providing various learning tasks with adequate difficulty levels while maintaining the challenge and competition of the game can arouse users’ competitive intentions and guarantee success for everyone. DGBL can attain individualized learning goals (Deubel, 2006).

Personal control

Digital games constitute interactions between humans and machines, presenting players with a sense of performing actions. Play provides a strong and passionate sense of participation. Game actions also enhance a part of the learner’s sensory control (Malone & Lepper, 1987).

Natural consequences

When learners successfully complete tasks and accumulate scores, the sense of winning naturally induces a sense of achievement.

Positive consequences

Designs where adequate feedback is obtained immediately following a key operation or decision can strengthen active interaction.

Equity

Users must play the game using the same rules, and their scores for completing tasks must be identical.

Table 3. Game characteristics applied in various design strategies

<table>
<thead>
<tr>
<th>Design strategies</th>
<th>Description</th>
<th>Game characteristics</th>
<th>Screen shots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story-based gaming processes and goals</td>
<td>In the game, users act as a factory operator. Accumulating profits using the production model is the user’s goal.</td>
<td>Goals, play, fun, and multimodal presentation and story</td>
<td><img src="image" alt="Screen shots" /></td>
</tr>
<tr>
<td>Role-playing</td>
<td>Users can select the character they desire in the beginning of the game, and generating a sense of identification with the character they like.</td>
<td>Multimodal presentation and story, curiosity, and role-playing</td>
<td></td>
</tr>
<tr>
<td>Learning task design</td>
<td>Game goals are presented using five tasks. The model style and difficulty required for each task varies. The virtual capital obtained when the model is completed increases with the difficulty level.</td>
<td>Problem solving, challenge, rules, and tasks</td>
<td></td>
</tr>
<tr>
<td>Score accumulation design</td>
<td>Whenever a player completes a task, the player obtains profit corresponding to the task’s difficulty. However, if an incorrect operation occurs during the completion of a task, the obtained profits decrease.</td>
<td>Competition, win state</td>
<td></td>
</tr>
<tr>
<td>Real-time display design</td>
<td>Our system applied health points to show real-time profits. The reward or monetary amount obtained in each task is directly proportional to the number of correct operations.</td>
<td>Outcomes and feedback, challenges, and competition</td>
<td></td>
</tr>
<tr>
<td>Time-limited game design</td>
<td>In the “aiming for the workpiece original point” phase, a timed game was designed to stimulate a sense of fun. The player operates the CNC platform to aim for the workpiece original within the time limits. Before playing the game, the player can select the difficulty level. Although harder levels possess shorter time limits, the profits for completing the game are more heavily weighted, thereby inducing the win state.</td>
<td>Control, challenges, outcomes and feedback, win state, adaptively, interactive, play, and fun</td>
<td></td>
</tr>
<tr>
<td>Online updating design</td>
<td>The server side performs online updates, such as the percentage of correctly manufactured models and the number of completed models. In addition, the system shows the top five players that have earned the most capital, creating competition and interaction between players.</td>
<td>Outcomes and feedback, competition, win state, and interaction</td>
<td></td>
</tr>
<tr>
<td>Humorous dialogue</td>
<td>Appropriate and timely feedback is presented with the character’s humorous dialogue.</td>
<td>Fun, role-playing</td>
<td></td>
</tr>
<tr>
<td>Exaggerated design</td>
<td>Feedback from the operating outcomes for height correction during the work piece original point phase is presented using sound effect and visual effect in dual models.</td>
<td>Fantasy, outcomes and feedback, and control</td>
<td></td>
</tr>
</tbody>
</table>
The OSF-CAM game used in this study included a CAM game-based module (Fig. 3) and a CAM multimedia module (Fig. 4; Figs. 3 to 5 are screenshots). The CAM multimedia module enables users to understand background knowledge of CAM and its operating procedures, such as the effects of cheats, which can enhance game-based learning (Royle, 2008). OSF-CAM was designed according to multimedia cognitive load theory (Fig. 1) proposed by Mayer (2001) and the game characteristics applied in Keller’s ARCS model (Tables 2 and 3). Strategies for reducing learners’ cognitive loads include pretraining, signaling, segmenting, aligning, synchronizing, and weeding. For example (as shown in Fig. 5), the planning of machining cutting tools is performed using shot representation to increase the dynamic effects and sense of presence during operations in the predesigned interface. First, the lens was rotated left and zoomed in to present the complete installation of the first cutting tool (Step 1). Subsequently, Steps 2 (rotated right and zoomed out) and 3 (complete installation of tools) were performed. According to the principles of synchronizing and weeding, the designed operating interface was modified using a single perspective and appropriate zooming out techniques (to streamline the steps). The complete dynamic process of planning and installing all tools was fully demonstrated from this perspective, excluding the processing of presentational holdings. This operational practice enables learners to determine how to perform tool planning correctly, including the suitable sizes and types of cutting tools to select and the appropriate arrangements of the positions and sequences of the tools. Because the CAM course content involves complex learning, we introduced pretraining to help students understand the relevant background knowledge and technical terms before using the game-based learning module. In addition, bite-sized segments were adopted as the principle for presentation design (Mayer, 2001; Squire, 2005).

![Figure 3](image-url)  
*Figure 3. CAM game-based module: operating a small factory, (a) portal page of the online game, (b) player appearance and relevant information*
Users log into the game online through the webpage browser. Learners can examine the CAM multimedia module (Fig. 4) to learn the background knowledge and operating procedures relevant to CAM. Next, learners enter the
CAM game-based module, as shown in Fig. 3. When an authorized user logs into the game, the server mode is activated (this updates the use contexts and online scores and shows the score rankings). Learners can begin to play the game by manufacturing process simulation of easy or complex models, depending on their personal abilities.

**Data collection**

CAM course instruction was provided for 3 h every week. The participants were second-year university students of the Department of Art and Design. The study researcher lectured on course-related matters for 8 weeks and provided the OSF-CAM to guide learning. Student responses were measured after completing the game training. At the end of the course, 63 sets of valid responses (from 48 females and 15 males, aged between 19 and 21 years, mean = 20.12, standard deviation (SD) = 0.47) were employed for data analysis.

**Motivation measurement**

Measurements of learning motivation were based on the IMMS and comprised the subscales of attention, relevance, confidence, and satisfaction. A symmetric 9-point Likert scale containing 36 items was used, with 12 items for attention, 9 items for relevance, 9 items for confidence, and 6 items for satisfaction (Keller, 2010, pp. 283-284). Minor modifications were incorporated to accommodate the DGBL setting (i.e., replacing “this lesson/material” with “OSF-CAM”). The Cronbach’s α for each subscale was .852, .819, .818, and .801, respectively, with 63 samples, and that for the overall survey was .914. All of the subscale Cronbach’s α values exceeded .8, confirming the high reliability of the tests used in this study.

**Measurement of cognitive loads**

The sum of extraneous and germane cognitive loads is assumed to be equal to the total cognitive load deducted from the intrinsic cognitive load. Because intrinsic cognitive loads cannot be manipulated using instructional intervention approaches, the primary purpose of instructional designs is to construct the optimal combination of extraneous and germane cognitive loads (Huang, 2011). Inappropriate instructional designs generally generate relatively higher extraneous loads but, nevertheless, leave sufficient room for learners to individually decide how much germane load they want to invest (Cierniak, Scheiter, & Gerjets, 2009). The germane cognitive load is the effective cognitive load, which demonstrates the efforts that learners invest in learning (Kalyuga, 2009).

Salomon (1984) reported that students’ perceived self-efficacy regarding learning material is significantly and positively correlated to their mental efforts and learning achievements. Because increased mental effort can enhance learning outcomes, mental effort appears to be correlated with germane cognitive load during the learning process. Both subjective and physiological measurements are non-interference measurements; however, mental effort measurements yield comparatively higher reliability and sensibility (Paas et al., 2003). Therefore, subjective mental effort measurements can be used as the main indicator for measuring learners’ overall cognitive loads and evaluating their differing cognitive loads. In addition, concentration is considered active content, and can replace neutral “mental effort” learning conditions. Concentration in the learning process reflects the student’s attention, is a cognitive ability required for various processes relevant to learning, and represents the schema for success (Cierniak et al., 2009). Consequently, the germane load scale is represented using the item “how much did you concentrate during learning?” Similar to the methods adopted by Ayres (2006) and Cierniak et al. (2009), the intrinsic load scale involves obtaining learners’ self-evaluations regarding the perceived difficulty of the learning content by using the item “how difficult was the learning content for you?” Learning content with high difficulty may lead to low learning outcomes, and relatively higher concentration during the learning process may generate superior learning outcomes, which indicates a possible learning performance relationship. In this study, the use of germane and intrinsic load scales involved adopting symmetric 9-point Likert scales.
Performance measurements

The CAM cognitive scale

This study used six cognitive levels, specifically knowledge, comprehension, application, analysis, synthesis, and evaluation (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956), as the scope of the cognitive domain, and established a two-way specification table with which the pretest questionnaire items were determined. Subsequently, three CAM experts reviewed the items before completing the pretest questionnaire (35 items). We distributed the pretest questionnaires to 39 third-year university students of the Department of Art and Design who had completed a course in CAM. The item difficulties ranged from .2 to .8 with a discrimination of larger than .25 set as the criterion for selecting items. Subsequently, 24 items with a Kuder-Richardson reliability of .817 were retained. The completed scale is presented as Appendix A.

The CAM skill scale

The CAM skill scale was devised according to the course content and learning goals, and comprised items for (a) programming: planning suitable process types and cutting tools (including selecting suitable sizes and types of tools, and arranging the sequences of tools) and setting machining parameters (cutting-in amount, cutting interval, feed rate, etc.); (b) machining: setting up workpieces, machine settings, and machine operations. Three CAM experts reviewed the scale, and 5 third-year university students who had completed a CAM course implemented the model. The average time required to complete the model was 121.2 min (the working time included the programming and machining time), and the scale is presented as Appendix B.

Results and discussion

Learning motivation

Table 4 lists the students’ responses regarding the ARCS motivation scale, for which the mean of the relevance subscale was the highest (6.37) and that of the attention subscale was the lowest (5.77). The means of the confidence and satisfaction subscales were 5.90 and 6.05, respectively. In addition, all of the motivational subscales attained standard deviations ranging from .67 to .78. The mean of overall learning motivation was 6.02. Although game characteristics, such as fun, fantasy, curiosity, and role-playing can attract learners’ attention, they are not necessarily directly relevant to learning. Therefore, Royle (2008) indicated that although actual learning occurs in the game, a certain portion of the game content must be external to the course to maintain game value. Digital games that contain multimedia features can be designed to attract players’ attention. However, Mayer (2001) contended that including appealing animation, which constitutes external materials, can increase learners’ cognitive loads. Consequently, this study suggests using the weeding strategy to resolve this problem, which can lead to the coherence effect. When extraneous but interesting materials are excluded, students can interpret multimedia-based explanations more clearly. However, reducing the number of extraneous but interesting materials weakens learners’ attention. This may be why attention received the lowest points among the four motivation subscales of the developed game. However, the attention value exceeded the median, primarily because the developed game retained the fun and curiosity designs attracting learners’ attention to a considerable degree. Previous studies have indicated that certain forms of cognitive load are desirable for prompting and enriching the challenge of a game (Ang, Zaphiris, & Mahmood, 2007).

Regarding the lack of corresponding game characteristics for the relevance subscale, this game involved adopting the following two design strategies (as shown in Table 3): (a) familiarity with product model examples, and (b) realistic simulation. Multimedia elements can enhance the degree of reality in learning environments to increase the familiarity of the learning environment (Astleitner & Wiesner, 2004). Furthermore, the relevance of learning environments can be increased by integrating students’ previous experiences. This may be the reason that the mean of the relevance subscale attained the highest score.

The mean of the confidence subscale (5.90) was only slightly higher than the value of the attention subscale. The reason may be, as indicated by Squire (2005), conventional school-based professional knowledge acquisition cannot be realized by successful students in a game-based learning environment. Such students do not believe that game-
based learning can benefit their performance in college entrance examinations or university classrooms. The mean of the overall motivation scale was 6.02, indicating that OSF-CAM stimulates students’ learning motivation.

<table>
<thead>
<tr>
<th>Item</th>
<th>N</th>
<th>Mean</th>
<th>SD/(CV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>63</td>
<td>5.77</td>
<td>.67</td>
</tr>
<tr>
<td>Relevance</td>
<td>63</td>
<td>6.37</td>
<td>.70</td>
</tr>
<tr>
<td>Confidence</td>
<td>63</td>
<td>5.90</td>
<td>.78</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>63</td>
<td>6.05</td>
<td>.72</td>
</tr>
<tr>
<td>Average of total ARCS</td>
<td>63</td>
<td>6.02</td>
<td>.55</td>
</tr>
<tr>
<td>Germane cognitive load</td>
<td>63</td>
<td>5.98</td>
<td>1.33</td>
</tr>
<tr>
<td>Intrinsic cognitive load</td>
<td>63</td>
<td>5.51</td>
<td>1.73</td>
</tr>
<tr>
<td>Cognition of CAM</td>
<td>63</td>
<td>60.98</td>
<td>15.58/(.26)</td>
</tr>
<tr>
<td>Working time (Completion times of operation)</td>
<td>63</td>
<td>111.84 (min.)</td>
<td>27.04/(.24)</td>
</tr>
</tbody>
</table>

Cognitive load

In the students’ self-evaluated reports of cognitive load (shown in Table 4), the mean/SD of the germane and intrinsic cognitive loads were 5.98/1.33 and 5.51/1.73, respectively. The intrinsic cognitive load reported by the learners suggests that the CAM course is a learning domain with a slight to moderate level of difficulty for students, whereas the SD (1.73) shows that learners possess individual differences. The mean of the germane cognitive load was 5.98. Germane cognitive load increases learning effectiveness and creates relatively more profound learning experiences (Kalyuga, 2009), indicating that the proposed game-based learning design positively affects learning enhancement. However, the results were not sufficient in showing that the learners had learning experiences that were more significant in higher-order cognition and skills. Therefore, although games are an effective learning tool (Papastergiou, 2009), multiple researchers have contended that game-based learning is an instructional approach that complements only conventional methods (Royle, 2008).

Performance

The mean and coefficient of variation (CV) of students’ cognitive performance were 60.98 and .26, respectively. The working time to complete model manufacturing represents the student’s skill performance. The mean and CV of the student’s skill performance were 111.84 min and .24, respectively. CV is independent of the unit. To compare data sets with differing units, the CV should be employed instead of the SD. Therefore, the CVs of cognitive performance and skill performance reflect the similarity in the degree of group dispersion. These performances are shown in Table 4.

Canonical correlation analysis of motivation, cognitive load, and performance

Table 5 lists the canonical correlation test results. In the DGBL of this study, motivation and cognitive load exhibited a significant canonical correlation with performance (Dimension 1, \( p < .05 \)). Dimension 1 showed that the two sets of variables possessed a canonical correlation of \( .446 \), and that learning motivation and germane cognitive load positively influenced skill and cognitive performance (i.e., substantial learning motivation and germane cognitive load result in comparatively shorter working times and greater cognitive achievements). Intrinsic cognitive load negatively influenced cognitive performance and had a directly proportional influence on the working time. Table 6 lists the canonical correlation analysis results of this study; the first canonical variable \( \chi_1 \) in Set X (canonical loadings \(-.446, -.263, \text{ and } .776\)) explains 39.8% of the variance in motivation, germane cognitive load, and intrinsic cognitive load. The first canonical variable \( \eta_1 \) of Set Y (canonical loadings: \(-.638, .505\)) explains 76% of the variance in cognitive performance and skill performance. The independent variables of motivation, germane cognitive load, and intrinsic cognitive load in Set X can explain 15.1% of the variance in the dependent variables of cognitive performance and the skill learning performance in Set Y. A path diagram of the canonical correlation is
shown in Fig. 6, in which motivation and germane cognitive load positively influenced skill and cognitive performance. This means that substantial motivation and germane cognitive load result in comparatively shorter working times (i.e., superior skill performance) and higher cognitive scores. Intrinsic cognitive load negatively influenced cognitive performance and had a directly proportional influence on the working time. This study verified that motivation and cognitive load are latent factors that influence performance in DGBL, as suggested in Keller’s MVP theory.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Canonical correlation</th>
<th>Wilk's</th>
<th>Chi-sq</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.446</td>
<td>.796</td>
<td>13,432</td>
<td>6</td>
<td>.037</td>
</tr>
<tr>
<td>2</td>
<td>.080</td>
<td>.994</td>
<td>.377</td>
<td>2</td>
<td>.828</td>
</tr>
</tbody>
</table>

Table 6. Canonical correlations

<table>
<thead>
<tr>
<th>Canonical loadings $x_1/\eta_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent variables (set X)</td>
</tr>
<tr>
<td>Motivation</td>
</tr>
<tr>
<td>Germane cognitive load</td>
</tr>
<tr>
<td>Intrinsic cognitive load</td>
</tr>
<tr>
<td>Variance extracted</td>
</tr>
<tr>
<td>Redundancy index</td>
</tr>
<tr>
<td>Dependent variables (set Y)</td>
</tr>
<tr>
<td>Cognitive performance</td>
</tr>
<tr>
<td>Working time (skill performance)</td>
</tr>
<tr>
<td>Variance extracted</td>
</tr>
<tr>
<td>Redundancy index</td>
</tr>
</tbody>
</table>

Figure 6. Path diagram of the canonical correlation

Conclusions

OSF-CAM digital learning games were completed by applying game characteristics that matched the ARCS motivational strategies and integrating multimedia research effects (Mayer, 2001). The results verified that CAM skill and cognitive learning can cultivate learner motivation in the aspects of attention, relevance, confidence, and satisfaction, and can increase germane cognitive loads. In addition, a significant relationship exists in the first dimension of the canonical correlation analysis for motivation, germane cognitive load, and intrinsic cognitive load related to cognitive performance and skill performance. Motivation and the germane cognitive load are proportional to cognitive performance and skill performance (shorter working times), whereas the intrinsic cognitive load is inversely proportional.

In summary, this study proposes a systematic DGBL design method that applies game characteristics that match the ARCS motivation strategy and integrates Mayer’s multimedia research effects. Based on the research results, when designing DGBL, designers should increase motivation and germane cognitive load to enhance learning effectiveness. In addition, the attention component must be compromised; however, relevance, satisfaction, and confidence do not conflict. For example, nonlinear presentations of textual information may increase learners’ cognitive load (Zumbach & Mohraz, 2008), and animated instructional messages may necessitate additional cognitive processing capacity from learners (Ayres & Paas, 2007). Both nonlinear presentations of textual information and animated instructional messages commonly feature in GBLEs because they attract learners’
attention and enhance their motivation; however, they also present ineffective cognitive loads. Furthermore, the experiment results confirmed that learners’ learning motivation, cognitive load, and performance possess latent correlations in DGBL, as suggested in Keller’s MVP theory. This is a case study for exploratory purposes. Future research with a larger sample size might be able to identify significant correlations in DGBL.

**Acknowledgments**

This work was supported in part by the National Science Council of Taiwan, under Grant NSC 101-2511-S-152 -013.

**References**


### Appendix A. The CAM cognitive scale

<table>
<thead>
<tr>
<th>items</th>
<th>content</th>
<th>categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The feed rate is ① feed amount per revolution × revolutions per minute (rpm) ② feed amount per revolution ÷ rpm ③ cutting depth × feed amount per revolution ④ cutting depth ÷ feed amount per revolution.</td>
<td>C2</td>
</tr>
<tr>
<td>2</td>
<td>If the milling depth increases, the axial rpm should ① increase ② decrease ③ remain unchanged ④ increase proportionally.</td>
<td>C3</td>
</tr>
<tr>
<td>3</td>
<td>The common unit for the feed rate of a CNC mill is ① mm/min ② mm/rev ③ rps ④ rpm.</td>
<td>C2</td>
</tr>
<tr>
<td>4</td>
<td>Generally, the 3D modeling file format that can be loaded into computer-aided manufacturing (CAM) software for tool path planning is ① .igs ② .stl ③ both of the above ④ none of the above.</td>
<td>C3</td>
</tr>
<tr>
<td>5</td>
<td>During general roughing, the machining parameters must set the cutting-in amount, tool path cutting interval, and ① finish margin ② tool length ③ machining time ④ all of the above.</td>
<td>C2</td>
</tr>
<tr>
<td>6</td>
<td>The cutting speed of the milling should be reduced under the condition of ① finishing ② the blade being worn within an acceptable range ③ disregarding the life of the milling cutter ④ softer workpieces.</td>
<td>C6</td>
</tr>
<tr>
<td>7</td>
<td>When milling using an end mill, which of the following is not the reason for abnormal vibrations? ① overly long shank ② overly short shank ③ insufficient rigidity of the mill shank ④ overly thin shank.</td>
<td>C6</td>
</tr>
<tr>
<td>8</td>
<td>When milling, which of the following is not the reason for cutting vibrations? ① excessive milling depth ② unfixed workpiece ③ tool wear ④ insufficient feed.</td>
<td>C6</td>
</tr>
<tr>
<td>9</td>
<td>When milling workpieces using an end mill 50 mm in diameter at a milling speed of 30 m/min, the axial revolution speed should be ① 150 ② 170 ③ 190 ④ 210 rpm/min.</td>
<td>C4</td>
</tr>
<tr>
<td>10</td>
<td>Which of the following end milling conditions is wrong? ① The tool rotations decrease as the diameter increases ② The tool rotations increase as the diameter increases ③ The tool diameter increases with increased cutting depth ④ The tool rotations increase when the roughing changes to finishing.</td>
<td>C3</td>
</tr>
<tr>
<td>11</td>
<td>When clamping a workpiece using a mill clamp, we found that the workpiece moves upward and cannot adhere closely to the parallel block. Which of the following is improper? ① loosening the mill clamp and hammering the workpiece downward using a soft hammer ② clamping the workpiece using pressure wedges ③ hammering the workpiece downward using a sledge hammer ④ adjusting the chute clearance of the mill clamp moving jaw.</td>
<td>C6</td>
</tr>
<tr>
<td>12</td>
<td>When should the tool length be corrected again? ① after completing the model machining ② when resuming suspended machining ③ when replacing damaged cutting tools or maintaining the tools ④ all of the above.</td>
<td>C6</td>
</tr>
<tr>
<td>13</td>
<td>Regarding the CNC milling operation of a Roland MDX650, when implementing workpiece original point settings for a workpiece, what approach should be adopted to move cutting tools closer to the workpiece when they are already approaching the workpiece? ① moving button on the X and Y axes ② up and down moving button on the Z axis ③ hand wheel ④ all of the above.</td>
<td>C3</td>
</tr>
<tr>
<td>14</td>
<td>When performing the workpiece original point settings, the original point is ① set individually according to the model characteristic ② must be consistent with the workpiece original point set in the tool path of CAM software ③ can be set randomly ④ none of the above.</td>
<td>C2</td>
</tr>
<tr>
<td>15</td>
<td>To temporarily halt tool movement during program operations, we should press ① the emergency stop switch ② the reset button ③ the pause switch ④ the cutting feed rate adjustment button, adjusting to 0%.</td>
<td>C1</td>
</tr>
<tr>
<td>16</td>
<td>During milling, when the diameter of a cutting tool is small, the axial rpm should be ① higher ② lower ③ unchanged ④ variable.</td>
<td>C3</td>
</tr>
<tr>
<td>17</td>
<td>If the cutting speed is 75 m/min, with a milling cutter diameter of 80 mm, the milling cutter rpm is ① 258 rpm ② 298 rpm ③ 358 rpm ④ 398 rpm.</td>
<td>C4</td>
</tr>
</tbody>
</table>
The common ψ10 end mill is not suitable for milling ① R4 rounds ② a level of 11 mm ③ groove widths of 12 mm ④ R4 fillets. C5

Which of the following tool or cutting tool materials possesses a higher toughness? ① high speed steel ② tungsten carbide ③ ceramic ④ diamond. C1

Take 19 mm from the main scale of the vernier caliper and divide it into 20 intervals on the vernier scale. The minimum reading of the vernier caliper is ① 0.01 ② 0.02 ③ 0.05 ④ 0.10 mm. C4

The graph above shows the name of each part of a vernier caliper. The name of the ( ) is ① sliding scale ② vernier scale ③ caliper ④ moving scale C1

The parallel blocks used for the mill clamp generally comprise ① ② ③ ④ blocks in each set. C1

When determining the margin of workpieces using a piece of thin oily paper, feeding is terminated immediately after the piece of thin paper is scratched off by the milling cutter. Which of the following actions should be conducted as the next priority? ① return this axis scale to zero ② eject the milling cutter from the workpiece ③ add cutting fluid to prepare for milling ④ cut off the power. C5

When determining the workpiece margin using a thin paper 0.07 mm in thickness, if the drill bit diameter is 5 mm and the distance between the intended hole center and the margin is 20 mm, the table moving distance should be ① 22.57 ② 25.07 ③ 17.57 ④ 15.07 mm. C5

C1/ knowledge, C2/ comprehension, C3/ application, C4/ analysis, C5/ synthesis, and C6/ evaluation

Appendix B. The CAM skill scale
Scripting For Collaborative Search Computer–Supported Classroom Activities

Renato Verdugo¹, Leonardo Barros¹, Daniela Albornoz¹, Miguel Nussbaum¹ and Angela McFarlane²

¹Department of Computer Science, School of Engineering, Pontificia Universidad Católica de Chile // ²Kew, Royal Botanic Gardens, United Kingdom // rverdugo@ing.puc.cl // libarros@puc.cl // dralborn@puc.cl // mn@ing.puc.cl // A.McFarlane@kew.org

*Corresponding author

(Submitted August 23, 2013; Accepted November 27, 2013)

ABSTRACT

Searching online is one of the most powerful resources today’s students have for accessing information. Searching in groups is a daily practice across multiple contexts; however, the tools we use for searching online do not enable collaborative practices and traditional search models consider a single user navigating online in solitary. This paper presents a three level conceptual model, called the Collaborative Search Procedural Model, which enables the implementation of collaborative search classroom activities based on multi-user collaborative search scripts. A software solution, CollSearch, which follows the Collaborative Search Procedural Model and offers a unified tool to enable collaborative searching computer-supported classroom activities, is also presented. Empirical evaluation of the tool with high school students as part of an English as a second language course shows that students’ outcomes improve when compared to non-scripted group search. Results show that by following the Collaborative Search Procedural Model students better appropriate the work they build together with their group. The OECD has highlighted the importance of collaborative work by the fact that PISA 2015 will assess collaborative problem solving; collaborative search is a fertile field for fostering better group work interactions. This paper shows that new tools that enable collaborative work dynamics in searching for information must be developed in order to address the educational challenges that today’s students are facing.

Keywords
Computers and education, Collaborative computing, Collaborative learning, Web search

Introduction

Collaboration in the search task

one of the most common tasks carried out on the Internet is searching for information. Under a traditional model, this task is conceived as executed by a single user, omitting interaction with other people (Twidale et al., 1997). Because of this, online search tools –search engines plus the browsers used to access them- are mostly designed for users to search individually (Broder, 2002). Education, work or social interaction are some of the many scenarios under which collaboration in the search process happens in our daily lives (Amershi & Morris, 2009); however search engines and browsers cannot handle this characteristic, forcing users to turn to complimentary methods and tools. Examples of this are the use of a single computer, with one user leading the search, and another looking over his shoulder, sharing search results via email, or coordinating joint searches through instant messaging systems (Morris, 2007). The distance between the functionality offered by the technologies we use, and the practices of users searching online for information translates into processes where collaboration is not only not supported but also discouraged by the many obstacles users face when attempting to search online together.

Group information searching and the lack of tools to support it have been studied in recent years. (Amershi & Morris, 2008) identified a series of limitations that emerge when users search online for information together without tools that have been specially designed to promote collaboration. Some of these limitations are:

- **Difficulties contributing.** There are multiple scenarios under which current search tools foster an environment where group members asymmetrically contribute to the search task.
- **Lack of awareness.** Dominating group members minimize the contribution of others, reducing the awareness of their ideas and suggestions.
- **Lack of hands-on learning.** Group members who do not have access to a shared computer’s input devices loose the opportunity to gain expertise interacting with search technologies.
- **Information Loss.** Multiple difficulties emerge when groups try to keep track of their findings.
These problems configure an initial research question; how can the search process be scripted to foster collaborative practices between users when they search online together? A script structures the interaction between individuals and determines collaboration and problem solving logistics (Nussbaum et al., 2009), as well as offering detailed sets of instructions for each part of the activity being conducted (O’Donnel & Dansereau, 1992). Scripting for collaborative searching requires us to distinguish between cooperation and collaboration, as the difference between them is often unclear. Cooperation refers primarily to the division of tasks within a group, where each member is responsible for his own actions, while collaboration is defined as the coordinated work of a group of individuals to solve a common problem together, where all members are responsible for the end result (Roschelle & Teasley, 1995; Dillenbourg, 1999). Cooperation is similar to what factory workers do in an automobile assembly line, where each worker is responsible for carrying out a specific action, and is only worried about completing said action successfully. Collaboration, on the other hand, can be compared to putting together a puzzle, where everyone helps on any part of the puzzle and is responsible for a misplaced piece (Szewki et al., 2011). When applying these concepts to group information search, it is possible to see how cooperative searching is actually a union of individual searches, and doesn’t necessarily offer a technological challenge to the tools we use today. On the other hand, collaborative searching faces us with the need to restructure our practices and tools to include the possibility for users to search and build solutions and answers to questions together.

The success of collaborative dynamics in-group work depends on six criteria that have been established in several studies, and were summed up in the work of E. Szewkis (Szewkis et al., 2011). These collaboration criteria are:

- **Common goal**: a common objective, shared by all the members of the group (Dillenbourg, 1999).
- **Positive interdependence**: correlation between peers’ work, so that the success of each member depends on the work of his teammates (Johnson & Johnson, 1999).
- **Coordination and communication**: interactions must occur in the right order and at the right time, avoiding the loss of communication and cooperation efforts (Raposo et al., 2001, Gutwin & Greenberg, 2004).
- **Individual accountability**: each member of the team is responsible before his teammates for the actions he carries out and their consequences (Johnson & Johnson, 1999).
- **Awareness**: each member of the group can obtain information about the state that the work is in regarding both the group work and his teammates’ individual work (Zurita & Nussbaum, 2004).
- **Joint rewards**: depending on the results of their work, the entire team receives the same evaluation, whether it is a reward or punishment (Zagal et al., 2006).

Additionally, collaborative search presents the challenge of division of labor (the way that the members of a group distribute the workload between themselves). Attempting to reduce unnecessary redundancy users work with parallel search patterns; this can only be achieved successfully through high coordination and awareness of teamwork (Morris, 2008). Faced with this task, strategies that distribute the workload enabling parallel work without affecting awareness of what others are doing must be found.

All of this adds up to the need to develop new search models that support collaborative behaviour among users. Our initial research question can be rephrased as, how can the search process be scripted to help align a group of users under a common goal, augment their positive interdependence, contribute to better coordination and communication, reinforce individual accountability of each member of the group, help them achieve better awareness of their teammates’ work and strengthen the division of labor between them when searching online together? The reformulation of the research question purposefully leaves out the criteria of joint rewards because this is external to the actual search process; the reward for a successful or unsuccessful search session depends on the context where this is being carried out and therefore does not correspond to the organic structure of the task of searching online for information.

**Collaborative learning and collaborative search**

In recent years, collaborative learning (CL) environments have gained importance and notoriety. By collaborating with their peers, students develop important communication and social skills as they learn to carry out multidirectional dialogues and submit their ideas to their classmates’ critical analysis (Nussbaum et al., 2009). Collaborative learning allows the members of a group to articulate their points of view and negotiate and exchange ideas; learning is achieved through a process of building knowledge (Infante et al., 2010; Zurita et al., 2005) where students interact with the source of information, their peers, and the teacher.
The Internet, with its growing availability in schools, is shaping up to be the main source of information for students. A large part of the information searches that are carried out within schools are based on group interactions (Large et al., 2002), which is why searching for information has the potential to become a powerful collaborative learning activity. Collaborative search allows students to share not only the results or final products of information searching, but also the process that led to those results (Twidale et al., 1997). Cooperative searching, by merely distributing the workload, encourages bad practices where students use only a fragment of knowledge that they later copy-and-paste to form a greater project; contrarily, collaborative search moves students to work together to build, as a group, the knowledge they need. Collaborative learning environments provide a fertile field where the previously stated research question gains practical applicability; how do collaborative search activities—those follow specific scripting—change the way students work together in groups when searching for information?

Effective collaborative search activities require proposing search models and activity scripting where the user is no longer viewed as an isolated individual, but as an active member of a group. The purpose of this paper is to present a conceptual model of the process of collaborative search and the scripting it requires to be used as a structure for teaching activities that revolve around group search for information and the collaborative building of knowledge. The following section presents a model for collaborative search that is articulated through three levels: a high level understanding regarding the search process that each user faces (abstract model), a general structure for collaborative search activities within the classroom (Macro-Script), and the specific ways in which Macro-Scripts are adapted to specific implementations (Micro-Script). Later we present the CollSearch tool, specially designed computer software that follows the Macro-Script we propose. We also present an empirical study conducted with high school students that evaluate students’ outcome when working with the CollSearch tool. The final section presents conclusions and future work opportunities regarding collaborative search as a collaborative learning tool.

Scripting for collaborative searching

The complexity of developing multiuser search models that promote collaboration among the members of a team is possibly one of the reasons that explain the lack of computational tools that help to carry out this task. In order to better understand the challenge, the problem can be divided into two levels: the user’s experience when faced with the task of searching (abstract model) and the sequence of steps or stages that a collaborative search activity follows (script). Scripting can be conceived on two levels: on one hand, we have the general structure of the activities (Macro-Script), and on the other we have the concrete steps that must be followed during a collaborative search activity within the classroom (Micro-Script). A Micro-Script is an instruction manual for the teacher and the student that adapts the Macro-Script to the specific subject and context in which the activity is being carried out (Dillenbourg & Tchounikine, 2007). In order to illustrate the difference between both scripts, we can picture a game of chess. The Macro-Script contains the rules of the game that determine the existence of two players that face each other, the goal for each player, the distribution of the pieces on the board, the movements allowed for each piece, the structure of turn-taking in the game, etc. Then a Micro-Script determines the implementation of the Macro-Script. This Micro-Script changes according to the type of implementation: for example, the game can be played with a physical board and pieces, by letter, through a computer simulator, etc. While the Micro-Script can change depending on the support tools that are used and the context of the game, the Macro-Script remains constant because each Micro-Script is an implementation of a game of chess. This allows us to see the logic of the activities and their specific implementations, separately.

Figure 1. Collaborative search model in three levels
The following model is articulated on the three levels described above (Figure 1). It uses Kulthau’s Information Search Process as an abstract model of the stages a user undergoes when facing a search task (Kuhlthau, 2010), it defines a procedural model for collaborative search and, proposes the issues that are relevant when transforming both the abstract model and the Macro-Script to a concrete implementation within the classroom, or Micro-Script.

Abstract model

Kulthau’s Information Search Process (ISP) models the user’s experience during the search task (Kuhlthau, 2010); it was initially developed in the 1980s and has been refined and updated since then, becoming a highly influential model (Cronin & Mehno, 2006) that has been used across multiple settings including educational applications through guided inquiry (Kuhlthau et al., 2007). Despite how much the information environment has changed since the first design of the model, it remains a valid way “for describing information behavior in tasks that require knowledge construction” (Kuhlthau et al., 2008). The model offers a general understanding of the different stages that a person goes through during the process of searching for information (Figure 2). Despite the fact that Kuhlthau’s ISP is based on the perspective of a single person searching for information, there are studies that show that it would be possible – without major modifications- to apply this model to a group search (Hyldegard, 2009).

Macro-script

In the case of collaborative search activities within a collaborative learning context, we propose a Macro-Script called Collaborative Search Procedural Model (CSPM). The model considers the existence of two roles: student and teacher. The student acts as a member of a search team who must collaborate with his peers, search for and contribute new information and document the search process, as well as the results. The teacher is a facilitator of the activity and must monitor each group’s and students’ performance to know when and how intervention is required. This model presents the collaborative search activity scripted as a series of steps where the individual process that each student goes through –described in Kuhlthau’s ISP- is inserted into a group dynamic where the collaborative atmosphere aims to take advantage of the interaction among peers, so as to reduce negative feelings –uncertainty, frustration, etc.- and boost positive feelings –optimism, confidence, etc. The CSPM proposes decreasing interventions from the teacher, as far as guiding the search process, so it is gradually left in the group’s hands as they learn to focus their work. In order to achieve this, the steps that are mainly exploratory are guided by the teacher, while the steps where knowledge is built and documented call for more independent work by the students.

Figure 2. Collaborative search procedural model and its connection to Kuhlthau’s ISP
The CSPM’s Macro-Script proposes four linear stages: (1) Motivation and domain definition, (2) Search term selection, (3) Search and construction –made up of the sub-stages of personal search, personal build, personal discover and describe, and group build- and (4) Group discover and describe. The structure of the CSPM—including stages and sub-stages- is summarized in the diagram presented in Figure 2. Each stage and sub-stage of the CSPM is defined by a high-level procedural description and by collaboration goals determined according to the criteria laid out in the introduction of this paper (Table 1).

**Table 1. Stages of the CSPM**

<table>
<thead>
<tr>
<th>Macro-Script Stage</th>
<th>Description</th>
<th>Collaborative objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation and Domain Definition</td>
<td>The subject of the investigation, as well as its general objectives, reach and focus must be determined with the intervention of the guide (teacher). Additionally, the rules to be followed during the activity are established, such as the way in which the work will be divided, and what platform or medium will be used to build the final hand-in which could be a written document, a video, a presentation, etc.</td>
<td>A common goal must be established for all students. Rules of coordination and communication must be determined. The common rewards that the students will receive must be made explicit.</td>
</tr>
<tr>
<td>Search Term Selection</td>
<td>Lead by the activity guide (teacher) the students suggest key words or queries for the group search, which are shared with all the other students. The facilitator does a brief discussion about query construction and the search tool being used; in-group discussions about the search terms to be used are conducted to help students informally distribute the work.</td>
<td>The collective formation of queries increases each student’s awareness regarding others’ work (Morris et al., 2006). By establishing an initial work distribution students can avoid redundancy and help establish each student’s individual responsibility regarding the activity.</td>
</tr>
<tr>
<td>Personal Search</td>
<td>In this stage, each student searches for information independently –using queries related to the keywords and sub-domains assigned in the previous stage. Each result must be filtered, evaluated, and valued.</td>
<td>Personal searches allow students to feel that part of the work belongs to them, increasing their perception of individual responsibility. By ensuring personal processes, we can avoid a single student taking over the work and ignoring others’ work, which increases positive interdependence. Parallel work contributes to a distribution of the work without duplicates, while at the same time forcing an increase in awareness, coordination and communication.</td>
</tr>
<tr>
<td>Personal Build</td>
<td>The searching users (students) summarize every result that comes up through their web search. They are free to organize and categorize these summaries how they prefer, so they understand the specific contribution that this information makes to the investigation.</td>
<td></td>
</tr>
<tr>
<td>Personal Discover and Describe</td>
<td>The student organizes his personal summary so it is coherent, and builds a macro-structure with the information, following the classification he determined in the personal build stage. In this way, he can articulate his ideas and knowledge before exchanging with his teammates.</td>
<td></td>
</tr>
<tr>
<td>Group Build</td>
<td>Using the results contributed by each member, as well as their respective summaries of key concepts, the information must be reorganized, so as to articulate the entire group’s contribution in a first draft of the answer to the initial question. Group building can happen through the reclassification of group summaries under new criteria defined by the group, or by linking ideas from the summaries, so as to create a “map” of the knowledge that the group has built.</td>
<td>The integration of each member’s contribution increases positive interdependence and forces students to work in a coordinated manner. Ideas presented by teammates increase awareness regarding others’ work. Building a single final presentation that belongs to the entire group –as opposed to belonging to any one student- allows students to better understand why there are joint rewards.</td>
</tr>
<tr>
<td>Group Discover and Describe</td>
<td>The members of the group work together to build a final answer to the question that was the</td>
<td></td>
</tr>
</tbody>
</table>
Describe object of the collaborative search. This answer includes all the angles that were studied individually, but organized in such a way as to allow the group to articulate and transfer the joint knowledge. In this way, the final result belongs to the entire group, and not any individual member. The format of this answer will depend on what was initially proposed by the guide.

Following the collaborative search model in Figure 2, each of the stages of the Macro-Script in Table 1 must be implemented through a specific Micro-Script. Determining these instructions requires a series of design decisions that are specific to the implementation scenarios under which the final deployment will be done; because of this, there is no unique Micro-Script for the model.

CollSearch, a software tool based on the CSPM

The CSPM and the Micro-Script that implements it propose a schoolwork situation where a group of students can investigate a certain subject assigned by their teacher, and then collaboratively build a final hand-in. The first two stages of the CSPM, which are mostly directed by the teacher, present the topic and help students learn about query construction and search techniques. The following three stages are individual research phases, where the student finds information that is relevant to the topic, builds his own point of view and knowledge on the matter and prepares to share his results with his teammates. The final two stages allow students to exchange knowledge –fostering learning among peers- and work together to build a final hand-in containing ideas proposed by all the members of the group.

The implementation of the CSPM in a computational tool requires the elaboration of a Micro-Script that transforms each step of the previously presented Macro-Script into a concrete sequence of activities. It also requires a communication and interaction model that allows the members of each group to interact among themselves and with the system, while the teacher monitors each group’s progress.

Micro-script for CollSearch

Because the CSPM is a Macro-Script, it can be applied to many different scenarios through different Micro-Scripts. CollSearch’s Micro-Script establishes a workflow that enables a group research activity within a classroom environment; each step of the Micro-Script is aimed at promoting collaboration between the group’s members.

Table 2 shows the implementation of each step of the CSPM through specific sequences of activities that compose the Micro-Script, broken down according to the two roles considered by it: teacher and student.

<table>
<thead>
<tr>
<th>CSPM Stage</th>
<th>Specific Micro-Script Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation and Domain definition</td>
<td>TEACHER: The teacher introduces the topic of the investigation, explaining its general objectives, reach, and focus. In the computer system, he defines the parameters of the activity and uploads the class list, in order to randomly form groups.</td>
</tr>
<tr>
<td>Search Terms Selection</td>
<td>The teacher asks the students to suggest search terms that pertain to the topic. He receives the students’ contributions in the computer system and might add more or delete some if necessary. The teacher must then lead a brief discussion about each search term and how they relate to the topic. Additionally, brief explanations about terms that the teacher considered off-topic (and deleted) can also be useful. During this discussion, the teacher must also address query construction techniques and how queries impact the quality of our research.</td>
</tr>
<tr>
<td>Personal Search</td>
<td>The teacher monitors each group’s progress using his computer interface. When he detects the need, or a group calls him, he can physically approach them and provide help or guidance.</td>
</tr>
<tr>
<td>Personal Build</td>
<td>The student can access his bookmarks and make small annotations about the information he has found (Figure 3d). Each note is a summary of the site’s content, and it is stored along with the source’s URL.</td>
</tr>
<tr>
<td>Personal Discover and Describe</td>
<td>Students work as a group to organize the information they collected individually. Each student checks the contents of the Virtual Work Table (Figure 3f) complementing his teammates’ information with his own knowledge, with the possibility of adding comments to a teammate’s notes, checking sources (web references), modifying the contents of the note or the outline. All the notes are reorganized, forming a group outline that expresses the vision and knowledge of all the members of the group.</td>
</tr>
</tbody>
</table>
Communication and interaction model

One of the complexities involved in the collaboration criteria of Coordination and Communication (Gutwin & Greenberg, 2004), and Awareness (Janssen et al., 2004) is the fact that in order to be updated regarding their teammates’ work, students may be constantly interrupted in their own work; this makes it necessary to implement technological support for the CSPM that provides transparent communication mechanisms among the members of the group, i.e., mechanisms that don’t interrupt the student’s work flow. For CollSearch we determined the communication needs that the system had to fulfill in each stage of the CSPM (Table 3).

Table 3. CollSearch communication needs

<table>
<thead>
<tr>
<th>Stage</th>
<th>Communication needs</th>
<th>Technological support (See Figure 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation and Domain Definition</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Search Terms Selection</td>
<td>Coordination between students to determine an initial work distribution.</td>
<td>Group chat, which is always visible.</td>
</tr>
<tr>
<td>Personal Search</td>
<td>Awareness regarding the bookmarks created by other members of the group, the type of information they have gathered and how advanced their work is.</td>
<td>Group chat, which is always visible. “My Teammates” tab, with a summary of the work carried out by others.</td>
</tr>
<tr>
<td>Personal Build</td>
<td>Awareness regarding the number of notes created by other members of the team, as well as their content</td>
<td>Group chat, which is always visible. “My Teammates” tab, with a summary of the work carried out by others. Shared Virtual Work Table.</td>
</tr>
<tr>
<td>Personal Discover and Describe</td>
<td>Awareness regarding how advanced other members of the group are in building personal summary outlines</td>
<td>Group chat, which is always visible. “My Teammates” tab, with a summary of the work carried out by others. Shared Virtual Work Table.</td>
</tr>
<tr>
<td>Group Build</td>
<td>Shared view of the outline or conceptual map that gathers the group’s ideas.</td>
<td>Group chat, which is always visible. “My Teammates” tab, with a summary of the work carried out by others. Shared Virtual Work Table.</td>
</tr>
<tr>
<td>Group Discover and Describe</td>
<td>While one student writes the essay, the rest must be able to view what he is writing, offer feedback, give additional information, etc.</td>
<td>Group chat, which is always visible. Shared Virtual Work Table. Interface that allows students to view the Final Document.</td>
</tr>
</tbody>
</table>
Empirical evaluation

To evaluate the impact that the CollSearch tool has on students’ work, we conducted an experimental trial with 60 students from the eleventh grade of a private non-government supported high school in Santiago, Chile.

Method

The activity was part of an English as a second language course and asked students to write a letter to the International Olympic Committee (IOC) nominating Santiago, Chile as the host city for the 2024 Olympic Games. Students were asked to include a minimum of three of the following topics in their documents: Santiago’s sports infrastructure, Santiago’s hotels and accommodation infrastructure, Santiago’s crime rate and safety, Santiago as a tourist attraction, Santiago’s transport system, or Santiago’s climate. The goal was for them to write in English a compelling argument that would convince the IOC to consider Santiago as a possible host city.

Students were randomly split into groups of 3 and then each of the 20 groups was randomly assigned to one of the following independent activities:

- Control Group: Each student was given access to a personal computer with an internet browser and word processing software. Written instructions were provided and each group was free to organize their work in whatever way they chose. The activity facilitator (teacher) conducted the introduction to the topic and an initial search term discussion (similar to the one the CSPM considers). Later during the activity, the teacher was available to answer questions and guide students through the activity if they asked for her help.
• Experimental Group: Each student was given access to a personal computer with the CollSearch tool. A week before the activity took place, students assigned to this group had an initial 90 minute introduction to the tool and worked in a preliminary activity with a completely different topic and different groups than the experimental session. This initial activity was meant to teach students how to use the CollSearch tool and control for any changes in performance between the experimental and control groups that could be explained by the fact that the CollSearch tool was unknown to students while the browser and word processing software used by the control group was not. During the experimental activity, students were given the same set of written instructions given to the control group students. The activity facilitator (teacher) was available to answer questions and guide them through the activity if they asked for her help.

Each group was given 90 minutes to complete the activity and had to hand in only one letter per group. After they had done this, they had to access an online form where each student was asked to rewrite their letter according to what they remembered from their group’s work; during this phase of the activity students were forbidden to interact with each other and no longer had access to the document they had built as a group. The purpose of this second part of the activity was to evaluate how much of the group work had each student appropriated by participating in the group work. An external teacher who did not know which documents belonged to the experimental or control groups conducted a blind evaluation of each of the groups’ work. The evaluation measured the overall quality of the letter considering the groups’ argumentation and use of facts. A second evaluation was then conducted in which the external evaluator was asked to assess how each individual letter compared to the group’s letter; this was done by comparing how much of the information contained in the group’s work was reflected in each individual letter. The external evaluator assigned a percentage from 0% to 100% to each individual letter according to how complete it was compared to the group’s letter. This meant that any individual letter that contained as much (or more) information than the group’s letter would be assigned a 100%, independent of the first evaluation that measured the overall quality of the group’s letter.

Results and discussion

Table 4 shows the results of the experimental evaluation of the group work and the comparison between the individual and group letters.

<table>
<thead>
<tr>
<th></th>
<th>Percentage of accomplishment of the activity’s objectives</th>
<th>p</th>
<th>Percentage of the group work reflected in the individual work</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Group</td>
<td>87%</td>
<td>0.0064</td>
<td>61%</td>
<td>0.0087</td>
</tr>
<tr>
<td>Experimental Group</td>
<td>99%</td>
<td></td>
<td>77%</td>
<td></td>
</tr>
</tbody>
</table>

Results show a statistically significant difference between the performance of the control and experimental groups. The evaluation of the level of achievement of the activity’s goals raises when students work using CollSearch instead of freely choosing their work methodology (Cohens’d=1.24). When each student’s individual work is compared to their group’s work, there is a drop between the amounts of information that the group incorporates into their work versus the information that each student is able to recall when working individually. Despite this being true in both groups, there is a statistically significant difference between the experimental and control groups; students who followed the CSPM script within CollSearch can individually recall more information from their group’s work than students that could freely choose their group’s work methodology (Cohens’d=0.63).

The differences observed between the control and experimental groups can be explained by the fact that the CSPM fosters collaboration by forcing students to be aware of their teammates’ work and to build their final answers together. The activity facilitator observed that most of the students from the control group chose to work in cooperative ways in which each student was responsible for one part of the letter, and then the final document was built copying and pasting each part together; CollSearch through the CSPM prevents students from doing this by forcing them to discuss as a group their work and to build together their final document. This might also explain why students from the experimental group are able to recall more information from their group’s work than students from the control group; instead of simply giving a chunk of text to their group’s work, CollSearch made them be aware of their teammates’ research and contribute their work in a way that appropriately fit their group’s final document.
Conclusions and future work

The tools we use today to search the Internet do not offer the possibility for groups of users to work collaboratively. The articulation of Kuhlthau’s ISP with the macro and Micro-Script structure proposed by the Collaborative Search Procedural Model is a first approximation to supporting collaborative search practices.

Students who face collaborative search activities while following the CSPM when compared to those who can freely determine their work methodology show better appropriation of their group’s work. Empirical evaluation shows that the CSPM increases students’ performance and improves the overall quality of their work. Future work must be conducted to understand the impact of systematically using CollSearch as a collaborative learning tool and to study how the observed differences change over time as students and teachers better appropriate the model. An exploration of how different Micro-Scripts and implementations of CollSearch might have bigger or smaller impacts on student performance might also provide further insight on the ways in which the CSPM can have a positive impact on collaborative learning activities. Additionally, from a technical point of view, instead of being built from scratch, new versions of CollSearch could be built using previously existing online tools mashed-up into a unified tool.

Finally, it is important to note that Collaborative Search is a scarcely explored and emerging research field; the approach followed in the project presented in this paper is one of many possible alternatives. The Program for International Student Assessment (PISA) has emphasized the importance of the collaborative component, and starting in 2015 it will measure the capacity and willingness of students to solve problems through interaction among themselves (Davidson, 2012; De Jong, 2012); because of this, we encourage the computer science research community to explore other models that promote collaboration in the search task. Additional elements that must be addressed in future work are the dangers of over-scripting (Dillenbourg, 2002). Scripts provide the necessary scaffolding to enable users to effectively collaborate, but when the overhead introduced by following specific Micro-Script instruction sets surpass the benefits of the new work dynamics, users’ performance and throughput drops. Finding strategies to build effective macro and Micro-Script structures is a challenge the community must face through more empirical investigations. Future work must focus on ways to prevent over-scripting and on how to balance the need for a structure to guide activities and the need for students to be able to determine the structure of their workflow. Although this paper presents empirical evaluation conducted with high school students, there are multiple educational settings in which collaborative search can prove useful to improve student outcomes; the use of CollSearch and the CSPM in academic and university settings provides a fertile field to explore in future work.

Acknowledgments

Research supported by the Center for Research on Educational Policy and Practice, Grant CIE01-CONICYT.

References


(Ed.), Three worlds of CSCL. Can we support CSCL (pp. 61-91). Heerlen, the Netherlands: Open Universiteit Nederland.


Effectiveness of an Electronic Performance Support System on Computer Ethics and Ethical Decision-Making Education

Serhat Bahadır Kert, Çiğdem Uz and Zeynep Gecü

Department of Computer Education and Instructional Technologies, Yildiz Technical University, Istanbul, Turkey

Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from the editors at kinshuk@ieee.org.

(Submitted September July 8, 2013; Revised December 24, 2013; Accepted January 20, 2014)

ABSTRACT

This study examined the effectiveness of an electronic performance support system (EPSS) on computer ethics education and the ethical decision-making processes. There were five different phases to this ten month study: (1) Writing computer ethics scenarios, (2) Designing a decision-making framework, (3) Developing EPSS software, (4) Using EPSS in a case-based computer ethics education process, and (5) Analyzing all of the data collected from the implementation. A total of 15 computer ethics scenarios were written by the researchers and revised in accordance with the opinions of 12 experts from different universities. Barger’s (2008) ethical decision-making model was adapted to a flow chart for generating a decision-making framework for the system. Quantitative and qualitative research methods, multivariate analysis of variance and semi-structured interviews were used together to investigate the effectiveness of the system. The results showed that the EPSS that was developed improved the decision-making skills of the students in the experimental group who were faced with different computer ethics scenarios during the implementation of the study. To present a balanced view both positive and negative comments made by the students regarding the EPSS were recorded and are presented in the study.

Keywords

Media in education, Interactive learning environments, Pedagogical issues, Higher education, Computer ethics

Introduction

Computers play an important role in modern societies. We communicate with our friends, read articles, share documents and even find new social groups via computers. If a technological tool is so nested within the daily life of the human being, it is then inevitable that new social problems will emerge related to its usage. Some of these emerging issues consist of different ethical cases. Barger (2008) pointed out that old ethical problems have evolved into new complex ones because of the remarkable operating capacities of computers. However, ethical principles are not based on punishments as is the case with the legal system or on changeable rules as is the case with personal morality. The most important goal in ethics is to regulate social relationships among individuals by using a global perspective. Mason (1986) categorized the ethical problems of the information age under four headings: Privacy, accuracy, accessibility and property. The attributes of intellectual property, the rules of access to common knowledge, the certain limits of individual or official privacy and the problems arising from the transmission speed of wrong data are a few of the subtopics discussed under these headings. Specific ethical approaches, concepts, policies or values related to computers are studied in the field of computer ethics. The most effective and scientific way to cope with ethical issues is by using the ethical decision-making models proposed by different researchers. However, using these ethical decision-making models to solve the computer-related ethical issues faced in everyday life can be tough and include complex processes for individuals to follow. In this regard, using an EPSS in the role of the facilitator of an ethical decision-making process can be an effective and easy way to reach logical solutions to ethical issues.

In the literature, it can be seen that there have been a few attempts to support ethics or computer ethics education processes through the technological tools. Coldwell (2000), with his study titled "It is possible to teach computer ethics via distance education", investigated the possibilities of accessing computer ethics content by using Internet connections. At the end of his research, he mentioned that if the required technological infrastructure is not used then providing an online computer ethics course would be harder than doing so through traditional methods. In 2005, an online computer ethics course was given by Ferhan Odabaşı and Abdullah Kuzu (Uysal & Odabasi, 2013), which was a project financially supported by European Union. In this 6-week course, case-based methods were used for online ethical decision-making and evaluation processes. In another study, Biggerstaff (2008) used the online environment to give a social work ethics course. She used a case-based ethics education model in her study.

ISSN 1436-4522 (online) and 1176-3847 (print). © International Forum of Educational Technology & Society (IFETS). The authors and the forum jointly retain the copyright of the articles. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear the full citation on the first page. Copyrights for components of this work owned by others than IFETS must be honoured. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from the editors at kinshuk@ieee.org.
In a new attempt, an online EPSS named BILEP was developed in this study to support computer ethics education and ethical decision-making processes. The effectiveness of the system was examined over the span of the 3-month experimental study in a computer ethics course. During this experimental study, an attempt was made to find answers to these specific research questions: (a) What could be the positive effects of an EPSS on a theoretical computer ethics education? (b) What could be the positive effect of an EPSS on an ethical decision-making education? Additionally, at the end of the experimental process, the opinions of the participants regarding the BILEP system and the use of such software in ethics education were recorded in semi-structured interviews.

Literature review

Theoretical bases of computer ethics

The impact of the use of technological tools has increased dramatically in recent decades. As a result, the ethical use of computers and information technology has also become a subject of great interest. Grant, Stahl and Rogerson (2009) stated that computer ethics, which is a relatively new discipline, brings together people such as computer scientists, engineers, teachers, academicians, philosophers and psychologists in order to discuss ethical issues. The impressive foundations for computer ethics were laid down by Norbert Wiener between the 1940s and 1950s; the significant efforts of Donn Parker, Joseph Weizenbaum and Walter Maner in the 1970s; James Moor’s and Deborah Johnson’s contributions in the 1980s and Krystyna Gorniak’s hypothesis in the 1990s (Bynum, 2001).

In his groundbreaking article “What is Computer Ethics?” Moor (1998) stated that computer ethics is a special field of ethical research and application. He also mentioned that computer ethics involves the analysis of two areas: “the analysis of the nature & social impact of computer technology” and “the corresponding formulation and justification of policies for the ethical use” (Moor, 1998). Like James Moor, Walter Maner (1996) made reference to computer ethics as a unique field of study. This statement was based on his view that the “involvement of computers in human conduct can create entirely new ethical issues, unique to computing, that do not surface in other areas” (Maner, 1996, p.139). On the other hand, Johnson did not defend the issue of the uniqueness of computer ethics and proposed that ethical issues surrounding computers are “new species of old moral issues” (as cited in Moor, 1998). According to Barger (2008), like Johnson, computer ethics are not qualitatively different from media ethics or legal ethics or any other kind of professional ethics.

According to Gotterbarn and Rogerson (1998), both Moor (1985) and Gorniak (1996) demanded that the impact of the computer is pervasive in all aspects of society. While Moor (1985) concluded that people need to apply ethics to the policy vacuums created by computers, Gorniak questioned the applicability of traditional ethics to global ethics (Gotterbarn & Rogerson, 1998). Like Gorniak (1996), Johnson (1999) stated that computer ethics will disappear as a separate branch of applied ethics. In addition, Johnson said that information technology would become very commonplace and it would be perceived simply as an aspect of everyday life (as cited in Bynum, 2001). Also, in order to study computer ethics, people should be aware of the nature of moral judgment and the limitations of ethical decision making in computing (Stamatellos, 2007).

Ethical decision-making models

It can be said that the most important part of the developed EPSS is an “expert support system” in which user can find a step-by-step ethical decision-making platform. Before designing this part of the software, authors conducted research on the fundamentals of ethical decision making and decision-making models. The details of this research are presented below:

Ethics provides reasons for how humans should act and it addresses the process of ethical decision making. Ethical decision making is the process of evaluating and choosing among alternatives by the taking into consideration of ethical principles. In making ethical decisions, it is important to discriminate and eliminate unethical options and select the best and most appropriate ethical alternative. This process requires individuals to make hard choices. There are a variety of models available in the effort of to explain and predict the process of ethical decision making (Corey et al. 1998; Mattison, 2000; Robinson & Reeser, 2000; Welfel, 2006; Cottone, 2001; Stephenson & Staal,
These models can make the decision-making process easier and more systematic. The heuristics approach, which is a stepwise procedure, is commonly used in ethical decision-making models.

Ethical decision-making models can be developed for specific domains such as marketing, social science, clinical psychology, law and computers or can be employed in general frameworks. For instance, the Canadian Psychological Association (1991) have developed a model consisting of seven steps for psychology, Corey et al. (1998) have developed a model for law and Barger (2008) and Maner (1999) have developed a model for computer ethics.

In this study, the model proposed by Barger (2008) is used to specify the general framework of an electronic support system. Barger (2008) suggested a model consisting of steps involving the answering of questions in order to resolve computer ethics dilemmas. The steps are:

- Briefly describe the ethical issues
- Identify the stakeholders and state what each of them would like to offer as the solution to the problem.
- Define three solutions to the case, mark them and select the solution you would choose.
- Decide if the solution you chose would be acceptable for everyone.
- Determine if this solution is accordance with what is natural.
- Decide if there would be a majority agreement that this solution is the most efficient.
- Make a decision if this solution is the one to which you feel most committed in your own conscience.
- Determine which philosophy you feel was the most influential in your solution to this case.

These steps were used for designing the algorithm of the performance support software developed by researchers. As the first step, all items were converted to an algorithmic structure, and later, a flowchart of the software was created in accordance with the phases of this algorithmic text.

Why EPSS for ethical decision making?

Performance support systems are performance-oriented tools which are used individually and intended for specific problems. The main aim of an EPSS is to support the user during an operation and try to prevent the interruption of the process. There are various definitions of electronic performance support systems. Raybould (1990) defines EPSS as "a computer-based system that improves worker productivity by providing on-the-job access to integrated information, advice, and learning experiences". These systems can be designed for various purposes. For instance, Van Schaik et al. (2002) used EPSS within the domain of 'quantitative research methods' to facilitate learning. Sleight (1993) stated that EPSS has a two-part characteristic. The first part is to provide access to the specific information and tools needed to perform a task. The second part is providing access to the information and tools at the time the task is to be performed. Levin (1994) stated that EPSS facilities can offer many different benefits for an organization. Nguyen and Klein (2008) examined the effect of EPSS and training on user performance, time on task, and time in training and positive effects of EPSS usage were detected on user performance and time in training in this study. However, Bastiaens et al. (1997) developed an EPSS for a large Dutch insurance company and found that EPSS did not produce the expected benefit of an increase in productivity. According to Raybould (1990), it is composed of an Advisory System (expert system), an Information Base (interactive documentation), Learning Experiences (support) and Productivity software.

The main characteristics of EPSS are that they are computer-based, they provide access during tasks, they are used on the job, they are controlled by the worker, they reduce the need for prior training, they are easily updated, they provide fast access to information, they do not display irrelevant information, they allow for different levels of knowledge in users, they allow for different learning styles, they integrate information, advice, and learning experiences and that they possess artificial intelligence (Sleight, 1993).

Considering the remarkable features of EPSSs, it is believed that the specific and unique attributes of decision-making processes for ethical cases can be supported through the use of an EPSS in different ways. In this study, by the taking into account of these kind of potentials for EPSS usage, the possible effects of this software on ethical decision making as well as theoretical computer ethics education were investigated within the span of the 3-month experimental study.
Methodology

Rationale

Case-based computer ethics and ethical decision-making processes consist of very subjective topics such as open-ended scenarios or philosophical comments. Therefore, talking about only a theoretical content related to ethical issues would be insufficient for computer ethics education. On the other hand, finding a practical solution to an ethical problem is more important than learning scientific content for individuals. In this respect, the rationale for the development of BILEP system can be explained in terms of combining the basic scientific content and a solution-oriented tool in the same software. Hereby, it was aimed to propose a new approach to the provision of online computer ethics education as an alternative to traditional methods.

Writing computer ethics scenarios

In order to create a case-based educational environment, 15 different ethical scenarios related to computer ethics were designed by researchers. All content was selected from actual problems in computer ethics and written in accordance with the four important ethical areas outlined by Mason (1986): Privacy, accuracy, property and accessibility. A sample from the scenarios can be seen below:

Title: Security Threat
Ali chooses improvement of commercial systems and security threats as a research subject for his Informatics Ethics course. He develops an algorithm for the purpose of implementing his project in a practical process. He can find the security vulnerabilities of the companies’ network systems with this algorithm and decides to use these security vulnerabilities in his project. His advisor professor approves Ali’s project and gives him a very high project grade. When one of the companies that Ali has entered searches for the source of the attack, they find university laboratory as a source of this attack and informs the chancellery of the university.

After having written the first manuscripts for the scenarios, an evaluation form was prepared in order to check some key features concerning the structures of the scenarios. Each manuscript was sent to three different scientific experts in order to be evaluated and the last revisions were made using the responses and comments of the experts. These revisions were discussed under four titles: the grammatical mistakes of the text structures, the relationships of determining roles to the cases, the relationships of the cases to computer ethics and the details concerning the cases that were still required. At the end of the process, some detailed revisions were made to the scenarios based on the responses from the experts and the final versions were uploaded to the BILEP system.

Designing flowchart of expert system

Before starting to develop the expert module of BILEP, it was thought by researchers that the only way to adapt a theoretical model to the software was to design a logical and clear flowchart. To this end, an elaborate flowchart was designed in accordance with Barger’s (2008) eight steps ethical decision-making model. Each step of the model was written in short commands onto a systematical and organizational flowchart. The first part of this detailed flowchart can be seen in Figure 1.

As seen in Figure 1, some steps of the model, which were too abstract to adapt to a programming language, were represented with digits in the flowchart. For example, in order to represent the statement “Identify the stakeholders and state what each of them would like to offer as solution of the problem”, a loop counting by the number of the roles was created and it was required for us to write possible solutions for each role. Additionally, these possible solutions for the roles were scored out of 100 by comparing the total score of users’ own solutions at the end of the process. Hereby, users had the opportunity to make comparisons between different solutions and choose one of them.

After designing the process, the coherence of the flowchart with the underlying model was checked for each item and the last structure of design was shaped by taking into account the thoughts of the scientific experts.
Developing EPSS software

ASP.NET 4.0 software, the SQL Database management system and Adobe Flash CS5 authoring language were used for developing different parts of the BILEP EPSS software. BILEP is composed of five modular sections: (1) Knowledge support system, (2) Expert support system, (3) Ethical scenario forum, (4) User portal and (5) Communication panel. The whole organization scheme of BILEP can be seen in Figure 2.

All data used in different sections were recorded in a database file which could only be managed by the administrator of the system. The expert support system was used to support the ethical decision-making steps for the participants. Theoretical information about computer ethics could be accessed through the knowledge support system. 15 ethical
scenarios were designed and added to the ethical scenario forum section during the development process of the BILEP system. Additionally, a synchronous communication panel, namely a ‘chat panel’, was placed in the system to communicate with the expert in cases of problems concerning the process. The homepage of BILEP system, which includes all menu items and links provided in order to reach the modules, can be seen in Figure 3.

Considering the aim of the study, an attempt was made to develop BILEP with a tidy interface and simplicity was a very important part of the development process. Users opening the page can view the login section at the top of the page and items linked to modules on the left side. Furthermore, the final comments on ethical scenarios, together with the usernames of the authors, can be seen on the right side of the page.

Participants

A total of 40 university students who attended the “Computer Ethics” course for the first time in the 2011-2012 spring semester were participants in this study. Participants were between 19 and 22 years of age; 23 females and 17 males. Their department was the Department of Computer Education and Instructional Technologies in the Faculty of Education. It can be said that this group was not a representative sample as all of the subjects were third-year students who were taking the computer ethics course for the first time. All students participated in the study during the experimental implementation. Computer Ethics was a third-year undergraduate course including two theoretical and two practical classes each week. All participants were divided into two groups which composed the experimental and control groups of the study, respectively. The groups in the study were assigned randomly as experiment and control. In order to eliminate the possible negative effects of random assignment: (1) Prior knowledge of the participants was taken under control by getting their pre-experimental academic scores through a pretest exam. (2) The ages of group members were compared (Average ages of Experimental and Control group members were 21.2 and 19.7, respectively). (3) Researchers checked the gender differences between two groups (it was seen that there were 12 females and 8 males in experimental group, 11 females and 9 males in control group). The students in the experimental group were supported only with BILEP EPSS while the others were supported by the lecturer of the course during the three months of the implementation. Additionally, at the end of the experimental process, seven students among the experimental group who volunteered to take part in the qualitative study were interviewed about the usage of the BILEP system and its future implementations.

Experimental implementation

The experimental part of the study commenced after the design and development processes of the BILEP system were complete. During these three months, the experimental and control groups of participants were given computer ethics lectures by the same lecturer. All educational content was the same for both groups. The only difference between the groups was that the students in the experimental group used the BILEP system as a support tool through the Internet connection, while the other group was supported by the lecturer in practical classes. Pre and post academic success tests related to theoretical ethics and ethical decision-making processes were used to compare the educational growth of the groups. The test questions were prepared according to the conceptual framework by undertaking a literature review and including course content concerning relevant topics such as computer ethics.
software piracy, basic ethical issues and Internet ethics, security options for Internet. After that, the questions were revised with respect to the opinions of experts in the field before their implementation. The test-retest reliability coefficient of the academic success test used in the study was 0.79.

Semi-structured interviews

After the experimental process, semi-structured interviews were conducted with seven students from the experimental group. All of these students volunteered to be interviewed. Each participant was individually interviewed and interviews were conducted in Turkish. Throughout the interviews, the researchers sought to obtain students’ thoughts on computer ethics education, their comments on the effectiveness of the process and suggestions for possible additions to BILEP.

Results

Quantitative results

In order to examine the academic growth of the students taking the computers ethics course, an academic success test was developed by the researchers. The developed test was composed of two different sections including different subtitles: one of them was computer ethics and the other one was ethical decision making. Due to the contextual differences between these two topics, the scores obtained were analyzed separately.

Before the three-month experimental implementation process, the prior knowledge of the students was tested. At the end of the process, the same academic success tests were conducted as posttests and comparisons were made between both the pretest and posttest scores of the students in different groups. The pretest and posttest results are shown in Table 1.

<table>
<thead>
<tr>
<th>Sub-categories of the academic success test on computer ethics</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Control Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethical decision making</td>
<td>19.00</td>
<td>6.20</td>
</tr>
<tr>
<td>Theoretical computer ethics</td>
<td>30.00</td>
<td>6.15</td>
</tr>
<tr>
<td>Experimental Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethical decision making</td>
<td>21.20</td>
<td>5.52</td>
</tr>
<tr>
<td>Theoretical computer ethics</td>
<td>33.40</td>
<td>6.65</td>
</tr>
</tbody>
</table>

As seen in Table 1, only very slight differences were found between pretest scores of the groups at the sub-categories level. The independent samples t-test method was used to analyze the significance level of the difference between pretest scores. The t-test results showed that there were no significant differences between the groups in terms of both of the two sub-categories (p > .05). After these results had been obtained, it was decided to use multivariate analysis of variance (MANOVA) to analyze the differences between the posttest scores of the groups. Three important assumptions of MANOVA (normal distribution of dependent variables, homogeneity of covariance matrixes and equality of error variances) were confirmed before using this analytical method. The MANOVA results of the posttest scores are shown in Table 2.

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Dependent Variable</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
<th>Eta Squared (ƞ²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>Ethical decision making</td>
<td>384.400a</td>
<td>1</td>
<td>384.400</td>
<td>8.536</td>
<td>.006</td>
<td>.183</td>
</tr>
<tr>
<td></td>
<td>Theoretical computer ethics</td>
<td>67.600b</td>
<td>1</td>
<td>67.600</td>
<td>1.005</td>
<td>.322</td>
<td>.026</td>
</tr>
</tbody>
</table>
It can be seen in Table 2 that no significant differences were found between the “Theoretical Computer Ethics” scores of the participants ($F(1,38) = 1.005, p > .05$). However, it was revealed that the “Ethical decision-making” scores of the groups significantly differed from each other in favor of the experimental group ($F(1,38) = 8.536, p < .05$). In other words, the analytical results showed that the BILEP system positively affected the improvement of the ethical decision-making knowledge of the students. Furthermore, it was found that whole test point was affected by the “Ethical Decision Making” scores ($\eta^2 = .183$) rather than “Theoretical Computer Ethics” ($\eta^2 = .026$). After the posttest comparisons, paired samples $t$-test analyses were conducted to reveal the level of in-group academic growths. The paired samples $t$-test results for the control (Table 3) and experimental (Table 4) groups’ scores are presented below:

<table>
<thead>
<tr>
<th>Group</th>
<th>Ethical decision making</th>
<th>Theoretical computer ethics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>N</td>
</tr>
<tr>
<td>Ethical Decision Making</td>
<td>31584.400</td>
<td>1</td>
</tr>
<tr>
<td>Theoretical computer ethics</td>
<td>55056.400</td>
<td>1</td>
</tr>
<tr>
<td>Ethical decision making</td>
<td>384.400</td>
<td>1</td>
</tr>
<tr>
<td>Theoretical computer ethics</td>
<td>67.600</td>
<td>1</td>
</tr>
<tr>
<td>Error</td>
<td>Ethical decision making</td>
<td>1711.200</td>
</tr>
<tr>
<td>Theoretical computer ethics</td>
<td>2556.000</td>
<td>38</td>
</tr>
<tr>
<td>Total</td>
<td>Ethical decision making</td>
<td>33680.000</td>
</tr>
<tr>
<td>Theoretical computer ethics</td>
<td>57680.000</td>
<td>40</td>
</tr>
</tbody>
</table>

| Table 3. Paired samples $t$-test results for control group scores |
|------------------------|-------------------------|
| Sub-categories of test | Mean                    |
| Ethical Decision Making | 19.00                   |
|                       | 20                      |
|                       | 6.20                    |
|                       | 19                      |
|                       | 4.17                    |
|                       | .00                     |
|                       | 25.00                   |
|                       | 20                      |
|                       | 5.78                    |
|                       | 19                      |
|                        | 4.65                    |
|                        | .00                     |
| Theoretical Computer Ethics | 30.00                   |
|                       | 20                      |
|                       | 6.15                    |
|                       | 19                      |
|                       | 4.65                    |
|                       | .00                     |
|                       | 35.80                   |
|                       | 20                      |
|                       | 7.28                    |
|                       | 19                      |

| Sub-categories of test | Mean                    |
| Ethical Decision Making | 19.00                   |
|                       | 20                      |
|                       | 6.20                    |
|                       | 19                      |
|                       | 4.17                    |
|                       | .00                     |
|                       | 25.00                   |
|                       | 20                      |
|                       | 5.78                    |
|                       | 19                      |
|                        | 4.65                    |
|                        | .00                     |
| Theoretical Computer Ethics | 30.00                   |
|                       | 20                      |
|                       | 6.15                    |
|                       | 19                      |
|                        | 4.65                    |
|                        | .00                     |
|                       | 35.80                   |
|                       | 20                      |
|                       | 7.28                    |
|                       | 19                      |

| Table 4. Paired samples $t$-test results for experimental group scores |
|------------------------|-------------------------|
| Sub-categories of test | Mean                    |
| Ethical Decision Making | 21.20                   |
|                       | 20                      |
|                       | 5.52                    |
|                       | 19                      |
|                       | 6.045                   |
|                       | .00                     |
|                       | 31.20                   |
|                       | 20                      |
|                       | 7.52                    |
|                       | 19                      |
|                        | 2.54                    |
|                        | .02                     |
| Theoretical Computer Ethics | 32.40                   |
|                       | 20                      |
|                       | 9.02                    |
|                       | 19                      |

The $t$-test results revealed that the computer ethics education process in each group significantly affected the academic growth of the students in the areas of “Theoretical Computer Ethics” (control group: $t = 4.65, p < .05$; experimental group: $t = 2.54, p < .05$) and “Ethical Decision Making” (control group: $t = 4.17, p < .05$; experimental group: $t = 6.04, p < .05$) levels.
Given the results of the whole quantitative analyses, it can be said that EPSS-supported practices positively affect the level of ethical decision-making knowledge. On the other hand, the same effect did not take place with regard to theoretical computer ethics education. This result is compatible with the performance-oriented characteristics of EPSSs. After the experimental implementation, semi-structured interviews were conducted to collect the comments and suggestions of the participants concerning computer ethics education as well as the usage of BILEP system.

**Interview results**

During the semi-structured interviews, different questions about computer ethics and using the BILEP system were asked of the participants. Due to the short and similar answers of the students, a brief theme analysis was performed on the answers instead of a detailed one. Interview data were categorized according to the questions. The most repeated and the most salient answers in each question are presented in the following section.

Firstly, the thoughts of the participants on the necessity of computer ethics education were requested. The importance of computer ethics education was emphasized by all of the students. Some of them explained their reasons with the following statements:

“In my opinion, computer ethics education is very important for individuals. We need to know how to make the right decision when we are faced with a complex ethical problem in our everyday life.” (Student C interview, 15 June, 2012).

“Nowadays, everybody is interacting with each other through online social environments and sometimes individual privacy is a big problem, that’s why I think that everybody in our digital age should get computer ethics education.” (Student F interview, 15 June, 2012).

“Yes, I believe that computer ethics education is necessary, because, we are using Internet and using online shopping sites more than before. Some people might abuse these kind of online implementations, in such cases, we need to know what can we do and how can we respond.” (Student E interview, 15 June, 2012).

Secondly, the opinions of the students regarding the most important points of an ethical decision-making process were recorded. It can be said that the approaches of the students could be divided into two general categories: the realistic approach and case-based approach. Examples of some statements expressing the realistic approach were as follows:

“Empathy should be the first step of the process, it is important to find a conciliatory resolution by looking at things from the different points of view of stakeholders in the ethical issue” (Student B interview, 15 June, 2012).

“Problems should be analyzed and the stakeholders in the problem should be determined, and later, we must ask this question to ourselves “If I were them, what I would want to be done?”. By answering this question we can make the right decisions.” (Student G interview, 15 June, 2012).

“I think that the most important point of the process is to be able to be impartial while making a decision on an ethical issue.” (Student D interview, 15 June, 2012).

One student made a case-based response to this question:

“We must think in accordance with the case, ethical issues can be different from case to case. Therefore, a specific ethical decision-making process must be composed for each case.” (Student B interview, 15 June, 2012).

After being asked general questions related to computer ethics, students were asked for their evaluations of using the BILEP system as an ethical decision-making assistant. All of the responses were positive in different ways, some examples are written below:
“The steps in the BILEP system provided us with the opportunity to think more elaborately than before, to write comments for the scenarios and discuss the different comments with our friends.” (Student C interview, 15 June, 2012).

“The steps of the ethical decision-making system lead us to think about cases more elaborately and, in addition, we could see the various comments made by other friends on the same case and discuss them within this system.”(Student A interview, 15 June, 2012).

Lastly, the suggestions from the participants concerning possible additions and improvements to the BILEP system were recorded. One student suggested the following improvement:

“When I first used the system, the steps of the expert module were very complex for me. I think it is better if a popup window on which there is a short description of how to use the expert system would appear before the process, so novice users can use the system without any confusion.” (Student G interview, 15 June, 2012).

Technical problems faced while using the system were seen as the major problems for the students. An example comment is presented below:

“I was faced with some technical problems while using the expert system. When I wanted to switch to the previous step, my session was closed automatically, it was really frustrating for me to log in the system again and again.” (Student A interview, 15 June, 2012).

It must be said that such technical problems mentioned by the students resulted from the hardware and Internet connection problems of the computer laboratory rather than the software structure of BILEP. Considering the content of all comments concerning the computer ethics implementation and usage of the BILEP system, it can be stated that a high level of satisfaction among the participants regarding the usage process of BILEP was recorded at the end of the study. However, in order to provide more convenient usage to users, it would be required for us to design less complex modules than those that currently exist. Additionally, it would be useful to provide an information window which appears on the first page of the BILEP expert module, so users can be informed about the possible technical problems they may encounter during usage. Moreover, during the interviews, the emerging consciousness regarding the importance of computer ethics education and the ethical decision-making process in minds of the students show that the experimental process ended with positive results in a qualitative, as well as a quantitative, sense.

Discussion

In this study, the effectiveness of EPSS on computer ethics education and the ethical decision-making process was examined. According to the results of the study, although there is no significant difference between the “Theoretical Computer Ethics” scores of the participants, there is a significant difference in terms of the “Ethical Decision-Making” scores of the groups. This indicates that the EPSS was successful in supporting students when performing tasks related to ethical decision making. When the related literature is considered, it could be stated that this finding is consistent with the study carried out by Raybould (1990) who stated that EPSS can be used in an advisory role for an ethical decision-making process. EPSS provides both specific information and access to the information at the time the task is to be performed (Sleight, 1993) so this result is also compatible with the characteristics of EPSS. Van Schaik et.al (2002) developed an EPSS for the domain of ‘quantitative research methods’ as taught within a psychology degree course and found that EPSS was successful in supporting students when performing tasks related to quantitative research methods. Levin (1994) stated that EPSS users have a high motivation to learn, they are in control of the learning process and their need to know the specific information is higher than in traditional training. Furthermore, Nguyen and Klein (2008) examined the effect of EPSS and training on user performance, time on task, and time in training. The results showed that participants receiving only EPSS and those receiving training and EPSS performed significantly better on a tax preparation procedure than participants who received only training. In this study, it was found that EPSS-supported practices positively affect the level of ethical decision-making knowledge.

However, it was also seen that the findings of this study are not supported by Kalota and Hung (2012) who claimed that EPSS did not improve participants’ learning. Bastiaens et al. (1997) developed an EPSS consisting of information, advice and learning material about products for a large Dutch insurance company which sells insurance products to their clients. They compared EPSS with the traditional environment. However, some older employees who were willing to work with the system preferred the traditional way. Bastiaens et al. (1997) stated that EPSS is not as successful as the literature often states.
Semi-structured interview questions were asked to investigate the students’ views on computer ethics and the BILEP system. First of all, the thoughts of the students on the necessity of computer ethics education were requested. The students interviewed highlighted the significance of computer ethics education. The students also shared their thoughts about the most important points of the ethical decision-making process in terms of realistic and case-based approaches. The students interviewed also stated their evaluations and recommendations concerning the BILEP system. Some of the students talked about BILEP positively and stated that the system required them to think about the cases systematically. As stated before, EPSS can provide support and improve and enhance performance in the learning process (Raybould, 1990; Milheim, 1997). Moreover, some students mentioned technical problems that they faced in the process. One of the students also talked about the complexity of the expert module in the system. This result supports the study of Kalota and Hung (2012) who stated that while a majority of the participants had positive feedback about EPSS, a couple of the participants found it confusing. However, Van Schaik et al. (2002) could not find significant correlation between students’ performance in terms of success in using the EPSS. Students in this study suggested in the semi-structured interview that social interaction could be integrated into the system. This could represent a possible future addition to EPSS.

Furthermore, according the study, the significant difference in terms of ethical decision-making scores is supported by qualitative results. The volunteered participants in the experimental group, who shared their thoughts about BILEP system as an ethical decision-making assistant, said that the systems made them to think more elaborately and systematically. Also they stated that their knowledge about the ethical decision-making process was increased. It might be said that the BILEP system helped students to think more deeply in ethical decision-making process as they stated. Moreover, some of the participants talked about looking at situations from different perspectives and stressed the importance of empathy. These opportunities that were provided by the BILEP system might be other impacts that affected the difference between the groups in terms of ethical decision making.

Many instructional methods have been used for the instruction of ethics in different domains such as medical ethics, work ethics, teacher education ethics, etc.; however, computer ethics is a relatively new field of study and it is difficult to envisage a variety of instructional methods at the present time. EPSS is used to reduce the cost of training staff while increasing productivity and performance and has a wide variety of usage in different domains. Furthermore, the positive effects on the performance of users can be seen in previous studies (Van Schaik et al, 2002; Levin, 1994; Nguyen & Klein, 2008). In this study, the results also showed that a developed EPSS can improve the decision-making skills of students and feedback from the students was considerably positive.

Conclusions

The present study showed that an EPSS can be used as an effective tool in order to support practical education processes related to computer ethics. It can be stated that the most important part of such kinds of implementation is having a first step to use an EPSS within ethical education. As adults of this technological age, we are supposed to prepare the next generation to be individuals who are conscious of the computer ethics issues that they will possibly face in the future. It is believed that the BILEP online system developed in the scope of this study will be very useful for both students taking computer ethics courses and adults who are confused about the computer ethics that manifest in daily life. Additionally, improving the capacity of the system with new scenarios, members and comments can lead to people living all around the world becoming aware of the importance of ethical issues related to technological tools.

References


Lost in Interaction in IMS Learning Design Runtime Environments

Michael Derntl$^1$, Susanne Neumann$^2$ and Petra Oberhuemer$^{2,3}$

$^1$RWTH Aachen University, Advanced Community Information Systems (ACIS), Informatik 5, Aachen, Germany // $^2$University of Vienna, Center for Teaching and Learning, Vienna, Austria // $^3$University of Vienna, Teaching Affairs and Student Services, Vienna, Austria // derntl@dbis.rwth-aachen.de // susanne.neumann-heyer@univie.ac.at // petra.oberhuemer@univie.ac.at

*Corresponding author

(Submitted March 15, 2013; Revised October 11, 2013; Accepted December 5, 2013)

ABSTRACT

Educators are exploiting the advantages of advanced web-based collaboration technologies and massive online interactions. Interactions between learners and human or nonhuman resources therefore play an increasingly important pedagogical role, and the way these interactions are expressed in the user interface of virtual learning environments is assumed to greatly impact the learning process. This paper presents an evaluation of how interactions for learning are expressed in the user interface of an IMS Learning Design (LD) runtime environment. Specifically, we selected Reigeluth and Moore’s set of interactions for learning was selected and tested how these interactions are visually expressed and supported in the user interface of the Service Based Learning Design (SLeD) player. The findings show several drawbacks in the current way of visualizing units of learning during runtime. For instance, the synchronization of participants in corresponding learning and support activities is not clearly expressed, the current role remains unclear in multi-role settings, and the nested display of unit-of-learning contents impedes navigation. Drawing from these findings, the paper offers recommendations regarding future development of IMS LD runtime user interfaces.

Keywords

Learning design, Learning interactions, IMS LD, SLeD player, Virtual learning environments

Introduction

Learners today use a multitude of web-based tools and services to support their interactions during learning experiences. While interactions for learning have always been an important aspect of education, several recent trends complicate the teachers’ roles as providers and orchestrators of learning interactions today. An example of such a trend is the rapid pace of development and advancement of web-based technologies that support interaction. Since the advent of Web 2.0 and successful web-based business models, new collaboration tools and services become available to web users on a daily basis for free. Another recent trend on a global scale, propelled by rapid innovation in web technology and web pedagogy, is the offering and support of learning opportunities that build on openness and interaction on a massive scale, e.g., in massive open online courses (Yuan & Powell, 2013). A decade ago, the learning-technology world was much simpler, limited to classic forms of interaction support such as chat rooms and asynchronous message boards to support learners and their interactions with other learners, tutors, and teachers. At that time, the IMS Learning Design (LD) specification was developed as an interoperability instrument that allowed course authors to describe a pedagogical approach in a so-called “unit of learning” (Koper & Olivier, 2004) using a well-defined set of concepts specified in the IMS LD information model (IMS Global, 2003). A unit of learning can be a formal model of a course, a seminar unit, a self-study unit, or any other teaching and learning activity. IMS LD differentiates between the design time (when the learning design models are being authored) and the runtime (when these models are instantiated and executed). This enables the transfer of designs between different learning design authoring tools and the reuse of designs and materials in any IMS LD-compliant runtime environment. This kind of interoperability of units of learning that are specified with IMS LD enables learners and teachers to use the unit of learning in the virtual learning environment (VLE) of their choice, as long as this system is implementing the IMS LD specification. On the one hand, the need today to preserve and transfer the designs of units of learning (e.g., successful MOOCs) is stronger than ever. On the other hand, the specification and its conceptualization of interaction was set in stone in 2003 and never revised since then, and in the light of recent developments towards massive online interactions for learning it is indicated to analyze the support of interactions for learning in existing IMS LD runtime environments.

This paper puts the frequently discussed design-time issues of IMS LD aside and focuses exclusively on the runtime perspective of IMS LD. One general challenge that comes with the strict separation of design time (authoring) and
runtime (execution) in IMS LD is the insecurity during authoring about how the designed unit of learning will look like in the VLE during execution, which was previously pointed out as a major problem by Neumann and colleagues (2010). Traditionally, teachers design the learning environment directly within the VLE, where they can immediately see the effects of their design choices because they are reflected in the user interface. For instance, when a teacher adds a new assignment in a Moodle course, s/he will immediately see the effects on the course page and can react accordingly. When using IMS LD software, design decisions have to be made beforehand without knowledge of the VLE, where the unit of learning will be deployed for learning.

Adding to these conceptual challenges, which can nonetheless be met by unit-of-learning authors (see Derntl, Neumann, Griffiths, & Oberhuemer, 2012) and appropriate runtime management tools (e.g., Leony, de La Fuente Valentín, Pardo, & Delgado Kloos, 2008), a severe shortcoming on the runtime side is that there are only few VLEs that are able to read and execute IMS LD units of learning. There have been various attempts to implement an IMS LD runtime in mainstream VLEs. An implementation in Moodle was discussed (e.g., in Burgos, Tattersall, Dougiamas, Vogten & Koper, 2007) but never realized. An IMS LD implementation as an extension of .LRN called GRAIL succeeded (Escobedo del Cid, de la Fuente Valentín, Gutiérrez, Pardo & Delgado Kloos, 2007); however, the resulting tool has not been adopted widely. Most of the other existing IMS LD players are either commercial products like CLIX (http://www.im-c.de/germany/en/solutions/learning-management/clix-learning-suite) or have been developed in European R&D projects, for example, the SLeD player (http://sourceforge.net/projects/lidplayer/) and the more recently developed Astro Player (http://tencompetence-project.bolton.ac.uk/lidruntime/index.html).

To identify how interactions for learning are expressed in current IMS LD runtime systems, we analyzed a representative IMS LD player, the SLeD player, according to relevant pedagogical principles. One goal of the analysis was to understand interaction issues in IMS LD runtime systems beyond pure technical interpretation and visualization of IMS LD units of learning. Drawing from this analysis, a second goal was to distill recommendations for IMS LD runtime environment developers.

Our paper is structured as follows. In the following section, we present the theoretical background of the study and the study objectives. In the third section, we present the methodology to obtain and analyze IMS LD units of learning at runtime. In the next section, we present a detailed account of the results obtained during runtime analysis. Following that, we discuss the technological restrictions and remedies to support runtime interaction with IMS LD. In the final section, we outline findings and recommend actions for future work.

**Background and objectives**

In this paper, we assume that the reader is familiar with the main concepts of the IMS LD specification. An excellent introductory article by Koper and Olivier was previously published in this journal (Koper & Olivier, 2004). Since then, a considerable body of research has been produced regarding the design time, that is, the authoring of IMS LD units of learning. (See Griffiths, Blat, García, Vogten and Kwong [2005] for an early discussion and Lockyer, Bennett, Agostinho and Harper [2009] for a more recent, comprehensive overview.) Given the current forces in web-based education as outlined in the introduction, it is equally important to view IMS LD from the runtime perspective. Less research was published on this aspect, since it is primarily an implementation effort. Because IMS LD, with its focus on roles and activities, uniquely set up among learning-technology specifications, it is particularly suited to be tested for the expression of pedagogical aspects in VLEs. For this reason, Reigeluth and Moore’s framework for comparing instructional strategies was chosen as an analysis framework (Reigeluth & Moore, 1999). This framework is appropriate because the contained aspects are well accepted and can be classified more precisely than with other frameworks, such as Reeves’ pedagogical dimensions of computer-based education (Reeves, 1997). Of the six aspects in Reigeluth and Moore’s framework, “interactions for learning” was selected for the evaluation because it appeared to be most meaningful for evaluating VLEs’ ability to deal with multiple roles and active engagement, which are considered to be strengths of IMS LD compared to other learning activity-focused specifications (Koper & Olivier, 2004; Griffiths & Liber, 2008).

In the framework proposed by Reigeluth and Moore (1999) human interactions are categorized into human and nonhuman interactions. The types of interactions that students engage in during the learning process are distinguished as follows in the framework, with the first three representing human and the last four representing nonhuman interactions:
• Student-teacher: interacting with the teacher or instructor
• Student-student: working with or utilizing other students as resources, individually or in a group
• Student-other human: interacting with a community member, parent, or other individual (or group)
• Student-tools: using tools that enable completion of tasks
• Student-information: working with, and making sense of, the information that is available or found
• Student-environment/manipulatives: utilizing and working with resources and simulations, both within and outside the classroom environment
• Student-other nonhuman: working with any other conceivable nonhuman resource.

Each of these interactions pursues a pedagogical purpose such as creating dependence between students for student/student interactions, going beyond the traditional classroom for learning challenges in student/environment interactions, or producing high quality products during student/tools and student/information interactions.

This analytical study was set up to analyze IMS LD units of learning (from here on simply referred to as “units of learning”) according to the (visual) expression of the above-mentioned interactions for learning in VLEs. This is not a test of the IMS LD specification itself (e.g., for its ability to express certain pedagogical aspects), but rather for the ability of VLEs to support and express pedagogical aspects contained in IMS LD units of learning and whether the chosen expression fosters or hinders the learning process.

### Methodology

#### Material selection

Units of learning were solicited from several European organizations that were part of the ICOPER consortium, a European Commission-funded best-practice network that investigated, among other issues, contemporary learning design standards and interoperability specifications (cf., Simon, Pulkkinen, Totschnig, & Kozlov, 2011). These organizations had built units of learning for specific disciplines, such as architecture, or for specific learning settings like collaborative learning. Criteria for eventual selection among collected units of learning were that they needed to exhibit characteristics relating to the interactions for learning. Six units of learning were selected that exhibited at least two different types of interactions (see Table 1). Two types of interaction listed in the previous section were not considered in the analysis (namely the “other” interactions) and the student-tools and student-environment or manipulatives interactions were subsumed under the student-tools interaction.

<table>
<thead>
<tr>
<th>IMS LD unit-of-learning title</th>
<th>Short description</th>
<th>Interactions*</th>
<th>SS H</th>
<th>ST H</th>
<th>ST N</th>
<th>SI N</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECONSTRUCTIVISM</td>
<td>There are two roles, learner and teacher. Learners are to explore a project called Villa dall’Ava, and then create their own project. The teacher supports learners when using the resources and gives them feedback on their project.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>MODERN ARCHITECTURE</td>
<td>There are two roles, learner and teacher. Learners brainstorm their previous knowledge on modern architecture, then read some selected content and prepare a presentation on this content. The teacher supports learners when they use tools and resources. After learners’ presentations, the teacher gives feedback.</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>SKYSCRAPERS AND RESIDENTIAL</td>
<td>The unit of learning features three roles: “interested in architecture”, “interested in buildings” and “teacher,” but the teacher role was not assigned any activities. Depending on their assigned role, learners move through the unit of learning according to architects and their styles or according to famous buildings. Some reuse of learning objects and activities can be seen. Another special feature of the unit of learning is that it comprises two plays.</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>RESIDENTIAL—LEVEL A</td>
<td>The unit of learning features one role, which is a learner role. Learners first choose whether they wish to learn about skyscrapers</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>
or about residential homes. Each activity and interaction with learning content gives learners also a chance to reflect and summarize what they’ve learned. At the end, learners summarize everything they learned.

There are five roles. In addition to the teacher role, learners can either be Team A or Team B, and each team has an additional coordinator role. The teacher introduces the topic, then learners split into two groups A and B, around an assigned theme, for the project work phase. Each group selects a person to be the coordinator, who aggregates the group’s work. After group work, students answer a questionnaire. If they pass, the activity is finished. Otherwise, the group has to re-work its project to find the correct answers to all the questions. The teacher has the option to support students when needed.

There are two roles, learner and teacher. Learners select a topic from one of two lists, which the teacher provided. One list contains learning technologies; the other contains learning contexts. Students keep a blog on the chosen technology or context, and discuss with each other how to apply learning technologies in various learning contexts. In a final report, learners describe scenarios for applying learning technologies in specific learning contexts.

| SS<sub>H</sub> | … Student-student (human) interactions |
| ST<sub>N</sub> | … Student-tools and student-environment/manipulatives (nonhuman) interactions |
| ST<sub>H</sub> | … Student-teacher (human) interactions |
| SI<sub>N</sub> | … Student-information (nonhuman) interactions |

**Runtime player**

From the VLEs that are capable of interpreting IMS LD units of learning, the system that has been established the longest, the Service Based Learning Design player (SLeD, version 3.0) ([http://sourceforge.net/projects/ldplayer/files/SLeD%20Player/Sled%203.0/sled3.0.zip/download](http://sourceforge.net/projects/ldplayer/files/SLeD%20Player/Sled%203.0/sled3.0.zip/download)), was chosen for the analysis. The SLeD player is an enhanced version of the CopperCore LD player. The latter is also known as the player module of the CopperCore Run Time (CCRT) environment ([http://coppercore.sourceforge.net/documentation/ccrt.shtml](http://coppercore.sourceforge.net/documentation/ccrt.shtml)). While the CCRT-integrated player can be used to run an LD, the distribution page of CCRT clearly states that it was “not intended to be used in a real-life deployment.” The SLeD player, on the other hand, builds upon CCRT and enhanced the state of the art by “separating out the player functionality from the underlying engine” (McAndrew, Nadolski, & Little, 2005). SLeD is the only player to date that fully implements IMS LD based on CCRT and is freely available for educational and research purposes. To substantiate this, we performed pre-study trials with other available players. GRAIL requires a dotLRN host system and is visually similar in its setup to SLeD but architecturally more tightly integrated with the host VLE. We realize that an analysis using GRAIL may have provided different results than the ones presented in this paper. Trials with the other available players CLIX and the Astro Player showed that these systems did not support all features used in the units of learning used for the study, as presented in the previous sub-section. For instance, the Astro Player did not support the displaying of multiple activity descriptions, which is however possible in the IMS LD specification.

IMS LD players are usually set up in a similar manner. To access activities, participants click on one of the links in the navigation tree on the left (see Figure 1). When an activity is selected from the navigation tree, the content area of the browser window provides access to the activities’ descriptions and learning objectives. If the activities or learning objects contain several items, players can access them by selecting the activity or learning object in the navigation tree on the left and selecting 1, 2, 3, etc., within the main frame on the right. For instance, the currently selected learning object “Skyscrapers—Seagram Building” in Figure 1 contains three subordinate items as indicated by “1 2 3” under the Description tab.
Analysis procedure

Two IMS LD experts conducted the following analysis procedure independently. The analysis results were discussed and integrated in a face-to-face session. The interpretation of the results in the light of the research question tackled in this article was done by the authors. The analysis procedure proceeded as follows. First, the IMS LD conformant XML representation of the units of learning, including their packaged resources, were analyzed structurally in relation to the interactions for learning outlined in the previous section. That means, IMS LD elements used to provide these interactions were noted down. Then, one after the other, each unit of learning was played in the SLeD player. In each unit of learning each role was impersonated; all activities were performed in all possible permutations of the activity sequence through the unit of learning, and all included services and resources were accessed and used. While doing so we recorded how Reigeluth and Moore’s interactions for learning were expressed visually and/or metaphorically in the user interface of the SLeD player. The expressions of the interaction were then analyzed by the authors of this paper for the ability to support the interactions during runtime. The results of this analysis procedure are presented in the following section.

Results

The results are grouped by the interaction typology introduced before. “Student-other human” and “Student-other nonhuman” were excluded from further analysis as interactions since these were not featured in any of the units of learning selected for this study. These interactions, while representable with IMS LD, generally appear to be out of the scope of capturing and analysis in an IMS LD runtime player.

Student-student

DECONSTRUCTIVISM, SHARED OUTCOME, and BLOG COLLABORATION (Text in small caps refers to the units of learning presented in Table 1) featured student-student interactions. Learners in these units of learning could only identify a chance for student-student interaction if they received explicit instructions in the activity, and if the author of the unit of learning included corresponding services where the interaction takes place, such as a chat or forum. Students were at times asked to post an idea, link, or other information, but there were no instructions for working with or responding to one another. Next to the missing instructions for posting or interacting, it was not apparent who
else had access to and participated in the forum. What other students had access? Were forums meant for student access and discussion only, or did teachers have access as well? To support student-student interactions, it should be made clear with whom students are to interact. Most modern VLEs offer features to support awareness who has access to and who is contributing in a VLE. The SLeD player, a VLE by definition, does not display this information, although the IMS LD conceptual model would easily allow revealing it.

**Shared Outcome** featured four different learner roles: Team A, Team B, Team Leader A, Team Leader B. During the course of the unit of learning, the team members chose a team leader, who was then assigned this additional role. For a person assigned to the Team Leader A role, it was not clear in the user interface how and when to act as team leader and when to act as a regular member of team A (because the person likely acted in both roles). The player did not offer explicit support for this change of roles. Switching roles was not obvious because the player displayed all units of learning a person was currently assigned to in one drop-down box; the role the person took within each unit of learning was written in parentheses behind the unit of learning’s title.

**Student-teacher**

Student-teacher interactions were mostly realized using the IMS LD concept of support activities in the units of learning. In the player, this interaction was not well supported. The view offered to learners differed from the view offered to teachers: Each role saw only those activities that were assigned to this very role. **Deconstructivism**, **Modern Architecture**, **Shared Outcome**, and **Blog Collaboration** featured student-teacher interactions and thus encountered this problem. For illustration, see the screenshots below (Figure 2 and Figure 3), which were taken from the **Deconstructivism** unit of learning. As the two navigation trees indicate, there was no guidance for learners or teachers regarding when the teacher was to support the students during their activities. The player did not support the teacher in finding out about learners’ current status.

![Figure 2. Student navigation within Deconstructivism in SLeD](image)

![Figure 3. Teacher navigation within Deconstructivism in SLeD](image)

Having isolated views for each role’s activities created an artificial separation in the VLE. It required some explaining within the activity descriptions to overcome this separation, that is, to make clear that teachers and students were to work together at this point. This is a difference between currently available IMS LD players and non-IMS LD VLEs. In non-IMS LD VLEs, teachers and learners see the same interface in the learning environment, and all instructions in the interface are targeted towards learners, while the teacher implicitly assumes any support tasks. The SLeD player interprets IMS LD units of learning in a way that learners and teachers are presented with separate views onto the (same) unit of learning. Each role receives own activities with own instructions and is not aware of other roles and their activities. This makes it difficult to identify when the different roles will interact.

In the units of learning **Deconstructivism**, **Modern Architecture**, **Shared Outcome**, and **Blog Collaboration**, the teachers’ activities were IMS LD support activities. This special type of activity did not appear any different in the player than a learning activity. The teacher was not offered additional functions to execute this support activity and may thus not even know that s/he was involved in a support activity. What appears to create
further uncertainty is that in the IMS LD specification support activities are by definition not linked to another activity but to another role. The specification is unclear whether the indicated role is only valid during an act, during a play, or always. This leaves room for interpretation for player vendors and poses obstacles for student-teacher interactions.

**Student-tool and student-environment/manipulatives**

In a VLE, as opposed to a face-to-face learning environment, these two interactions are hard to distinguish because the line between the environment and its tools diminishes. These two interactions were thus observed together for the purposes of this analysis. Interactions with tools and environments can be regarded as interactions with the VLE itself and with provided resources as well as tools. Such resources or tools are sometimes external to the unit of learning, for example, external webpages. In the player, learners received no hint regarding what role they were assigned and what role they currently carried out (what the role is called and a possible description of the role). This information, however, influences the learners’ interaction with the tools and environment, and is thus important to display. For instance, the unit of learning SKYSCRAPERS AND RESIDENTIAL HOMES—LEVEL A features two different learner roles: “interested in architects” and “interested in buildings”. Learners had no idea that they were assigned one of these special foci in the unit of learning since the role’s title was not displayed. Acting in either one of the roles, however, affected how students interacted with the tools and the environment.

In the unit of learning SKYSCRAPERS AND RESIDENTIAL HOMES—LEVEL A, the nested activity structures are confusing and do not support navigational orientation. As can be seen from the screenshot of the unit of learning SKYSCRAPERS AND RESIDENTIAL HOMES—LEVEL A in Figure 4, keeping the overview of the different layers inside the activity structure could be quite hard. Adding to the structural complexity the activity structure “Selection: Architect Tracks” contained three nested activity structures, each containing additional activities and resources. In the unit of learning SKYSCRAPERS AND RESIDENTIAL HOMES—LEVEL B, navigation was further hampered because it was hard to distinguish between activities that were part of the activity structure “Sequence: Introduction into skyscrapers” and activities that were part of the regular sequence because they are nearly horizontally aligned (cp. Figure 5). Also, the distinction between activity structures of type sequence and of type selection should be better reflected in runtime interfaces so it is clear to learners what the difference is and what options for progression they have. In the screenshots, it might appear as if the SLeD player did actually make this distinction because “sequence” and “selection” were written in the navigation tree. In fact, unit-of-learning authors, not the runtime system, defined these titles; the distinction would not be clear without the explicit inclusion of “sequence” and “selection” in the activity structures’ titles.

**Figure 4.** Screenshot of the nested activity structures in the unit of learning SKYSCRAPERS AND RESIDENTIAL HOMES—LEVEL A

**Figure 5.** Screenshot of the nested activity structures in the unit of learning SKYSCRAPERS AND RESIDENTIAL HOMES—LEVEL B
Next to the nested activity structures, some activities included additional activity descriptions or learning objects, which in turn had further nested objects. For instance, the learning object “Explore content about early skyscrapers and major buildings” (cp. Figure 4) provided access to 12 subordinate objects (note that the actual subordinate objects were only displayed in the main frame and are not depicted in the navigation tree shown here). Some of these objects again featured links to other sites. Learners would have to navigate through up to five levels. The student-environment interaction was greatly impeded through this complicated setup. Although IMS LD supports the use of nested structures, one of the goals of VLEs should be to ease navigation by modifying the display of activities and learning objects.

**Student-information**

The difference between the information and the environment was not easily drawn because the information is highly embedded in the environment. For the purposes of this analysis, student-information interactions were interpreted as the interactions taking place with activities or learning objects. From a learner’s point of view, it was hardly understandable what the difference between learning objects and activity (descriptions) was in the player interface. The activity description stated to “do something” but the place to do this was not immediately provided where the activity description was displayed. The separation between activities, whose description was displayed in the main frame on the right, and the referenced learning objects, which were only accessible via the navigation tree on the left, introduced an additional layer of complexity. The use of activities and learning objects as seen in the units of learning under analysis suggests that IMS LD activities were at times used as an additional, yet nearly useless, container for referencing learning objects. If used this way, the line between the IMS LD concepts’ activity and environment (with contained learning objects) diminishes. This may also have been an inspiration to the developments of Simple Learning Design 2.0 (Durand & Downes, 2009), where only one of the concepts (the activity) is included, and the concept environment is not present at all.

**Interaction services**

The investigation so far has shown that interaction between learners and other human or nonhuman entities is an important pedagogical aspect of VLEs. IMS LD as a prominent interoperability specification should therefore be able to express such collaboration and communication interactions. This section discusses in more detail how this can be achieved in the runtime perspective.

In the IMS LD information model (IMS Global, 2003), activities take place in environments, which can contain learning objects, services, and further environments. A service element in an IMS LD unit of learning is an abstract declaration of the required facility or tool, for example, a synchronous conferencing service. During runtime, the IMS LD player will create and offer an instance of the service for use within the environment. Since different services like discussion forum, chat, announcement, etc., offer different functionalities and require different steps and parameters during setup and use, there is a dedicated description schema for each service. When the IMS LD specification was conceived, the authors decided to include the four “most widely implemented and used services” (Olivier & Tattersall, 2005, p. 32) into the specification. These are send-mail (for emailing), conference (synchronous, asynchronous communication, and announcements), monitor (observing property values), and index-search (searching in the unit of learning). From the teaching practitioner’s point of view, this limitation comes with drawbacks: The creation of units of learning with intense virtual communication and collaboration as in the study presented in the previous section may become problematic since the only services that are available for this purpose are send-mail and conference. There is no service description schema in IMS LD for currently popular applications supporting collaboration and communication like blogs and microblogs (Twitter), wikis (Wikipedia), shared virtual calendars, document sharing (Google Drive), polling (Doodle), and so forth. While these are gradually being integrated as external services in popular VLEs like Moodle, IMS LD (conceived a decade ago) was not built to cater to these new options.

A simple workaround that does not require any modification or extension of the IMS LD information model would be to provide a link to or description of the required service as well as the instructions for setup and use of the service in the IMS LD activity description element. There are several potential problems with such an approach. For one, the external service likely requires login with a separate account (or signing up for a new one). There is currently no way

339
to map the user’s account in the service with his/her account in the IMS LD player. Also, different users have different roles within the same activity/environment (e.g., moderator and participant in discussions). These roles need to be mapped to the roles offered by the service (if any), but the mapping cannot be enforced because there is no protocol for transferring any data between the IMS LD player and the external service. Additionally, the design-time provision of a static hyperlink to a service instead of an abstract service description means that different runs of a unit of learning will use the same instance of the service.

Other approaches to using services that are not offered by IMS LD potentially require a modification to the IMS LD information model and/or a customized interpretation of the existing XML binding and will therefore neither be portable to nor runnable on all existing IMS LD players. In recent years, several approaches have been proposed in this area. The perceived trend upon which these developments build is that of a move towards personal learning environments that are composed of small, autonomous, loosely coupled applications, for instance based on widget technology (cf. Wilson, Sharples, & Griffiths, 2009). Another avenue to broaden the spectrum of available interaction services is based on generic service integration (de la Fuente Valentín, Miao, Pardo and Delgado Kloos, 2008), which supports the semi-automatic identification appropriate services at runtime based on a design-time description of what a service is supposed to achieve during runtime, rather than describing which specific service to use.

As it currently stands, the IMS LD specification and, therefore, most runtime players offer only two services that can be used for collaboration and communication: the send-mail service for email communication and the conference service for synchronous and asynchronous communication/collaboration and one-way communication. The concrete appearance and handling of these services is left to the player. Unfortunately, there is currently no agreed way of providing richer communication and collaboration services in IMS LD. In the light of the large number of interaction services available and used for educational purposes, this shortcoming will continue to diminish the role of IMS LD.

Discussion and conclusion

The task of IMS LD compliant VLEs is to interpret a unit of learning so that the learning environment is set up and the learning process is best supported. The VLE must work with the information provided in the IMS LD unit of learning. What information can or should be derived from this unit of learning at runtime to support interactions for learning? In this paper, we explored answers to this question in an IMS LD runtime environment, namely the Service Based Learning Design (SLeD) player, and selected and adopted interactions for learning from Reigeluth and Moore’s (1999) set of pedagogical aspects. We demonstrated that the expression and implementation of interactions of learners with human actors and nonhuman elements in the learning environment are currently far from optimal and far from fit for current pedagogies and diverse communication channels during learning experiences. Crucial issues encountered include no explicit link between activity descriptions in the main content area and the environment objects in the navigation area, unless the author provides a descriptive linkage, which would however contradict the basic idea of IMS LD by requiring knowledge of the runtime user interface at design time. The first recommendation to developers, therefore, is to provide in-place access to information pertaining to the current activity, which would enhance student-information interactions by integrating the activity description with the learning object more closely in the user interface or by providing instant previews common in popular VLEs like Moodle.

The second recommendation is that items that conceptually belong together be displayed together in the user interface to facilitate the learners. This is the case for activities and environments that are linked to one another as well as for their subordinate items, such as multiple activity descriptions and multiple learning objects. Currently, the display of conceptually connected items is separated in the observed runtime environment: The participant has to switch back and forth between the navigation tree and the main frame to get the wanted items to display. The hierarchical nesting as the display metaphor of all unit-of-learning content/activities additionally impedes easy navigation since there is little learning-process-related information contained in such a structure. The third recommendation, therefore, is that IMS LD player developers devise a meaningful depiction of nested activity structures in the user interface to guide learners in progressing along the learning path. Although we don’t know why, this kind of nested navigation structure is a visual design flaw present in all IMD LD players to some extent. Lambe acknowledges that nested depiction of the unit of learning’s structure becomes hard to navigate beyond three levels of depth (2007).
We revealed that it is not clear at runtime which role or roles the user is currently impersonating, and there are missing visual cues on supported and/or collaborating roles for the current activity. Thus, the fourth recommendation is that roles should be explicitly displayed in the user interface. Also, the runtime system does not indicate places to interact with other roles (even if designed that way), such as when and how does a supporter actually support learners? Missing indication of interaction and awareness of other participants was also the case with any type of conference service included in the units of learning. Participants had no idea who was interacting with them within the conference service or widget. The fifth recommendation, therefore, is that the runtime system make participants aware of who is interacting (or meant to interact) with them during an activity or in an environment.

The concluding overall recommendation in light of these issues is that developers of IMS LD VLEs take more caution in observing the user perspective. Instead of exposing the XML hierarchy defined in the IMS LD unit of learning to the users, runtime environments should display units of learning in a way that lessens the cognitive load currently needed for navigating and understanding how the different parts of a unit of learning relate to each other. Thus far, many crucial elements that set IMS LD apart from other learning-technology specifications — such as roles, activities, and environments — are not expressed with their full potential in the runtime. Especially a lack of support regarding human-to-human interactions can be observed. Currently, the responsibility lies with the author of a unit of learning to make instructions for interactions explicit at design time, requiring prior knowledge about how the unit of learning will be rendered at runtime. The runtime environment did not offer support in this regard as the technical differentiation of concepts such as activities, environments, or activity structures hardly affected their display in the runtime’s interface. Rather, there appeared to be a crucial connection between decisions the unit-of-learning author made at design time and their influence on the appearance of a unit of learning in the runtime system. Not knowing where the unit of learning will be deployed, however, may serve for poor design decisions.

This paper offered bits of advice to IMS LD VLE developers about what pitfalls to avoid in terms of student interactions. In future work, the expression of additional relevant pedagogical aspects and additional IMS LD players need to be taken into account. Understanding how pedagogy is expressed visually and navigationally in units of learning at runtime helps to identify shortcomings in the way runtime environments have been conceived to date. Building on knowledge about these shortcomings, work can then proceed towards interoperable runtime support of interactions for learning based on evolving pedagogies and contemporary web-based interaction technologies.

Acknowledgments

This work was partially funded by the European Commission in the ICOPER network (ECP-2007-EDU-417007); we are grateful to ICOPER partners for providing the units of learning. MD acknowledges further support by the European Commission through the TEL-Map support action (FP7-257822) and the Lifelong Learning Programme multilateral project METIS (531262- LLP-2012-ES-KA3-KA3MP).

References


Practising Arithmetic Using Educational Video Games with an Interpersonal Computer

Vagner Beserra¹*, Miguel Nussbaum¹, Ricardo Zeni², Werner Rodriguez³ and Gabriel Wurman¹

¹Pontificia Universidad Católica de Chile, Santiago, Chile // ²Universidade Estadual Paulista, Guaratinguetá, São Paulo, Brazil // ³Universidade de Costa Rica, San José, Costa Rica // vagner.beserra@gmail.com // mn@ing.puc.cl // jrzeni@feg.unesp.br // werner.rodrigues@ucr.ac.cr // gabow4@gmail.com

*Corresponding author

Abstract
Studies show the positive effects that video games can have on student performance and attitude towards learning. In the past few years, strategies have been generated to optimize the use of technological resources with the aim of facilitating widespread adoption of technology in the classroom. Given its low acquisition and maintenance costs, the interpersonal computer allows individual interaction and simultaneous learning with large groups of students. The purpose of this work was to compare arithmetical knowledge acquired by third-grade students through the use of game-based activities and non-game-based activities using an interpersonal computer, with knowledge acquired through the use of traditional paper-and-pencil activities, and to analyze their impact in various socio-cultural contexts. To do this, a quasi-experimental study was conducted with 271 students in three different countries (Brazil, Chile, and Costa Rica), in both rural and urban schools. A set of educational games for practising arithmetic was developed and tested in six schools within these three countries. Results show that there were no significant differences (ANCOVA) in the learning acquired from game-based vs. non-game-based activities. However, both showed a significant difference when compared with the traditional method. Additionally, both groups using the interpersonal computer showed higher levels of student interest than the traditional method group, and these technological methods were seen to be especially effective in increasing learning among weaker students.

Keywords
Cross-cultural projects, Intelligent tutoring systems, Shared display, Interpersonal computer, Arithmetic practice, Educational games

Introduction
Playing is, above all, a learning experience (Rosas, Nussbaum, Cumsille, Marianov, Correa, Flores et al., 2003). The use of computer games favors the development of complex thinking skills related to problem solving (Shih, Shih, Shih, Su, & Chuang, 2010), strategic planning (McFarlane, Sparrowhawk, & Heald, 2002), and self-regulated learning (Mayo, 2009). Computer games can also support different learning styles (Connolly & Stansfield, 2007), since speed and level of difficulty can be adjusted according to the player (Alcoholado, Nussbaum, Tagle, Gomez, Denardin, Susaeta et al., 2012).

In the past few years, parallel to the development of educational video games, strategies to optimize the use of technological resources have been developed with the goal of facilitating wide-scale adoption of technology in classrooms. In particular, the interpersonal computer stands out because of its low acquisition and maintenance costs (Kaplan, DoLenh, Bachour, Kao, Gault, & Dillenbourg, 2009). With an interpersonal computer, multiple users located in the same space share one output device, like a computer screen, but each user has their own input device that they use to interact simultaneously with the virtual world.

The interpersonal computer is very attractive for schools in developing countries, where computational infrastructure is an entry barrier (Trucano, 2010). Cost is a key element in the widespread adoption of technology in classrooms, which is the main reason why the interpersonal computer is such an attractive proposal: it centralizes resources by minimizing the amount of equipment and technical support required.

The use of multiple inputs has been studied by a number of researchers who have sought to demonstrate its effects on peers working with a single screen (Paek, Agrawala, Basu, Drucker, Kristjansson, Logan et al., 2004). The
interpersonal computer bolsters the learning process when teacher and student are in the same physical space, since the technology does not just capture student attention and motivate them, but also significantly mediates the construction of concepts (Smith, Gentry, & Blake, 2012). The results show that children controlling their own input devices in a situation with shared screens are more active and therefore demonstrate less boredom and fewer disruptive attitudes (Infante, Weitz, Reyes, Nussbaum, Gómez, & Radovic, 2010). A fundamental aspect that favors interactivity among the students, and particularly their level of motivation, is the fact that the activity makes each of the students work with their own objects; each student controls their own input device, which forces them to participate and become protagonists of their own learning (Infante, Hidalgo, Nussbaum, Alarcón, & Gottlieb, 2009).

When teaching arithmetic, Alcoholado et al. (2012) show that engaging in interactive practice on interpersonal computers facilitates learning more than using the traditional paper and pencil method. This suggests that this technology not only has economic advantages, but is also an effective educational tool.

Consistent with the above, and understanding the potentialities of using video games in the teaching-learning process, our first research question is: Within individual pedagogical activities in which arithmetic is practised using an interpersonal computer, what is the added value of a game in terms of student knowledge gained when compared with the same non-game-based activities and traditional paper-and-pencil activities?

On the other hand, according to Zaharias and Papargyris (2009), culture is a potentially important factor in deciding performance and user satisfaction with video games. Culture is also one of the factors that influence user preferences, according to Ramli, Zin, and Ashaari (2011). How to play and how to solve a video game challenge are activities commonly influenced by a player’s background and environment (Fang & Zhao, 2010), which are in turn considered cultural values (Schumann, 2009). Thus emerges our second research question: Is culture a factor that influences learning and/or student interest when interacting with game-based activities on the interpersonal computer?

In the following section, we present a video game that makes use of an interpersonal computer to practise basic mathematics and study its performance compared to a non-game-based activity in three different cultural contexts. The experimental design is presented in Section 4, the results obtained in Section 5, closing with conclusions.

**Game-based pedagogical mathematics activities for the interpersonal computer**

**Interpersonal-computer structure**

Each student is assigned a cell on a shared screen (Figure 1), where they work individually, with no possibility of leaving the assigned cell. Through data persistence the system stores a record of each stage that a student completes. This allows each student to work at their own pace over various sessions. To facilitate the process of identification, each student is assigned a unique symbol, which appears on the screen (Alcoholado et al., 2012).

![Figure 1. Shared screen](image-url)
Considering that teacher mediation is essential, the system provides instant feedback of the state of each student via the same shared screen, which allows the teacher to identify which students need help and how they are doing. The teacher can support a specific student by moving his cursor freely across the screen to the workspace of said student. When the symbol identifying that student is chosen, it will show the student’s name. This enables the teacher to interact with the student, either by using their cursor to guide the student on their individual screen or by working face-to-face with the student to help with the detected problem. It is the teacher’s responsibility to review the student’s incorrect responses.

**Game-based pedagogical design**

The video game, designed to help students practise arithmetic, shows the student exercises according to their rate of progress. Given that its goal is to reinforce the student’s mathematical knowledge and skills, this video game was designed to fulfill the characteristics of a test educational game (Li, 2012). It occupies a rule-based system: 18 for addition, 18 for subtraction, 13 for multiplication, and 16 for division (Alcoholado et al., 2012). The order of these rules follows the sequence defined by the Chilean (Mineduc, 2009), Brazilian (MEC, 1997), and Costa Rican (MEP, 2005) curricular frameworks, respectively. The number of exercises assigned for each rule depends on the skill level of each student. In order to move on from a concept, that is, a rule, the student must successfully solve at least ten consecutive exercises or eight exercises out of at least fifteen, with the last three exercises being correct (Alcoholado et al., 2012).

According to Sweetser and Johnson (2004), a video game’s narrative is important in attracting players to the game and keeping them immersed, since it provides the players with a plot and background and makes them feel like part of the story. The game’s narrative, used as much to develop the story (Qin, Rau, & Salvendy, 2009) as to solve problems, can be divided in two parts:

- **Story**—At the beginning of the activity, the teacher presents the story and general objective to the students (in a magical country a wizard stole their magical rocks, burying them deep in the ocean and forming small islands; the challenge is to explore the different islands with the goal of finding the stolen magic rocks and re-establishing geographic order.)

- **Specific objectives**—The objectives of the game should be transmitted clearly and directly to the player (Pagulayan, Keeker, Wixon, Romero, & Fuller, 2003). As the students explore the islands, they encounter various challenges represented by six different games related to the main story (Figure 2).

![Figure 2. Pedagogical Activities](image-url)
Each individual’s game space is composed of seven elements (Figure 3):

- Identification: Each player is assigned a unique colored symbol to identify them.
- Equation: Area displaying the mathematical equation that the student must complete.
- Player’s pointer: Represents each child’s cursor, which can only move within its own cell.
- Mechanics of the games: The goal of the game is for the child to choose the correct answer from a set of alternatives. For example, in Figure 3, the child had to choose the correct answer, 10, for the stated problem: 8 + 2. By maintaining the same simple mechanics for all of the games, we ensured that students needed only minimal skills to be able to play.
- Game rewards:
  - Immediate results reward
    Players should always be able to see their results and their progress in the game with immediate access to all information regarding each player’s actions. (Fu, Su, & Yu, 2009). This is achieved by using three animations that indicate when an exercise is answered successfully (Figure 4a), when a level of the game is completed successfully (Figure 4b), or when an exercise is answered incorrectly (Figure 4c). In this last case, the question can be re-submitted until the student responds correctly.
  - Immediate progress reward
    To improve their mastery of the game, feedback is given to represent the child's progress (Federoff, 2002; Fu, Su, & Yu, 2009). After each player’s action, the game displays a positive or negative reward so that the player can observe their own progress and see how close they are to reaching the goal (Fu, Su, & Yu, 2009). For example, in Figure 3, progress in the game is indicated by increasing or decreasing pieces of the bridge in relation to whether the player answered correctly or incorrectly. In Figure 2, “Capturing Fish,” the fish either jump into the boat (correct answer) or fall out of the boat (incorrect answer). The level ends when the student accumulates seven pieces of the bridge in the first example or catches seven fish in the second example.
  - Accumulated reward
    The total rewards earned in each session are awarded for each correct answer (coins) and for each completed level in the game (medals). For example, in Figure 3, 89 coins have already been accumulated.

- Pedagogical feedback: Information mainly for the teacher. Represents the number of questions answered in the session, with colors indicating achievement: green indicates 66.6% or more correct answers, yellow indicates 33.3%–66.6% correct answers, and red indicates 33.3% or less correct answers.
- Alert: In the case of a child who does not take part in the activity, that is, is inactive for more than 120 seconds, a sleep symbol is displayed in their cell (Figure 4d).
At the beginning of the first session, each student chooses their avatar, i.e., their character for the game, which, according to Annetta (2010), allows the players to identify themselves as individual members of the community of game-players. There are three types of avatars available: male, female, and animal (Figure 5), all drawn in the style of a children’s cartoon. This style was chosen following a study of 45 school children from the same grade level as the students in this project. Findings showed that 90% of children preferred the cartoon style to a more realistic one, while 80% preferred a more child-like aesthetic to an adolescent one.

![Figure 4](image)

*Figure 4(a) Positive feedback, (b) Game progress feedback, (c) Negative Feedback, and (d) Inactivity*

![Figure 5](image)

*Figure 5. Avatars*

Considering the reduced size of each cell, and taking into account the suggestions made by Federoff (2002), the interface was developed to be as simple and unobtrusive as possible to facilitate use of the video game and reduce unnecessary cognitive processing (Pilke, 2004). The more intuitive the video game-player interaction is, the greater the possibility that “flow” will exist.

A video game is adequate when it creates flow, which allows a child to maintain interest in the game (Csikszentmihalyi, 1990; Annetta, 2010). An educational video game should strike a balance between fun and educational value (Prensky, 2003). The teacher is in charge of regulating the game-based dimension, since the child progresses through the different levels based on the teacher’s observations of the student’s needs, while the pedagogical requirements are regulated by the previously established system of rules. In order to do so, the teacher must decide which of the secondary narratives (i.e., the various games) to choose and when.

**Non-game-based pedagogical activity for practising arithmetic using the interpersonal computer**

The non-game-based individual interactive pedagogical activity using the interpersonal computer (Figure 6) allows up to 49 students in the classroom to work at their own pace, simultaneously, using a PC, a projector, and one mouse per child. Both the game-based and non-game-based activities were designed for practising arithmetic; both generate activities in line with the progress of each student by using the same intelligent tutoring system (Alcohohado et al., 2012). Each of the non-game-based applications’ individual spaces consists of six elements (Figure 7):

- **Equation**: Area showing the mathematical equation that the student must complete.
- Answer construction: The child must assemble each digit by increasing or decreasing the indicated number.
- Player’s pointer: Represents each child’s cursor, which can only move within their own cell.
- Identification: A unique colored symbol is assigned to every player. This serves two purposes: identifying the student’s work area, and serving as a button with which to confirm the final answer.
- Individual answer feedback: Once a student has confirmed their answer, feedback appears in the middle of their cell. There are three types of feedback: correct answer (Figure 6, column 2, row 4), incorrect answer (Figure 6, column 1, row 2), and medals that correspond to the completion of the level and appear once the last question in the level is correctly answered (Figure 6, column 3, row 5).
- Alert: If a student’s mouse remains motionless for 60 seconds, a symbol denoting inactivity appears (Figure 6, column 2, row 1). If this continues for more than 120 seconds, the cell’s background turns the same color as the symbol (Figure 6, column 2, row 3).

![Figure 6. Shared screen from the non-game-based application](image)

As with the game-based version, this application uses data persistence to store the point reached by each student once their respective session draws to a close. This allows each student to work at their own pace throughout the different sessions. The status of each student is also visible, (Figure 6, last column in the right) enabling the teacher to mediate either individually or across the entire group, as they deem necessary.

![Figure 7. Individual space from the non-game-based application](image)
Experimental design

We designed a quasi-experimental study about teaching arithmetic to third-grade students in order to answer our research questions. Three types of groups were created, one per class, based on the type of technology used to practise arithmetic, as seen in Table 2. The first group used the game-based version (Figure 3), the second group used the non-game-based version (Figure 7), and the third group used the traditional method (with paper exercises). Because not all of the participating schools had three third-grade classes, the game-based version was prioritized over the non-game-based version and the latter over the traditional method.

In order to respond to our second research question, regarding the influence culture has on knowledge acquisition and student interest in a technology-mediated learning process, we must first note that we are operating under the UNESCO definition (2001): “Culture should be regarded as the set of distinctive spiritual, material, intellectual and emotional features of society or a social group, and that it encompasses, in addition to art and literature, lifestyles, ways of living together, value systems, traditions and beliefs.” From this broad definition, we have identified two characteristics that affect the interaction between a social group and mediating technology: 1) socioeconomic and rural/urban context and 2) exposure to technology.

In regards to socioeconomic context, three Latin American cities were selected for the study: Santiago, Chile; Guaratinguetá, Brazil; and San Ramón, Costa Rica. In each of these cities, both a rural and an urban school were selected. However, this was not possible in Brazil, since in the rural location there were no schools with at least two third-grade classes representative of the region. Table 2 indicates the class characteristics of all participating groups. These countries were selected based on their varying experiences with the aforementioned characteristics (ECLA, 2010; ITU, 2013).

Table 1 illustrates how the participating schools from rural areas stand out for their prevalence of agricultural labor, whereas the urban areas mainly depend on service industry jobs. There were computer laboratories in all the urban locations (and in the rural Chilean school), but only enjoyed frequent use in the cases of Chile and Brazil. In this context, we considered the existence of computer labs an indicator of the pervasiveness of technology. However, the technology required for the experiment was provided by the authors. As sampling bias can affect the quality of an experiment’s results, care was taken that each of the six participating schools fell within the average range for national educational quality.

<table>
<thead>
<tr>
<th>Table 1. School and location data</th>
<th>Chile</th>
<th>Brazil</th>
<th>Costa Rica</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rural</td>
<td>Urban</td>
<td>Rural</td>
</tr>
<tr>
<td># of students in the school</td>
<td>600</td>
<td>1690</td>
<td>120</td>
</tr>
<tr>
<td># of classes at the school</td>
<td>18</td>
<td>47</td>
<td>6</td>
</tr>
<tr>
<td>Computer laboratory</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Frequency of computer laboratory use (per week)</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Professional profile</td>
<td>Farmworkers</td>
<td>Workers and housewives</td>
<td>Farmworkers</td>
</tr>
</tbody>
</table>

For every school, a total of eight sessions of approximately 45 minutes each were conducted, distributed over time according to the school’s availability (between three and six weeks). The study was broken into eight experimental sessions in order to approximate the context in which students practise arithmetic; in the participating schools most students have close to eight arithmetic practice sessions in a one-month period. In all the participating classes, a pre-test was applied at the beginning of the experimental study. This consisted of an individual evaluation of each student’s previous knowledge of mathematics, for a maximum of 45 minutes. This evaluation contained 45 addition, subtraction, multiplication, and division exercises designed to identify the third-graders’ skills (Alcoholado et al.,
At the end of the eight sessions, a post-test was conducted with the same instrument and under the same conditions. In order to ensure the reliability of these (pre- and post-test) instruments, they were analyzed for each of the participating classes according to the Cronbach’s alpha criteria, Table 3. According to Bland and Altman (1997), a value of higher than 0.6 for Cronbach’s alpha indicates that the applied test is acceptable for classifying students, based on the content delivered. Furthermore, these instruments were carefully reviewed by each of the teachers whose classes participated, so as to confirm both the validity of the pedagogic content and its effectiveness in terms of measuring knowledge acquired.

### Table 2. Class characteristics

<table>
<thead>
<tr>
<th>Type</th>
<th>Chile Rural</th>
<th>Chile Urban</th>
<th>Brazil Rural</th>
<th>Brazil Urban</th>
<th>Costa Rica Rural</th>
<th>Costa Rica Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students in the class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game-based</td>
<td>36</td>
<td>32</td>
<td>28</td>
<td>25</td>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td>Non-game-based</td>
<td>(15, 13)</td>
<td>(12, 13)</td>
<td>(18, 7)</td>
<td>(14, 11)</td>
<td>(9, 11)</td>
<td>(12, 8)</td>
</tr>
<tr>
<td>Participating students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game-based</td>
<td>25</td>
<td>25</td>
<td>22</td>
<td>19</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>Non-game-based</td>
<td>(10, 9)</td>
<td>(10, 9)</td>
<td>(10, 11)</td>
<td>(9, 8)</td>
<td>(7, 9)</td>
<td>(10, 8)</td>
</tr>
<tr>
<td>Hyperactive student ratio</td>
<td>0.08 0.06</td>
<td>0.09 0.06</td>
<td>0.06 0.08</td>
<td>0.09 0.11</td>
<td>0.21 0.13</td>
<td>0.04 0.08 0.12</td>
</tr>
<tr>
<td>Mean student attendance</td>
<td>7.26 6.6</td>
<td>- 7.66</td>
<td>7.55 6.85</td>
<td>- 6.74 6.81</td>
<td>6.39 -</td>
<td>6.17 6.05</td>
</tr>
</tbody>
</table>

The participants, both girls and boys (indicated in the second row of Table 2, with the number of girls being the first number in the pair), were taken from third-grade classes and ranged in age from eight to ten years old. The discrepancy between the number of students in each class and the number of students that participated (in the experiment) was caused by the way “participant” was defined; only those students who took the two assessments (pre and post) were considered participants.

This experiment also took into account the presence of hyperactive students, classified as such by their respective teachers. These students stand out from the rest of their classmates based on their high activity levels within the classroom. The hyperactive student number in Table 2 indicates the hyperactive student/total student ratio in each class. This information is relevant to the study because hyperactivity can influence the behavior of the rest of the class and lead to disruptions.

To ensure that the intervention was similar for all the students, we calculated the mean student attendance (Table 2), which corresponds to the average number of sessions attended by the students in each class. This value was obtained from the data recorded by the system during each session. The difference between the minimum and maximum number of sessions attended by the students in each of the groups was not significant.

The longest the experiment lasted was six weeks, in the rural Chilean school, because of the number of national holidays during the experimental period. The shortest time was three weeks, which occurred in the urban Costa Rican school, due to the restrictions imposed by the school that hurried the conclusion of school activities. In the rural Costa Rican school, from the fifth session onwards, a significant number of students did not attend the activities due to the start of the coffee-harvesting season. This made it impossible to apply the post-test, which impeded the non-game-based experiment from being fully concluded. Interestingly, however, at the same rural school in Costa Rica where the non-game-based experiment was cancelled, it was possible to carry out the game-based experiment in conditions similar to the urban Costa Rican school. The staff from all the classes indicated that in general, attendance was above average compared to a regular day.

Quantitative and qualitative data collection for this study was done in three different ways: through paper instruments (pre-test, post-test, and questionnaires); through information regarding work the students completed in both technology groups, which was stored by the system; and finally, through in-person observations of the students’ behavior and interactions. This was handled by three observers in each country, who used the same observation
Results

Quantitative results

At the end of the experimental study, a comparative analysis was conducted with the data collected by the instruments described above. The results of the pre- and post-test are shown in Table 3, together with the t-test and the increment between groups (Δ% group). The latter corresponds to average student progress for each group, calculated to evaluate the difference between the pre- and post-test results.

For all samples, the applied pre-tests obtained a Cronbach’s alpha greater than 0.74, except for the rural Brazilian group, which only reached 0.52 and was therefore removed from the sample.

<table>
<thead>
<tr>
<th>Country</th>
<th>Location</th>
<th>Type</th>
<th>t-Test</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Δ% Group</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>Rural</td>
<td>Game-based</td>
<td>&lt; 0.001</td>
<td>26.70</td>
<td>6.85</td>
<td>30.26</td>
<td>8.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-game-based</td>
<td>&lt; 0.0001</td>
<td>28.24</td>
<td>9.58</td>
<td>31.68</td>
<td>9.91</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>Traditional method</td>
<td>&lt; 0.001</td>
<td>12.35</td>
<td>5.99</td>
<td>14.39</td>
<td>6.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-game-based</td>
<td>&lt; 0.0001</td>
<td>13.64</td>
<td>5.31</td>
<td>16.36</td>
<td>4.67</td>
</tr>
<tr>
<td>Brazil</td>
<td>Urban</td>
<td>Traditional method</td>
<td>&lt; 0.001</td>
<td>29.83</td>
<td>9.68</td>
<td>33.78</td>
<td>8.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Game-based</td>
<td>&lt; 0.001</td>
<td>23.06</td>
<td>9.48</td>
<td>27.06</td>
<td>9.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-game-based</td>
<td>&lt; 0.001</td>
<td>28.19</td>
<td>9.64</td>
<td>32.63</td>
<td>9.24</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Rural</td>
<td>Game-based</td>
<td>&lt; 0.0001</td>
<td>18.94</td>
<td>5.68</td>
<td>23.71</td>
<td>5.11</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>Traditional method</td>
<td>&lt; 0.00001</td>
<td>19.15</td>
<td>8.98</td>
<td>22.25</td>
<td>8.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Game-based</td>
<td>&lt; 0.001</td>
<td>17.06</td>
<td>8.56</td>
<td>20.50</td>
<td>8.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-game-based</td>
<td>&lt; 0.001</td>
<td>19.67</td>
<td>5.64</td>
<td>23.62</td>
<td>7.40</td>
</tr>
</tbody>
</table>

Statistical significance was measured for each group by the t-test, which found significant results between the pre- and post-tests, as shown in Table 3. A one-way analysis of covariance (ANCOVA) was undertaken to compare the effectiveness of the various types of work groups (game-based, non-game-based, and traditional method). This analysis indicated a significant difference between the three groups ($F = 3.292, p < .038$). Upon analyzing the ANCOVAs of the paired off groups in Table 4, we observed a significant difference between the game-based and traditional method groups and the non-game-based and traditional method groups, but found no significant difference between the game-based and non-game-based groups. In Table 4, we indicate in parentheses the sample size, degrees of freedom, effect size, and the differences between the adjusted averages obtained from the paired-off groups’ ANCOVAS, in cases of significant difference. We conclude that both technologies had a similar impact, with a significant difference arising in the traditional method group in terms of the student knowledge acquired.

<table>
<thead>
<tr>
<th>Group</th>
<th>Non-game-based</th>
<th>Traditional method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game-based</td>
<td>$F = 0.968, \ p &lt; .326$ (191, 189, -, -)</td>
<td>$F = 3.865, \ p &lt; .048$ (170, 167, 0.22 small, 3.185)</td>
</tr>
<tr>
<td>Non-game-based</td>
<td>-</td>
<td>$F = 6.723, \ p &lt; .010$ (147, 144, 0.30 small, 4.887)</td>
</tr>
</tbody>
</table>

Given that there was not a significant difference between the two technology groups, we aimed to analyze the learning impact based on previous knowledge. To do so, students from both experimental groups (game-based and non-game-based) were separated according to whether their performance in the pre-test placed them in the top or bottom half of their group, thus resulting in two groups: top-half and bottom-half. A one-way analysis of covariance (ANCOVA) indicated a significant difference between these groups ($F = 6.150, \ p < .014$). By analyzing the
difference between the two groups’ adjusted pre- and post-test averages, we observed a difference of 7.813 points in favor of the bottom-half group over the top-half group, when compared to the average progress of both. As such, we conclude that the learning impact was significantly greater among students who were initially in the less knowledgeable half of the group.

Finally, the impact of the technological resource’s effectiveness regarding knowledge acquired in each one of the studied countries (Brazil, Chile, and Costa Rica) was investigated as a way of measuring cultural differences. A one-way analysis of covariance (ANCOVA) indicated a significant difference between the three countries \( F = 6.251, p < .002 \). This result has been repeated to analyze the ANCOVA of the three countries paired off against each other (Brazil/Chile \( F = 4.961, p < .027 \); Brazil/Costa Rica \( F = 4.019, p < .048 \); Chile/Costa Rica \( F = 7.603, p < .006 \)). The differences between the adjusted ANCOVA averages of three paired-off countries indicate that the impact of student knowledge acquired was greater for children in Brazil, followed by Costa Rica, and finally Chile (Brazil/Chile—7.144; Brazil/Costa Rica—1.249; Costa Rica/Chile—5.894). It should be noted that these findings are coherent with the average difference between each country’s pre- and post-test, Table 3.

In all three countries, the technological groups (game-based and non-game-based) consistently obtained better results than the traditional method groups. For all the technological groups, the pre-test results had a −0.85 correlation with the final results (Table 3, Δ% group); the lower the scores obtained on the pre-test (Table 3, pre-test) by the group, the higher their scores post-test, which indicates that the application was more effective for the weaker groups.

Graph 1 illustrates each child’s progress by the end of the experience, measured in pedagogical rules (Y axis), for each one of the schools studied (X axis). Three values are indicated for each school: the greatest is the average of the 25% of those students who, at the end of the experiment, had made the most progress in the number of pedagogical rules, while the lowest corresponds to the average of the 25% of students who made the least progress in the pedagogical rules. The value located between the previous two values (next to a triangle) corresponds to the median of this group. As the results in Table 3 illustrate, the medians of the game-based and non-game-based classes of comparable pairs are similar. Looking at the totality of the medians, there exists a correlation of 0.77 with the pre-test; the higher the pre-test score, the greater the class median.

The time required to learn how to use each pedagogical activity, whether game-based or non-game-based, is calculated by the number of arithmetical operations successfully applied per session, and is directly related to the number of pedagogical rules undertaken per session. Graph 2 shows the progress at the end of the first session.
measured in pedagogical rules (Y axis), for each one of the participating educational institutions (X axis). The format of the graph used is the same as Graph 1. We observed that, in comparative contexts, that is, between urban school classes or rural school classes within the same country, the students using the game-based application progressed in the lower to medium range, which is similar to those using the non-game-based application. The exceptions being urban Brazil, which presented the greatest difference between the experimental groups’ pre-tests, and urban Chile, which presented the lowest pre-test scores. The case of Brazil can be explained by the differences observed in the pre-test, whereas in Chile it could be due to the differences observed in the exposure to technology (see Table 5) as much as it is to the low pre-test scores when compared to all the other classes. We conclude that the time required to learn how to use both technologies was similar. Therefore, the necessary required abilities were similar to both technologies under study. However, this time period is also related to the arithmetic and technological ability of the children. It should be noted that of the 45 minutes available in the first game-based session, approximately ten of these were used to introduce the game-based narrative, which decreased the time available for practising arithmetic activities and, therefore, negatively affecting students’ progress.

Graph 2. Distribution of pedagogical progress levels during the first session

Qualitative results

A questionnaire composed of four questions was given at the beginning of the activity to evaluate the prevalence and use of technology between the different countries and schools, Table 5 first four questions. This test was validated previously by applying it to 85 students at two Chilean schools (one urban and one rural), with the goal of defining the necessary set of questions and possible answers needed to classify the students.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Chile Rural Game-based</th>
<th>Chile Rural Non-game-based</th>
<th>Chile Urban Game-based</th>
<th>Chile Urban Non-game-based</th>
<th>Brazil Urban Game-based</th>
<th>Brazil Urban Non-game-based</th>
<th>Brazil Rural Game-based</th>
<th>Brazil Rural Non-game-based</th>
<th>Costa Rica Urban Game-based</th>
<th>Costa Rica Urban Non-game-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was the first time you played a video game before you were six years of age?</td>
<td>45%</td>
<td>48%</td>
<td>55%</td>
<td>65%</td>
<td>52%</td>
<td>59%</td>
<td>41%</td>
<td>48%</td>
<td>44%</td>
<td></td>
</tr>
<tr>
<td>Do you play video games at least six times a week?</td>
<td>66%</td>
<td>69%</td>
<td>72%</td>
<td>75%</td>
<td>65%</td>
<td>73%</td>
<td>44%</td>
<td>47%</td>
<td>53%</td>
<td></td>
</tr>
<tr>
<td>Do you have a computer at home?</td>
<td>44%</td>
<td>35%</td>
<td>67%</td>
<td>75%</td>
<td>73%</td>
<td>71%</td>
<td>33%</td>
<td>57%</td>
<td>56%</td>
<td></td>
</tr>
<tr>
<td>Do you have an Internet</td>
<td>32%</td>
<td>36%</td>
<td>60%</td>
<td>66%</td>
<td>61%</td>
<td>53%</td>
<td>37%</td>
<td>23%</td>
<td>25%</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Exposure to technology and interest in the activity
connection at home?
Did time pass quickly while I was playing?
Did I forget I was in class while playing?
Do I want to keep playing?
Did I have a good time while I was playing?
Is it fun to play?

It can be observed that at least 41% of the students had their first experience with video games before they were six years of age, and that, except for Costa Rica, at least 65% of the students use this technology six times or more per week. Costa Rica presented the lowest values in all the indicators, and the rural locations presented a lower prevalence of computers.

A second questionnaire applied at the end of the experience studied student interest in the activity (Table 5, last four questions). We can see that the classes with the game-based activities had a higher level of interest when compared to non-game-based classes, for all the locations. There exists a –0.87 correlation between the exposure to the technology questionnaire results and the results from the student interest questionnaire. This shows that the less exposure a group had to technology, the more interested they were in the activity. Additionally, the observers noted an important difference in the reactions of students with greater and lesser exposure to technology. While the former reacted normally, the latter showed greater astonishment. When considered alongside the previous results, this suggests that the more accustomed students are with technology, the more demanding they will be.

In order to measure social interactions, the following student behaviors during each experience were quantified and registered: disruptive behavior; proactive responses to received feedback; lack of understanding of the activity; cooperation/assistance among peers, and satisfaction at having participated in the activity. The first four behaviors were registered the moment they occurred, while the last one (satisfaction) was recorded at the end of each session. Table 6 shows the average number of observations registered per child, and the children’s most representative quotes for each of the defined student behaviors.

Taking into account how cultural differences can influence perspective, the observers in each of the three countries were trained to maintain standardized observation criteria by analyzing various initial test videos. These videos showed children in Chile and defined which aspects of behavior to observe and their corresponding rubrics. The person in charge of these trainings also acted as a supervisor during the observations, and was present in all the activities in each of the three countries.

Table 6. Observed social behaviors (average occurrence per child in the experiment).

<table>
<thead>
<tr>
<th>Observed behavior</th>
<th>Chile</th>
<th>Brazil</th>
<th>Costa Rica</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R GB</td>
<td>R NGB</td>
<td>U GB</td>
</tr>
<tr>
<td>Disruptive behavior: intentional interruption of the flow of their own activity or that of their peers.</td>
<td>3.9</td>
<td>6.4</td>
<td>3.1</td>
</tr>
<tr>
<td>Examples:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proactive response to receiving feedback: the student’s positive and spontaneous expression immediately after receiving the system’s feedback.</td>
<td>13.7</td>
<td>8.8</td>
<td>18.7</td>
</tr>
<tr>
<td>Game-based examples:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion regarding the activity: doubt</td>
<td>1.7</td>
<td>3.8</td>
<td>1.4</td>
</tr>
<tr>
<td>Examples:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

354
expressed in a loud voice when presented with a pedagogical task.

Peer cooperation/assistance: peer communication with the goal of helping or asking for help with a pedagogical task.

<table>
<thead>
<tr>
<th>Peer cooperation/assistance: peer communication with the goal of helping or asking for help with a pedagogical task.</th>
<th>Examples:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A student teaches another how to answer a question.</td>
</tr>
<tr>
<td></td>
<td>A student demonstrates a different way to answer a question.</td>
</tr>
<tr>
<td></td>
<td>“That’s not how you do it . . . Look, . . .”</td>
</tr>
<tr>
<td></td>
<td>A student gives an answer to another student.</td>
</tr>
<tr>
<td></td>
<td>“Put 54! 54!”</td>
</tr>
<tr>
<td>Game-based examples:</td>
<td>“Who is the red star?”</td>
</tr>
</tbody>
</table>

Examples:

- A student teaches another how to answer a question.
- A student demonstrates a different way to answer a question.
- “That's not how you do it... Look, . . .”
- A student gives an answer to another student.
- “Put 54! 54!”

Game-based examples:

- “Who is the red star?”

The teachers commented that the children tended to be very concentrated on their own personal space, presenting lower levels of disruption than in traditional classes. In addition, comparing Table 2 and 6, we observed that the groups with the game-based activity had the lowest average occurrence of disruptive behaviors in each of the countries—except in the Brazilian urban school, which had a higher ratio of hyperactive students. We also observed a greater number of proactive responses to the feedback received in the game-based classes. Using feedback as a narrative element enabled the students to become more involved in the pedagogic activities. In the game-based activities there were also fewer expressions of confusion because students were more immersed in the activity. It is worth noting that in Brazil there was a greater number of expressions of confusion, but considering that the median in the pedagogic rules completed by each student was greater, the children were exposed to more complex activities that required greater student-teacher interaction. Correspondingly, there is more frequent cooperation/assistance among peers in the game-based activities. The cooperation/assistance among peers has a 0.78 correlation with the questionnaire about exposure to technology. That is, the more familiar students are with technology, the greater their tendency to cooperate. In accordance with the above and with the correlation obtained between exposure to technology and interest in the activity (Table 5), a correlation of –0.87 resulted between exposure to technology and degree of satisfaction with the activity.

**Conclusions**

Our first research question asked: Within individual pedagogical activities in which arithmetic is practised using an interpersonal computer, what is the added value of the game in terms of student knowledge gained when compared with the same non-game-based activities and the traditional paper and pencil activities? In this paper we have proved:

1) that there was no significant difference between the final results of the game-based and non-game-based systems. This shows that the game’s narrative and graphics did not impact learning when compared with the technology used in the other technological activity. 2) The game-based activity did, however, distinguish itself by capturing greater student interest, which led to greater involvement in the activity. 3) The traditional method group, with no technology, was shown to produce consistently inferior acquired knowledge results when compared to the technological groups. This was principally due to the fact that the technological systems could provide immediate and individualized feedback for each student. 4) In terms of knowledge acquired, the technological systems proved to be more beneficial for the students who were considered weakest at the beginning of the experiment.
Our second research question asked: Is culture a factor that influences learning and/or student interest when interacting with game-based activities on the interpersonal computer? Regarding acquired student knowledge, we were able to observe a significant difference between countries. Such difference may be explained by the different technology appropriation and usage levels in the respective school contexts, factors that correspond to the degree of fluidity that the students have as they interact with the application. Only in Chile did we observe children quickly familiarizing themselves with the game's narrative and, without being prompted, expressing excitement when completing the different stages. In Costa Rica, the number of student interjections regarding “Proactive responses to received feedback” and “Satisfaction of participating in the activity” was greater, but the intensity of the commentaries received was greater in Chile. This led us to conclude that children with less exposure to the game and technology achieved greater immersion; however, involvement is greater when the narrative is closer to the socio-cultural context of the child. Through the observations, we also noted that children from the game-based groups in Costa Rica, after finishing each of their sessions, asked each other about how far they had got in the game and ranked one another. This shows that in Costa Rica, the two game-based groups (rural and urban) transcended the narrative and naturally created an environment of healthy competition. We leave as an open question how a greater level of competitiveness among children could improve the results in the various cultural contexts. It should be noted that it was not possible to analyze the game-based component’s added value by country due to the small sample size in each country.

Our results cannot be generalized; the samples were neither random nor representative of their countries. We leave as future work to develop a year-long randomized representative study to observe if this greater involvement translates into improved learning.

In the non-game-based version, the children had to construct the answers digit by digit, while the game-based version used multiple-choice. This led to a significant increase in students using trial and error. We leave as future work the development of game-based activities where the answers are not multiple-choice, to observe whether or not immersion is maintained and learning improved. An objective that was not fulfilled was the disconnection of students from their surroundings, whereby players become less conscious of their reality and of themselves (Fu, Su, & Yu, 2009). The interpersonal computer limits the game’s mechanics (due to the low-pixel count that corresponds to each child and the lack of sound) and, therefore, more studies are needed to determine ways of creating more engaging activities for this platform.

The level of involvement and commitment of the children in the game-based groups after taking the post-test (on paper) was lower than the non-game-based group. One possible cause is the change of work environment, from video game to paper work. We leave as future work discerning how to evaluate the video game group with video games to observe if this improves their post-test. However, at the same time it is necessary to observe if these children are capable of transferring what they have learned to other mediums, such as paper.

Acknowledgments

This study was partially funded by LACCIR, Microsoft Research, CONICYT- FONDECYT 1120177.

Reference


Languages for Specific Purposes in the Digital Era
(Book review)

Reviewers:
Nuria Otero de Juan
Researcher
Universidad de Alcalá, Madrid, Spain
n.otero@uah.es

Jesús García Laborda
Associate Professor
Universidad de Alcalá, Madrid, Spain
Jesus.garcialaborda@uah.es

Textbook Details:
Language for Specific Purposes in the Digital Era
Edited by Elena Bárcena, Timothy Read and Jorge Arús, 348 pages,
Published by Springer,

As the book states at the beginning, “research on CALL applications aimed at the teaching and learning of LSPs is a thriving area within the field of linguistics” (p. 5), and we agree that due to the enormous development of technological tools and applications, an analyse of the current situation and knowledge gained up to now from a better perspective is required. It would also allow us to reflect on the best way to make technological teaching and research innovation. In a way, this is what the book aims to do.

Despite of the wide range of different information covered in the book, it is clearly divided into four different parts, each of which comprising four chapters. The first part introduces some general issues about language learning and computing. The second part explains how language skills and therefore, linguistic competences, can be enhanced through computer-assisted experiences. The third part focuses on corpus-based applications for teaching and processing specific language domains. And finally, the fourth part deals with natural language processing.

Its sixteen chapters are -in general- easy to read and clear, making it an understandable book for people without a deep knowledge of programming or computer science. In order to get a more in-depth understanding of the contents of the book, a more detailed overview is given for each chapter, analysing the outstanding premises from a critical perspective:

Chapter 1 is an overview of the relationship between Information Technology and Languages for Specific Purposes identifying some areas of interest (Arnó et al., 2006). This part is extremely interesting, since it sums up the changes brought about by the Bologna Process resulting in the new European Higher Education Area, presenting in a nutshell the options and challenges for the knowledge society and showing the Quantum LEAP Project as an example of how to take advantage of these opportunities. Chapter 2 introduces the InGenio online authoring shell, as a tool that overcomes the difficulties found by the teachers when creating online language learning materials. According to Blin’s statement (2005, p. 33), “language learning environments [...] does not require the constant intervention of a teacher”, and being aware of the important role of independence in online learners, the authors provide some examples of initiatives that promote autonomy in the online learning process. Chapter 3 defends online blended learning -combining online and face to face work (Sharma, 2010) -, as exemplified by the efficacy of the I-AGENT Project, especially in the development of the foreign language learners’ oral skills. A classification of the different Virtual Learning Environments (VLEs) is presented establishing the difference between Content Management Systems (CMS) and Learning Management Systems (LMS), such as Moodle, Blackboard and WebCT. Chapter 4 states how assessment is a specific phase within the whole language experience (p. 69), and must be taken into account in online language learning as well, as InGenio System does with two different modalities (tutor/self-assessment).

Part 2 begins with the attractive chapter 5, which explains some of the requirements - the advantages were explained by Gasparetti et al. (2009, p. 287) - for specialized dictionaries as a pedagogical tool, and wonders if web-based
systems are appropriate for educational environments considering some of their features. It concludes, interestingly, that “Internet pedagogical dictionaries are much more than electronic versions of printed dictionaries” (p. 103). Chapter 6 details two different university activities carried out in the context of Teaching Legal English using Moodle glossaries. This chapter outlines the importance of motivation in collaborative tasks among students which enhances the learning process, and such motivation may be reached through online writing (Breeze, 2005; Davoli et al., 2009). In chapter 7, the author argues that in contrast to former times, vocabulary learning is receiving increasing attention in the foreign language learning process and “address the challenges of ICT-based pedagogy to the English for Specific Purposes classroom in general, and the ME [Maritime English] classroom in particular” (p. 139). Chapter 8 focuses on the use of Wikis as a collaborative and educational tool for Business English learning.

Part III starts with chapter 9, which states that a corpus of specialized full-text discourse documents can be a useful tool in the LSP class (Tribble, 1997), as the authors show by presenting the genre-based GENTT experience. Chapter 10 presents the use of corpora in a Russian Business class, and acknowledges that using corpora allows tuition to be offered in several LSP domains and curricula to be tailored the industry requirements, thereby increasing future opportunities of graduates (p. 198). However, as the authors make clear, tutors need to create materials and develop corpus-based exercises. Chapter 11, begins with the following question: “Is it possible to find specific characteristics in specialized text different from their discourse conditions or terminology they have?” and goes on to provide a consistent answer (p. 224). The last chapter in part III introduces the idea of teaching as being flexible in the sense that should allow students to cope with unexpected future situations since the amount of technical knowledge and tools around grows exponentially (Kastberg, 2002), and in combination with this idea, it also explores how efficient the use of corpora in the design of Danish translation courses is.

The first chapter in part IV, explains very clearly how knowledge can be represented in a user-friendly tool such as EcoLexicon. Chapter 14 shows to what extend can corpus-based studies are useful for the teaching of dubbing and subtitling, an activity which is likely to continue to gain research prominence in the future, since audiovisual translation is already a subject at Spanish universities. Chapter 15 introduces OntoLingAnnot, an useful annotation tool for linguists with a “comprehensive, global, detailed and general” approach (p. 340) since it covers the pragmatic level in a deep way and considers six different ontologies. The final chapter, written by the editors, plays its role as a conclusion. One of the weaknesses of such a compilation is stated here, and we agree with the idea that the real state-of-the-art in a field such as applied technologies can never be reached due to their steep evolution.

The book has a pedagogical approach, from an experiential teaching and learning perspective, always trying to adopt a communicative approach to language learning. One of the strengths of it is the list of recognized contributing experts, who have carried out both applied and theoretical research in this field. It is not only oriented to academic readers, or even to readers interested in technology applied to Languages for Specific Purposes, but to anyone with an interest in the way in which technology and the process of language teaching and learning have become almost inseparable. Furthermore, it will be found to be essential for all of those working in the defined area of specific linguistics domains.

References


Breeze, R. (2005). The sight and sound project: Developing an online film magazine as part of a university-level ESP course. In S. Posterguillo (Ed.), Language @t work: Language learning, discourse and translation studies in internet (pp. 57-70). Castellón, Spain: Universitat Jaume I.


