Educational Technology & Society
An International Journal

Aims and Scope

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

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

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 The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to Educational Technology & Society and three months thereafter.

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

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

Editors

Kinshuk, Athabasca University, Canada; Demetrios G Sampson, University of Piraeus & ITI-CERTH, Greece; Ashok Patel, CAL Research & Software Engineering Centre, UK; Reinhard Oppermann, Fraunhofer Institut Angewandte Informationstechnik, Germany.

Editorial Assistant

Barbara Adamski, Athabasca University, Canada.

Associate editors

Nian-Shing Chen, National Sun Yat-sen University, Taiwan; Alexandra I. Cristea, Technical University Eindhoven, The Netherlands; John Eklund, Access Australia Co-operative Multimedia Centre, Australia; Vladimir A Fomichev, K. E. Tsiskovskiy Russian State Tech Univ, Russia; Olga S Fomicheva, Studio "Culture, Ecology, and Foreign Languages", Russia; Piet Koomers, University of Twente, The Netherlands; Chul-Hwan Lee, Inchon National University of Education, Korea; Brent Muirhead, University of Phoenix Online, USA; Erkki Sutinen, University of Joensuu, Finland; Vladimir Uskov, Bradley University, USA.

Advisory board

Ignacio Aedo, Universidad Carlos III de Madrid, Spain; Luis Anido-Rifon, University of Vigo, Spain; Rosa Maria Bottino, Consiglio Nazionale delle Ricerche, Italy; Mark Bullen, University of British Columbia, Canada; Tak-Wai Chan, National Central University, Taiwan; Sherry Chen, Brunel University, United Kingdom; Darina Dicheva, Winston-Salem State University, USA; Robert Farrell, IBM Research, USA; Brian Garner, Deakin University, Australia; Roger Hartley, Leeds University, UK; J R Isaac, National Institute of Information Technology, India; Mohamed Jenni, University of Tunis, Tunisia; 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; Rudy Lelouche, Universite Laval, Canada; David McConnell, Lancaster University, UK; Roy McCreath, 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; Hiroaki Ogata, Tokushima University, Japan; Toshio Okamoto, The University of Electro-Communications, Japan; Thomas C. Reeves, The University of Georgia, USA; Gilly Salmon, University of Leicester, United Kingdom; 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; Brian K. Smith, Pennsylvania State University, USA; J. Michael Spector, Florida State University, USA; Chin-Chung Tsai, National Taiwan University of Science and Technology, Taiwan; Stephen J.H. Yang, National Central University, Taiwan.

Assistant Editors

Sheng-Wen Hsieh, Far East University, Taiwan; Taiyu Lin, Massey University, New Zealand; Kathleen Luchini, University of Michigan, USA; Dorota Mularczyk, Independent Researcher & Web Designer; Carmen Padrón Nápoles, Universidad Carlos III de Madrid, Spain; Ali Fawaz Shareef, Massey University, New Zealand; Jarkko Suonen, University of Joensuu, Finland.

Executive peer-reviewers

http://www.ifets.info/

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 a 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, they can also be sent via email to (Subject: Submission for Educational Technology & Society journal): kinshuk@ieee.org. In the email, please state clearly that the manuscript is original material that has not been published, and is not being considered for publication elsewhere.
Table of contents

Full length articles

Effect of Computer-Based Video Games on Children: An Experimental Study
Tsung-Yen Chuang and Wei-Fan Chen 1-10

Using Blogs to Enhance Critical Reflection and Community of Practice
Shih-Hsien Yang 11-21

Automated Inattention and Fatigue Detection System in Distance Education for Elementary School Students
Kuo-An Hwang and Chia-Hao Yang 22-35

The Instructional Effect of Online Reading Strategies and Learning Styles on Student Academic Achievement
Pei-Hsuan Hsieh and Francis Dwyer 36-50

MEAT: An Authoring Tool for Generating Adaptable Learning Resources
Yen-Hung Kuo and Yueh-Min Huang 51-68

Qualitative assessment across language barriers: An action research study
Johannes C. Cronjé 69-85

An Exploration of Pre-Service Teachers’ Perceptions of Learning to Teach while Using Asynchronous Discussion Board
Lasisi Ajayi 86-100

A Capstone Experience for Preservice Teachers: Building a Web-Based Portfolio
Leah M. Herner-Patnode and Hea-Jin Lee 101-110

Evaluating an Intelligent Tutoring System for Design Patterns: the DEPTHS Experience
Zoran Jeremic, Jelena Jovanovic and Dragan Gasevic 111-130

Digital Students in a Book-Oriented School: Students’ Perceptions of School and the Usability of Digital Technology in Schools
Yifat Ben-David Kolikant 131-143

A Markov-based Recommendation Model for Exploring the Transfer of Learning on the Web
Yueh-Min Huang, Tien-Chi Huang, Kun-Te Wang and Wu-Yuin Hwang 144-162

An Educational Mobile Blogging System for Supporting Collaborative Learning
Yueh-Min Huang, Yu-Lin Jeng and Tien-Chi Huang 163-175

Affective e-Learning: Using “Emotional” Data to Improve Learning in Pervasive Learning Environment
Liping Shen, Minjuan Wang and Ruimin Shen 176-189

Wireless Handhelds to Support Clinical Nursing Practicum
Cheng-Chih Wu and Chin-Yuan Lai 190-204

A Coursework Support System for Offering Challenges and Assistance by Analyzing Students’ Web Portfolios
Liang-Yi Li and Gwo-Dong Chen 205-221

Intra-action, Interaction and Outeraction in blended learning environments
Wu-Yuin Hwang, Jung-Lung Hsu, Alexei Tretiakov, Huey-Wen Chou and Ching-Yuan Lee 222-239

CogSkillnet: An Ontology-Based Representation of Cognitive Skills
Petek Askar and Arif Altun 240-253
<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engaging students in multimedia-mediated Constructivist learning – Students’ perceptions</td>
<td>254-266</td>
</tr>
<tr>
<td><strong>Mai Neo and Tse-Kian Neo</strong></td>
<td></td>
</tr>
<tr>
<td>Comparing Learning Performance of Students Using Algorithm Visualizations Collaboratively on Different Engagement Levels</td>
<td>267-282</td>
</tr>
<tr>
<td><strong>Mikko-Jussi Laakso, Niko Myller and Ari Korhonen</strong></td>
<td></td>
</tr>
<tr>
<td>Developmental Progression of Referential Resolution in Comprehending Online Texts</td>
<td>283-287</td>
</tr>
<tr>
<td><strong>Yu-Fen Yang and Ya-Chen Hung</strong></td>
<td></td>
</tr>
<tr>
<td>Attitudes and Satisfaction with a Hybrid Model of Counseling Supervision</td>
<td>298-306</td>
</tr>
<tr>
<td><strong>Steven R. Conn, Richard L. Roberts and Barbara M. Powell</strong></td>
<td></td>
</tr>
<tr>
<td>Modeling Academic Education Processes by Dynamic Storyboarding</td>
<td>307-333</td>
</tr>
<tr>
<td><strong>Yoshitaka Sakurai, Shinichi Dohi, Setsuo Tsuruta and Rainer Knauf</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Book review(s)**

<table>
<thead>
<tr>
<th>Book Review</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology-mediated Learning Environments for Young English Learners (L. Leanne Parker)</td>
<td>334-336</td>
</tr>
<tr>
<td><strong>Reviewer:</strong> Griff Richards</td>
<td></td>
</tr>
<tr>
<td>Tips for Teaching with CALL: Practical Approaches to Computer-assisted Language Learning (Carol Chapelle and Joan Jamieson)</td>
<td>337-338</td>
</tr>
<tr>
<td><strong>Reviewer:</strong> Ferit Kilickaya</td>
<td></td>
</tr>
</tbody>
</table>
Effect of Computer-Based Video Games on Children: An Experimental Study

Tsung-Yen Chuang¹ and Wei-Fan Chen²

¹Information and Learning Technology, National University of Tainan, 33, Sec. 2, Shu-Lin St. Tainan, Taiwan 700 // chuangyen@mail.nutn.edu.tw // Tel: +886-6-2133111 Ext: 940
²Information Sciences and Technology, The Pennsylvania State University, 03 Bell Atlantic Center for Technology, Lehman, PA 18627, USA // weifan@psu.edu // Tel: +1-570-675-9142

ABSTRACT
This experimental study investigated whether computer-based video games facilitate children’s cognitive learning. In comparison to traditional computer-assisted instruction (CAI), this study explored the impact of the varied types of instructional delivery strategies on children’s learning achievement. One major research null hypothesis was tested: there are no statistically significant differences in students’ achievement when they receive two different instructional treatments: (1) traditional CAI; and (2) a computer-based video game. One hundred and eight third-graders from a middle/high socio-economic standard school district in Taiwan participated in the study. Results indicate that computer-based video game playing not only improves participants’ fact/recall processes ($F = 5.288, p < .05$), but also promotes problem-solving skills by recognizing multiple solutions for problems ($F = 5.656, p < .05$).

Keywords
Game learning, Play theory, Cognitive processes, Educational technology

Introduction
The use of multimedia in education has significantly changed people’s learning processes. Results from a number of research studies indicate that appropriately designed multimedia instruction enhances students’ learning performance in science, mathematics, and literacy (Gee, 2003). Previous studies indicate that computer-assisted instruction (CAI) programs have important factors that can motivate, challenge, increase curiosity and control, and promote fantasy in children (Tzeng, 1999). Despite the fact that computer and video games have the same multimedia capability as CAI programs, their potential learning impact is often discounted by parents and educators. Recently, computer-based video games’ presence and popularity have been ever-growing, and game developers and researchers have started to investigate video games’ impact on students’ cognitive learning (Begg, Dewhurst, & Macleod, 2005; Squire, 2003; Vaupel, 2002). For example, Pillay commenced a study investigating the influence of recreational computer games on children’s subsequent performance on instructional tasks (Pillay, 2002). While game-playing is regarded somewhat negative in educational settings, particularly for young children, re-scrutinization of its influence in a teaching and learning context is crucial.

This study investigated whether computer-based video games facilitate children’s cognitive learning achievement. In comparison to traditional CAI programs, this study explored the impact of the varied types of instructional delivery strategies on children’s learning achievement.

Research results from the Kaiser Family Foundation and the Children’s Digital Media Centers (Rideout, Vandewater, & Wartella, 2003) found that children in the United States are growing up with media and are spending hours a day watching television and videos, using computers, and playing video games. According to the findings, today’s children are starting to be exposed to technologies and media at a much younger age than previously thought. Therefore, educators’ investigations become critical concerning the impact of technologies and media on children’s development. This study investigates two main questions: (1) Can computer-based video games be instructional tools in early childhood education? (2) Should instructional strategies be modified to fit into young children’s media experience?

Most previous research studies related to computer-based video games focused on the discussions of psychological study and child behavior (Provenzo, 1991; Squire, 2003). In psychological study, research results indicated that video games can promote hand-eye coordination, visual scanning, auditory discrimination, and spatial skills (Johnson, Christie, & Yawkey, 1999; Lisi & Wolford, 2002). For child behavior, evidence showed that violent video games may raise children’s aggressive play and violent behaviors (Funk, 2001). Separate from previous research,
this study discusses computer-based video games from an educational perspective by exploring the following issues: (1) How might some of the motivating aspects of computer-based video games be harnessed to facilitate learning? (2) How might motivational components of popular computer-based video games be integrated into instructional design?

**Research null hypothesis**

Based upon the aforementioned purpose of study, one major research null hypothesis may be drawn. That is, there are no statistically significant differences in students’ achievement when they receive two different instructional treatments: (1) traditional CAI; and (2) a computer-based video game.

**Methods**

**Participants**

The chosen participants were from a middle/high socio-economic standard school district in Tainan City, Taiwan. One hundred and eight third-graders participated in the study, 61 male students and 54 female students. The learning prerequisite for participants was the ability to use basic computer tools and Internet-browsing resources.

**Instructional materials**

The instructional topic used for this study was fire-fighting. Instructional content covered the basic knowledge about the danger of fire and fire-fighting techniques. The content was chosen because it covers a variety of educational objectives (facts, concepts, analyses) and is not related to the participants’ courses, thus avoiding potential threats to internal validity of the study.

Two different instructional materials: computer-assisted instruction and a computer-based video game were used in this study. For the computer-assisted instruction, a unit of text-based instructional materials was developed. For the computer-based video game, Fire Department 2: Fire Captain, a 3D, real-time strategy game developed by Monte Cristo Games, was chosen for this study. This computer game has been assessed by the Pan-European Game Information (PEGI) age rating system, and has been approved for children seven years and older.

This 3D, real-time computer game applied the design principles derived from Keller’s (1987) attention, relevance, confidence, and satisfaction (ARCS) model of motivation to facilitate children’s learning. Below is a description of how the game’s design strategies related to the ARCS model of motivation.

- **Attention:** The game created a 3D multimedia environment that generated multiple sensory stimuli to keep learners’ attention focused. It also created an active learning environment by using simulation and role-playing strategies.
- **Relevance:** The game addressed the importance of fire-fighting and related safety issues before learners started to play the game. Also, a clear description of learning objectives was presented to connect the current topic with their prior knowledge and skills.
- **Confidence:** The game provided a variety of options according to learners’ competence levels. For example, there is a tutorial game designed particularly for beginners not familiar with game playing.
- **Satisfaction:** After learners completed the game, several important statistics in the mission were reported, such as the total time spent on the mission, the number of people injured, the number of cars damaged, etc. This design strategy elicited learners’ actual performance in the game and motivated them to play it again.

**Independent variables**

One independent variable was examined in this study: type of instructional treatments. Two groups (the computer-assisted instruction and computer-based video game playing) were identified for the variable of instructional treatment.
Dependent variables

The dependent variables used in the study were the outcome of participants’ post-test scores and the results of the three sub-categories in the post-test. A post-test containing 28 items was designed for the study. It has a total of 100 points. The following details the three sub-categories in the post-test:

(1) Part 1: Multiple-choice (16 items): Part 1 measures knowledge of specific facts, terms, and definitions by asking participants to answer multiple-choice questions that relate to the topic of fire-fighting. The objectives measured in this test require an understanding of factual materials and concepts.

(2) Part 2: Matching (6 items): Part 2 evaluates participants’ ability to associate specific terms from a list of appropriate context. The objectives measured in this test also focus on recalling facts and identifying factual information. The difference between part 1 and part 2 is that for matching up specific terms, participants need to analyze and compare the similarities and differences in the descriptions.

(3) Part 3: Application (6 items): In part 3, participants are required to demonstrate their understanding of how to put out a fire as a firefighter in an emergency scenario. This part requires that participants thoroughly understand the danger of fire, fire safety information, and fire-fighting processes during a rescue scenario. The test measures a higher-level cognitive task that shows understanding of what is being taught and its use in other circumstances.

Instructional treatments

The experimental study required 115 third-grade students to compare two different types of instructional treatment by looking at their learning performance. By using similar content, this study investigated which instructional treatment better affects students’ cognitive learning (the computer-assisted instruction versus computer-based video game playing). Different instructional treatments were designed and selected for this study.

Control group

Participants in the control group (CG) learned by interacting with computers individually. The contents of the computer-assisted instruction in this study used a text-based format. The fire-fighting content was presented in a single web page without hyperlinks. The content included elements of fire, fire categories, fire-fighting techniques, and fire safety information. Participants determined their own reading pace by dragging the scroll bar on the right side of the web browser. No teacher was involved in this study, and the researcher provided only technical support to the participants when needed.

Experimental group

The purpose of this study was to determine whether computer-based video games support students’ cognitive processes. With this hypothesis, a game called Fire Captain was chosen for the experimental group (EG). This computer game used similar instructional content as the control group. The primary goal of this game is to master all kinds of knowledge and skills to prevent fire. Although five single-player campaigns with increasingly complex missions constitute this game, the tutorial mission is one that participants have to play first in the study. In the tutorial mission, the participants learned everything they need to know to be a firefighter through each relevant and inclusive introductory task. Because Fire Captain is a real-time strategy game, the players needed to put out a fire at the end of tutorial mission, and their performance was graded by the game. To avoid interference with others, no teacher was involved in the experiment. As in the control group, the researcher provided only technical support to the participants when needed.

Experimental procedures

After obtaining approval from the Office for Regulatory Compliance from the researchers’ universities, the researchers contacted the principal of the selected school in order to obtain a formal agreement for conducting this study. Two weeks prior to the formal study, parental consent forms were distributed to 136 students. These parental consent forms needed to be signed by the recruited students’ parents.
Sampling for preliminary study

One hundred and fifteen students agreed to participate in this study and provided parental consent. They were randomly assigned into two groups, the control group \( n = 58 \) and the experimental group \( n = 57 \).

Procedures

Participants were asked to take part in two instructional sessions. Estimates were that each session would take 40 minutes (the length of a regular class period in elementary schools in Taiwan), for a total of 80 minutes. In the first session, the control group received a unit of computer-assisted instruction about fire-fighting. The experimental group was asked to play the computer game Fire Captain. In the next session, all participants were required to take a quiz to allow assessment of their learning achievements.

As the experiment began, the researcher asked the participants to turn on their monitors. For the participants in the control group, their monitors immediately showed a text-only web page. The researcher informed the participants that they would have 35 minutes to read through the instructions. The participants could browse the contents at their own pace. As soon as they finished the text-based instruction, they raised their hands to indicate to the researcher completion of the task. However, anyone who finished early had to remain seated and wait for the other children to finish until the time was up.

For the experimental group at the beginning of the class, the researcher used five minutes to demonstrate the skills of mouse control for the computer game. After the demonstration, the experimental group turned on their monitors, put on headphones, and began playing the computer game. The main menu of Fire Captain appeared on their monitors. Participants also had 35 minutes to explore this game freely. However, they had to finish the tutorial mission as soon as possible. How fast they completed the tutorial mission was a significant factor for evaluating their ranking at the end of the mission. Participants had permission to raise their hands at any time to ask for any technical support during this session.

After receiving various instructional treatments, each participant was required to take a quiz during the second session. The quiz was designed and developed to assess participants’ learning achievements after finishing their instruction. From three categories of questions — fact differentiation/recall, analysis/comparison, and understanding/problem-solving — different types of learning outcomes were determined. The researcher graded and recorded all the post-tests.

Research design

This study investigated the effect of varied types of instructional treatments (computer-assisted instruction versus computer and video game playing) on individual learning achievement for third-grade students. This is a typical one-factor-with-two-levels experiment. Since the dependent variables are typically related statistically and conceptually (Yoder, 2002), the statistical correlation should be in the range of a low to moderate level (Tabachnick & Fidell, 2001). A multivariate analysis of variance (MANOVA) was performed to analyze the results. Furthermore, the real value of using MANOVA is in controlling Type I error (also known as false-positive error) while simultaneously analyzing multiple dependent variables.

Data collection

The data for this research was collected from a random sample of third-grade students enrolled in the selected school in Tainan City, Taiwan. One hundred and thirty-six third graders were expected to participate in this study. After contacting the participants’ parents, 115 participants were able to participate. After instructional treatments and conclusion of the post-test, a total of 108 participants had completed the experiment. The drop-out rate for the study was 6.1 percent.
The data used for statistical analysis was collected from two sources: (1) the results of the post-test scores, and (2) the results of the three sub-categories in the post-test. The reliability analysis of the post-test showed internal consistency of the test scores. By using Cronbach’s coefficient alpha, we reported an overall reliability of .88 for this study. Specific reliability for 16 items in Part 1: multiple-choice was .86, and for the six items in Part 2: matching, it was .71

**Statistical data analysis**

To test significant differences in learning achievement among the treatments (computer-assisted instruction versus computer and video game playing), a one-way multivariate analysis of variance (MANOVA) was performed. Multiple analysis of variance in Statistical Package for the Social Sciences (SPSS) was used to examine main and potential interaction effects of categorical variables on multiple interval/ratio dependent variables. That is, MANOVA tests for population group differences on several dependent variables simultaneously by creating vectors of the multiple dependent variables (Garson, 2005; Yoder, 2002). Thus, in conducting a one-way MANOVA for this study, we used various types of instructional treatments as independent variables (i.e., nominal scales); the outcome of students’ multiple-measure scores of the post-test were the dependent variables (i.e., interval/ratio scales). An alpha level of .05 was set to analyze the significant difference of hypotheses. Prior to using MANOVA, the researcher examined the test results and found the distribution to be fairly normal with no extreme outliers.

From the results of the post-test (see Table 1), group means showed significant differences between the web CAI group ($M = 70.05$) and the video game group ($M = 76.18$).

<table>
<thead>
<tr>
<th>Treatment groups</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part 1 (0–64 points)</td>
<td>Web CAI</td>
<td>48.8421</td>
<td>8.27170</td>
</tr>
<tr>
<td></td>
<td>Video game</td>
<td>52.2353</td>
<td>6.90098</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>50.4444</td>
<td>7.80766</td>
</tr>
<tr>
<td>Part 2 (0–18 points)</td>
<td>Web CAI</td>
<td>10.7368</td>
<td>4.70762</td>
</tr>
<tr>
<td></td>
<td>Video game</td>
<td>11.7059</td>
<td>4.08066</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>11.1944</td>
<td>4.42900</td>
</tr>
<tr>
<td>Part 3 (0–18 points)</td>
<td>Web CAI</td>
<td>10.4737</td>
<td>3.85035</td>
</tr>
<tr>
<td></td>
<td>Video game</td>
<td>12.2353</td>
<td>3.83452</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>11.3056</td>
<td>3.92561</td>
</tr>
<tr>
<td>Post-test total score (0–100 points)</td>
<td>Web CAI</td>
<td>70.0526</td>
<td>12.96068</td>
</tr>
<tr>
<td></td>
<td>Video game</td>
<td>76.1765</td>
<td>10.76978</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>72.9444</td>
<td>12.31049</td>
</tr>
</tbody>
</table>

**Correlations between multiple dependent variables**

MANOVA assumes that multiple dependent variables relate statistically and conceptually (Yoder, 2002). According to this assumption, correlations between multiple dependent variables should be checked before MANOVA is conducted. As Table 2 indicates, most of the dependent variables have statistically significant correlations with one another ($p < .01$) in the range considered appropriate by Tabachnick and Fidell (2001).

<table>
<thead>
<tr>
<th></th>
<th>Part 1</th>
<th>Part 2</th>
<th>Part 3</th>
<th>Post-test total score</th>
<th>Prior knowledge score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson correlation</td>
<td>1</td>
<td>.349**</td>
<td>.341**</td>
<td>.869**</td>
<td>.534**</td>
</tr>
<tr>
<td>Sig. (two-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>108</td>
<td>108</td>
<td>108</td>
<td>108</td>
<td>108</td>
</tr>
</tbody>
</table>
By examining the Levene’s test of equality of error variances, the results showed that most of the dependent variables had equal variances ($p > .05$). Based on the above testing results, a one-way MANOVA could be conducted for data analysis.

**Results of multivariate analysis of variance (MANOVA)**

Table 3 reports MANOVA for the effect of the studied independent variable on the dependent groups.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Wilks’ lambda</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.022</td>
<td>1568.560</td>
<td>.000*</td>
</tr>
<tr>
<td>Treatment groups</td>
<td>.927</td>
<td>2.739</td>
<td>.047*</td>
</tr>
</tbody>
</table>

* Significant at .05 level

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent variable</th>
<th>Type III sum of squares</th>
<th>$df$</th>
<th>Mean square</th>
<th>$F$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment groups</td>
<td>Part 1</td>
<td>309.911</td>
<td>1</td>
<td>309.911</td>
<td>5.288</td>
<td>.023*</td>
</tr>
<tr>
<td></td>
<td>Part 2</td>
<td>25.276</td>
<td>1</td>
<td>25.276</td>
<td>1.292</td>
<td>.258</td>
</tr>
<tr>
<td></td>
<td>Part 3</td>
<td>83.530</td>
<td>1</td>
<td>83.530</td>
<td>5.656</td>
<td>.019*</td>
</tr>
<tr>
<td></td>
<td>Post-test total score</td>
<td>1009.413</td>
<td>1</td>
<td>1009.413</td>
<td>7.036</td>
<td>.009*</td>
</tr>
<tr>
<td>Error</td>
<td>Part 1</td>
<td>6212.755</td>
<td>106</td>
<td>58.611</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part 2</td>
<td>2073.641</td>
<td>106</td>
<td>19.563</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part 3</td>
<td>1565.387</td>
<td>106</td>
<td>14.768</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-test total score</td>
<td>15206.254</td>
<td>106</td>
<td>143.455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Part 1</td>
<td>281344.000</td>
<td>108</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part 2</td>
<td>15633.000</td>
<td>108</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part 3</td>
<td>15453.000</td>
<td>108</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Post-test total score</td>
<td>590872.000</td>
<td>108</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Univariate analyses, on Table 4, reveal that the dependent variables had significant main effects on the independent variables in all the tests except for the Part 2 test on treatment groups.

Findings and discussions
To accurately discuss the significant differences between various instructional treatments, the results of this study were subdivided into four sections: multiple-choice, matching, application, and the post-test total scores.

Finding 1: Multiple-choice - Univariate analysis results for Part 1 show that significant differences between treatment groups exist ($F = 5.288, p < .05$). Therefore, for Part 1, the null hypothesis should be rejected. Compared to the control group, the experimental group on average had a significantly higher score.

Finding 2: Matching - Univariate analysis results for Part 2 show no significant differences between treatment groups ($F = 1.292; p > .05$). Therefore, for Part 2, the null hypothesis is retained. However, according to Table 1, the experimental group (mean = 11.71; standard deviation = 4.08) on average still had a slightly higher score than the control group (mean = 10.74; standard deviation = 4.71).

Finding 3: Application - Univariate analysis results for Part 3 show that significant differences between treatment groups exist ($F = 5.656; p < .05$). Therefore, for Part 3, the null hypothesis should be rejected. Compared to the control group, the experimental group on average had a significantly higher score.

Finding 4: Post-test total score - Univariate analysis results for the post-test score show that significant differences between treatment groups exist ($F = 7.036; p < .05$). Therefore, for post-test, the null hypothesis should be rejected. Analysis of the results shows that the experimental group, on average, had a significantly higher post-test score when compared with the control group.

Summary of findings
The statistical results clearly show a significant difference between computer-assisted instruction and computer-based video game playing in students’ learning achievement (see Figure 1). Based on this finding, playing computer-based video games was determined to be more effective in facilitating third-graders’ average learning outcome than text-based computer-assisted instruction. Therefore, it may be concluded that this null hypothesis should be rejected.

According to the results of post-test scores, computer-based video games can clearly facilitate students’ learning performance. This finding indicated that computer-based video game playing not only can improve participants’ fact differentiation/recall processes, but also promotes problem-solving skills by recognizing multiple solutions for problems. Additionally, more precise effects were detected by breaking the post-test scores into three sub-categories. When the researchers developed the items in Part 1, the objectives were to make the participants recall key facts, definitions, and concepts. As the results of Part 1 indicate, the participants in the experimental group have a better understanding of factual materials and concepts. Consequently, for finishing their mission, the participants in the experimental group know they have to memorize the key facts or concepts even if they were just playing a game.

The design of Part 2 is to measure the participants’ ability to associate and analyze similarities and differences. Unfortunately, the results in this part are not as impressive as expected. According to the results, there is no statistically significant difference in Part 2, and the scores of the experimental group are only slightly better than the scores of the control group. Consequently, computer-based video game playing did not much improve participants’ judgment. Two explanations are possible for this result: (1) insufficient descriptions or explanations of each key concept in the computer game, thus the participants in the experimental group did not have enough information to analyze and compare the differences, (2) the depictions of each item in Part 2 was not simple and easy enough for third graders to understand. Therefore, they were confused in making the right choice.
The results in Part 3 are the most remarkable finding in this study. Apparently computer-based video game playing is effective for improving critical thinking and higher-level cognition. Part 3 was designed to identify the participants’ comprehension, problem-solving, and strategy skills. The results indicate that participants’ achievement of learning comprehension knowledge was enhanced. A number of research studies have been developed and explore better approaches to prompt student’s higher-order thinking. With this finding, it may be concluded that computer-based video games can be considered an instructional approach to improve students’ higher-order thinking.

In conclusion, computer-based video games might improve students’ achievement in learning factual knowledge, problem-solving strategies, and higher-level cognitive thinking processes.

**Limitations**

Five limitations to this study became apparent when conducting the computer-based video game research and after reviewing the results.

First, all third-grade students had attended computer lessons and their ability to use basic computer tools and Internet-browsing resources were a prerequisite for participating in this study. One issue needs to be considered to ensure the validity of research design: for some students, Fire Captain was their first experience playing a computer game. Playing computer games is different from using an Internet-browser, and that could cause a problem for some students. Whether a student can successfully play the game in 40 minutes became critical for the willingness and ability of a student to finish this game.
Second, a 40-minute class for developing complex higher-order thinking and a post-test within 24 hours for measuring participants’ outcomes may be inadequate. The amount of time for instruction should be taken into consideration for children’s play in play theory. In particular, how much time a student spends on play or practice could be a major factor influencing children’s motivation for learning and achievement. After all, human motivation is a function of human cognition involvement (McDevitt & Ormrod, 2004, p. 462). However, this study did not intend to address this question, although it will be an important consideration for future studies.

Third, the content of instructional treatments for this study could have been more objective. Fire Captain is not a game designed for instructional purposes. Also, the possibility of designing a website that could perfectly match all content with a computer game is likely impossible. In particular, different hypertext developed by different designers may result in different formats. To ensure that the design of the hypertext format and the content was consistent with the game, the instrument’s validity needs to be examined further.

Fourth, from the results of this study, computer-based video games could function as powerful tools in assisting children’s cognitive processes. However, an unanswered question regarding what could be the improved major cognition factor arising from playing computer-based video games — for example, short-term memory or long-term memory — needs further investigation.

Fifth, one methodological weakness of this experiment is that a sufficiently larger number of participants would be better so that the statistical test of effect size would have a higher partial Eta-squared value. However, the sample size was limited by the number of computers in a computer laboratory and the consent of students’ parents. In Taiwan, the fact that some parents keep their children away from computer-based video games is well known. Without parents’ understanding and support, difficulty arises for having a sufficiently large number of participants for similar future studies.

Conclusions

Different from previous studies that examined computer-based video games, this study was developed to investigate the effects of computer-based video games on the cognition of elementary-level students. Based on the aforementioned findings and the limitations of this study, the following conclusions may be drawn.

First, computer-based video games were applied to cognitive improvement. The results from this study provided experimental evidence to support that the use of computer-based video games can facilitate students’ cognitive learning process. Few prior research studies dealing with computer-based video games explored the idea of learning. In fact, computer-based video games have often been criticized for reproducing violent action or behavioral aberration in most previous research reports. As this new form of technology has been treated as an entertainment device with negative influences on children, this study was attempted to restore computer-based video games to its original value: multifunction interactive media. In addition, given this approach, a different character emerges by suggesting multimedia’s functionality and instructional capability should be considered equally with CAI programs.

Second, of great theoretical and practical importance, children’s development and learning could benefit from play. Even without the emphatic educational intention such as CAI programs, children can still improve their cognition through playing computer and video games. Kids’ play with computer games exemplifies forms of knowing and being in the world that are irreducibly and simultaneously social, technical, material, and symbolic (Mizuko, 1997).

According to the findings of the study, future research should continue to investigate the impact of computer-based video games along with different instructional strategies on varied children’s learning achievements, such as facts, concepts, comprehension, problem-solving, or critical-thinking skills. Additionally, future studies should consider human factors in a gaming environment, such as learners’ individual differences, learning styles, preferences in learning visual/audio materials, field dependence/independence, etc. Many of the independent variables associated with the study of human-computer interaction should be taken into account in the design of computer-based video games.

While computer-based video games may be manipulated to positively influence children’s learning, particular attention must be given to guidelines derived from game design and experimental methodology, as well as to learner
characteristics and learning styles. Only by initiating a systematic program of investigation where independent variables are judiciously manipulated to determine their relative effectiveness and efficiency of facilitating specifically designated learning objectives will the true potential inherent in game design for learning be realized.

References


Tzeng, S.-C. (1999). *Optimizing challenges and skills in the design of an educational computer game and exploring adolescents’ gaming beliefs*, Unpublished PhD thesis, University of Georgia, USA.


Yoder, E. P. (2002). *Basic multivariate analysis of variance*, The Pennsylvania State University at State College, USA.
Using Blogs to Enhance Critical Reflection and Community of Practice

Shih-Hsien Yang
Department of Applied Foreign Languages, National Formosa University, YunLin, Taiwan // Tel: 886-5-631-5822 // Fax: 886-5-6365927 // Email: shiyang@nfu.edu.tw

ABSTRACT
Using the theories of critical reflection and community of practice, the aim of this paper was to explore the use of blogs as a reflective platform in the training processes of English as a Foreign Language (EFL) student teachers, who were learning to teach English for future employment in Taiwan. They made use of blogs as a platform to critically reflect on their learning processes as well as to gauge the impact of blogs on their own professional growth. Forty-three student teachers in two teacher-education programs at two science and technology institutions in central Taiwan participated in this study. Two instructors created a blog for use as a discussion forum so that the student teachers could engage in and examine their own reflection process. The data collected was qualitative in nature, consisting of student teachers’ posting messages and comments on the blog, surveys on the student teachers’ reflective experiences using blogs as reflection tools, and group reflective dialogues recorded by instructors in class meetings over the implementation of blogs during the course. The results showed that the student teachers actively discussed teaching theories and their implications through blogs. All of the 43 teachers who took part in this study were reflective, and some critically reflected on their thoughts and made significant comments; and the participants considered technology a useful platform for reflecting and communicating with each other. The positive implications for the use of blogs as a medium to provide and promote critical reflection for EFL teachers are discussed.

Keywords
Blog, Hypermedia instructional programs, Virtual community, Reflective journal for language learning, Technology for language learning, Computer-mediated communication

Introduction
Learning through reflection
Ng, Lan, and Thye (2004) point out that a professional teacher is one who regularly reflects on his or her teaching through critical thinking. According to Dewey (1933), reflective thinking is “the kind of thinking that consists of turning a subject over in the mind and giving it serious and consecutive consideration” (p. 3). Burnett and Lingam (2007) show that critical reflection helps teachers and administrators rethink the purposes of education and reshape the programs to meet students’ needs. Brandt (2008) also indicates that when feedback and reflection are integrated in the form of reflective conversations between teachers and students, both teachers and students could benefit from the reflective practice. Sockman and Sharma (2008) also show that through peer feedback, reflective journal writing, and reading, teachers could uncover the obstacles and discover how their teaching beliefs need to change in order to implement transformative teaching strategies. Therefore, reflective practice not only makes change possible, but also provides information needed to develop guidelines for setting new needs, goals, and plans.

The process of reflection includes reflective thinking and self-examination during or after teaching (Liou, 2001). Bailey, Curtis, and Nunan (2001) similarly divide reflective practices into reflection-in-action and reflection-on-action, where reflection-in-action teachers continually examine their teaching process to make any suitable changes in their pedagogy, and reflection-on-action teachers plan their teaching process in advance and then evaluate their teaching process afterwards. Critical reflection, as opposed to mere reflection, refers to how teachers learn to challenge their own teaching beliefs in a critical self-analysis and become responsible for their actions (Korthagen, 1993; Sockman & Sharma, 2008). Critical reflection on experiences allows teachers to develop a deeper understanding of themselves and their students. Liou (2001) expresses that critical reflection raises teachers’ awareness about their own teaching and triggers positive changes. Critical reflection fosters the most effective teacher interaction in a professional setting by encouraging teachers to take a stand through questioning and challenging others’ underlying assumptions, which is a way for teaching practices to be improved and for the conditions in which schooling takes place to be made more just (Carr & Kemmis, 1983; Manouchehri, 2002; Burnett & Lingam, 2007).
However, critical reflection and inquiry do not come naturally to most teachers, so the appropriate opportunities for reflection should be provided to teachers who are attempting to absorb knowledge at their initial stage of learning to teach and to bring newly learned knowledge back to their classrooms. Researchers also point out that peer collaborative reflection helps each individual develop his or her professional knowledge (Manouchehri, 2002). Writing journals or learning logs have gained popularity over the last decade as educational tools (Commander & Smith, 1996). Journal writing is effective in developing students’ metacognitive and reflective skills (Morrison, 1996). Furthermore, collaborative writing affords students the opportunity to share ideas and provide each other with feedback (Storch, 2005). A blog seemingly transpires to be a helpful tool for teachers to establish an encouraging and nurturing discussion space that enhances students’ reflection.

In summary, reflective teaching and reflective practices play important roles in teacher education. These processes prepare teachers to think critically and reflectively, enabling them to enter their careers with the capability for continuous professional development (Lester & Mayher, 1987; Mule, 2006). Lester and Mayher comment on critical inquiry and reflection as follows: “To be a professional is not to have all the answers. Rather, a professional is someone who can reflect on tentative solutions, collaborate with others on the possible avenues available, and risk making mistakes because mistakes are an inevitable part of building new roads” (p. 209).

Communities of practice and reflective practice

Recently, computer-assisted learning environments have provided both teachers and learners with an alternative avenue, unlike the traditional face-to-face meetings or discussions, to foster their personal knowledge development through meaningful negotiation and communication (Lord & Lomicka, 2004; Ahern et al., 2006). One of the most successful models is the application of community practices in which group members are able to explore their knowledge and exchange information through a synchronous discussion board, such as MOO and BBS (Godwin-Jones, 2003). The web-based technology supports collaborative learning that enriches learning performance, both for individual knowledge construction and group knowledge sharing (Liaw, Chen, & Huang, 2008).

Wenger (1998) submits an overall and clear conception of what communities of practice are. The definition of community of practice varies depending on what the community is about, how it functions, and what capabilities it has enabled. Members’ beliefs and interests exercise a profound influence upon the formation of communities of practice. In other words, a successful community of practice is one in which the contribution of each member is highly regarded. The nature of the community of practice, its goals, and its communicative procedures are all co-determined by the individual members in a way that allows for continuous change and self-development. In consequence, a community of practice usually serves as an effective platform for people to exchange knowledge and localize new information based on their personal needs and living environment. Furthermore, a community of practice tends to encourage every member to take responsibility for information-sharing and problem-solving, to develop their personal identities in the community, and to foster unification of the community. Given that reflective practice is “the relationship between an individual’s thoughts and action and the relationship between an individual teacher and his or her membership in a larger collective called society” (Leo, 1990, p. 204), a community of practice could be an alternative avenue for reflective practice among teachers. This was one of the key connections investigated in this study.

According to Zeichner and Liston (1996), by means of a community of practice teachers are able to 1) examine and attempt to resolve their teaching problems, 2) realize their teaching beliefs and goals, and 3) take responsibility for their professional development through continual participation in the community. Stiler and Philleo (2003) report the successful use of blogs for reflective practice among preservice teachers based on their findings from self-report questionnaires. They claim that reflection through blogs empowers teachers to give and receive more positive and immediate feedback from peers to resolve personal and professional problems. Therefore, teachers have more opportunities for critical reflection upon their own and others’ teaching and learning through the use of blogs.

Blogs

Due to the rapid expansion of technology over the past decade, hypermedia instructional programs have become commonplace in both educational and business markets (Kozma, 1991). Learners of all ages have increased access to
technology such as the World Wide Web, which provides a seemingly limitless amount of information. New technologies such as Mp3, podcasting, and social networking are booming. In contrast to more traditional environments, technology offers greater opportunities for interactivity and learner control (Kozma, 1991; Rodzvilla, 2002). There are more educators and language teachers using the Internet in language teaching as well (Godwin-Jones, 2003; Lord & Lomicka, 2004). Many computer applications, especially asynchronous computer-mediated communication such as email and electronic bulletin boards, promote interactive learning (Arnold & Ducate, 2006). With the booming growth of technology, blogs have become another learning platform for language teaching (Richardson, 2005).

A blog is an online journal that users can continuously update, in their own words, online (Matheson, 2004). Blogs utilize a simple interface to make it easy for any user to construct, without having to understand HTML or web scripting. Thus, anyone who can create a basic Microsoft Word document can create and maintain a blog. In addition, users can even add pictures or audio files to enhance their blog’s attractiveness. Furthermore, a blog is interactive (Rodzvilla, 2002) in the sense that readers can respond with comments in just a few steps.

Most blogs on the Internet are personal or journalistic (Godwin-Jones, 2003; Richardson, 2005). However, there have been increasing numbers of people using blogs in education (Richardson, 2005). Blogs are well suited to serve as online journals for users (Godwin-Jones, 2003; Richardson, 2005). In terms of language teaching and learning, “language learners could use a personal blog linked to a course as an electronic portfolio, showing development over time” (Godwin-Jones, 2003, p. 13). Hence, blogs could be used to monitor and assess students’ work as well as to encourage interaction among students and between teachers and students. The following sections outline the characteristics of blogs that make them useful tools for language teaching and learning.

Blogs can stimulate reading and motivate learning

Many studies have shown that it can be difficult to motivate language learners when it comes to reading texts (Kozma, 1991; Ho & Richards, 1993; Dewitt; 1996; Davis, 1997). Studies have also shown that Internet access motivates many students to read extensively (Rodzvilla, 2002; Stiler & Philleo, 2003; Liaw, Chen, and Huang, 2008). The use of blogs is a way to provide such motivation for reading in a language other than one’s mother tongue, through the interactive nature of the blog. One can read and also comment on what one reads in expectation of a little discussion and a quest for common interests and individual differences. By responding on blogs, people can get feedback from other audiences throughout cyberspace. Students have an opportunity to read things in which they are interested and write things they truly wish to write, thereby determining their own texts in language education and combining text with conversations in a very personal and stimulating way. Efimova and Fiedler (2003) characterize blogs as “personal diary-like-format websites enabled by easy to use tools and open for everyone to read” (p. 490). Godwin-Jones (2003) also indicates that through blogging, people are able to document their reflections about things relevant to their daily life experiences, sharing such things with their friends, families, and/or group members. He further points out that blogs and other social networking sites provide new opportunities and incentives for personal writing (Godwin-Jones, 2008). In other words, blogs allow people to exchange information without space and time constraints, to broaden their knowledge, and to meet personal needs and interests at the same time.

Community building through blogs

Computer users with access to the Internet can access some blog sites that are open to the public and welcome viewers to make comments or post messages. Language teachers can use blogs to ask students questions, share viewpoints, and encourage students to discuss issues and express their concerns. Learners are allowed to collaborate with others to establish a particular topic of mutual interest. More specifically, a blog is constructed by people who share mutual interests to collaboratively set objectives, regulations, and formats, and this is what distinguishes blogs from other forms of websites (Godwin-Jones, 2003, Richardson, 2005). A blog is like a small learning community (Efimova & Fiedler, 2003). Members tend to get more involved than they do in other pedagogic and web-based environments, thus producing a stronger sense of community (Wegner, 1998; Godwin-Jones, 2003; Efimova & Fiedler, 2003; Godwin-Jones, 2008). For example, a pair of readers was highly motivated to read different types of
novels, so they collected any related information and compiled it to share the information with novel lovers. A global community therefore forms through blogs.

**Blogs provide hyperlinks to other resources**

According to Udell (2001), blogs are a “genuinely new literary/journalism form” (p. 2). In addition, blogs can make use of other Internet resources for learning. Users and teachers can create more learning resources by adding hyperlinks within the blogs (Godwin-Jones, 2003). Efimova and Fiedler (2003) explain that the use of hyperlinks makes it possible for new readers to learn the previous course of the discussion more easily and to make the blog more enjoyable with pictures or sounds.

There are many ways in which users or teachers can create more learning resources by adding hyperlinks in their blogs. For example, teachers can have links to online testing sites to allow students more time to practise certain tests and exams. In addition, teachers can make good use of other sites embedded with audio and visual materials to enhance students’ learning motivation. As a result, students gain more knowledge through links to many sites if they choose to follow the hyperlink. The inclusion of these hyperlinks also allows for self-directed exploration within the topic (Efimova & Fiedler, 2003).

**Blogs provide a learning space**

Blogs do not merely establish scaffolding for beginners, nor do they merely allow students to learn from multiple perspectives or receive support from advanced students. They also create a relatively learner-centered environment that allows students to learn at their own pace (Efimova & Fiedler, 2003; Godwin-Jones, 2008). However, if the information that the students are learning is incorrect, then all of the students are being misinformed.

These points have already been made in the discussion above, but here we draw attention to the way in which blogs pull together several of the most recommended pedagogies from learning theory: scaffolding, student-centered learning, the incorporation of multiple perspectives, and the use of learning communities. However, the spreading of misinformation is self-governed within the community, and it’s the responsibility of people adding hyperlinked material to ensure that the material doesn’t contain misinformation.

In summary, blogs can be treated as virtual language classrooms. People from all over the world can share opinions and express ideas by using a language they all understand. Language teachers can use blogs as discussion forums to increase time and opportunities for students to learn the target language. While blogs are used in education, there is little research about the use of blogs for language learning and teaching in EFL contexts.

**The study**

**Participants**

This study focused on 43 EFL student teachers in two teacher training programs at two science and technology institutions in central Taiwan. Twenty-eight students were from one school and fifteen were from the other. The student teachers who were participants of the study were 22-year-old junior students who enrolled in the teacher training programs at two different universities in the fall semester of 2005. The courses they took were second- and foreign-language teaching methodology. The courses mainly covered major theories concerning English learning, teaching methodologies, and practical teaching.

**Setting**

The study was a self-study of two English teachers’ classrooms. Two instructors organized their class schedules similarly. Forty-three student teachers met weekly for two hours to learn language-teaching methodology in the first half of the semester (nine weeks) and practised teaching in class in the second half of the semester (nine weeks). The
assignments and the settings for these two classes were the same. These student teachers were required to design lesson plans, practise teaching, and keep reflective journals.

The instructors created a blog to be able to communicate with all 43 student teachers. In the first class meeting, the instructors explained the syllabus and introduced concepts such as critical thinking and reflection. Every class member was asked to record his or her thinking and reflections about each class, including the teaching methods they learned. After each class meeting and practical teaching session, participants were required to go online and write a reflection of their teaching experience, while electively making comments or expressing thoughts about their peers’ messages. The two instructors also commented on the student teachers’ postings and challenged their thinking by asking pertinent questions for the students to consider.

**Research questions**

The aim of the study was to investigate the reflection process in blog postings. In order to see if participants could critically reflect on what they have learned and enhance the effectiveness of community of practice, three specific research questions served as guides in the data analysis:

1. What types of reflection were involved in student learners’ reflection? Were they descriptive or critical?
2. What were the teacher trainers’ roles in the process of blogging?
3. How can a blog promote critical reflection and community of practice?

**Methodology**

**Data collection**

The primary data for this qualitative study consisted of: (1) student teachers’ posting messages and comments on the blog; (2) group reflective dialogues recorded by instructors in class meetings over the implementation of blogs during the course; and (3) an end-of-semester questionnaire given to each student teacher.

The end-of-semester questionnaire was given to all student teachers to gather their opinions about blogging following their experience with it. Other data were examined as well, including email dialogues among group members, other teaching information shared in class, and reflective journals from both the student teachers and the two instructors.

**Data analysis**

Postings on the student teachers’ blog were sorted into five categories and then analyzed with supplementary analysis of the other data sources in relation to the same five categories. The five categories were derived through a modification of Ho and Richards’ (1993) framework (See Appendix 1) for qualitative research on student teachers’ journals. Ho and Richards’ framework includes four categories that match topics that the two instructors in this study covered in their courses: (1) theories of teaching, (2) instructional approaches and methods, (3) evaluations of teaching, and (4) teachers’ self-awareness of strengths and weaknesses. However, as Ho and Richards’ framework focused on the in-service teachers’ views on the teaching theories, a modification of this framework to fit the interests of this study, focusing especially on the views from those EFL pre-service teacher participants, resulted in the following five categories:

4. Theories of teaching
5. Instructional approaches and methods used
6. Teaching evaluation methods and criteria
7. Self-awareness
8. Questions about teaching and requests for advice

Data was organized into these categories and analyzed. A more detailed explanation of each category is provided immediately below:
1. **Theories of teaching**: Comments made or posted by the student teachers about the second-language acquisition theories they were taught in the courses.

2. **Instructional approaches and methods used**: Comments and postings of the student teachers’ referring to their own use of methods in the classroom, as well as their expression of beliefs and knowledge related to these practices.

3. **Teaching evaluation methods and criteria**: The feedback and discussion provided by the whole class after each student practised teaching in front of the class.

4. **Self-awareness**: Comments, postings, and discussion based on self-awareness and self-evaluation.

5. **Questions about teaching and requests for advice**: All questions asked by the student teachers that pertained to teaching practice and theory, plus all requests for advice made by the student teachers.

In the results section below, the analysis based on this sorting procedure is presented in relation to the research questions.

### Results

The findings of this study are arranged by the three research questions presented earlier in this paper.

**What types of reflection were involved in student learners’ reflection? Were they descriptive or critical?**

Table 1 displays the topics reflectively addressed on the blog. Generally, there were both descriptive and critical reflections on their own teaching and on others’ teaching after observation. Table 1 shows the topics and number of student teachers discussions on the blog.

<table>
<thead>
<tr>
<th>Topic category</th>
<th>Total number</th>
<th>Descriptive</th>
<th>Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theories of teaching</td>
<td>324</td>
<td>243</td>
<td>81</td>
</tr>
<tr>
<td>Approaches and methods</td>
<td>106</td>
<td>73</td>
<td>33</td>
</tr>
<tr>
<td>Evaluation teaching</td>
<td>125</td>
<td>53</td>
<td>72</td>
</tr>
<tr>
<td>Self-awareness</td>
<td>354</td>
<td>202</td>
<td>152</td>
</tr>
<tr>
<td>Questions about teaching</td>
<td>68</td>
<td>31</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>977</td>
<td>602</td>
<td>375</td>
</tr>
</tbody>
</table>

Number of students = 43

* indicates where the number of critical reflections was greater than the number of descriptive reflections

From Table 1, it is obvious that the number of descriptive reflections far exceeded the number of critical ones. Additionally, there is evidence that students were more comfortable critiquing themselves than critiquing others. During the study, participants often questioned the applicability of the second-language teaching theories and methods used in the EFL contexts. Participants brainstormed a lot of solutions and adjustments for EFL classrooms. Some agreed that EFL teachers should combine the advantages from those teaching theories and create a new improved theory for the classes. Some even complained about the EFL environment for learning the target language. They pointed out the fact that once learners step out of the language classrooms, they don’t speak or practise the target language.

In addition to writing, because everyone could access and read the blog postings, many participants pointed out that their English writing skills improved greatly. Because they were afraid of losing face, they would double-check the content and grammar before posting on the blog.

**What were the teacher trainers’ roles in the process of blogging?**

During the study, every time these student teachers posted their reflections on the blog, the two class instructors would go online and read those reflections. Sometimes, the instructors would challenge student teachers’ thinking by
posting questions and asking for further reflection in order to raise participants’ critical reflection. Twenty student teachers reported that due to such challenges set by the instructors their thinking went deeper and became more critical. Table 2 displays the number of student teachers’ critical reflections in each category in these two classes and the increased number of reflections each time after the instructors intervened.

<table>
<thead>
<tr>
<th>Topic category</th>
<th>Total number</th>
<th>Class A ( n = 28 )</th>
<th>Class B ( n = 15 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theories of teaching</td>
<td>81 (21.6%)</td>
<td>33</td>
<td>48</td>
</tr>
<tr>
<td>Approaches and methods</td>
<td>33 (8.8%)</td>
<td>10</td>
<td>23</td>
</tr>
<tr>
<td>First instructor intervention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluating teaching</td>
<td>*72 (19.2%)</td>
<td>41</td>
<td>31</td>
</tr>
<tr>
<td>Self-awareness</td>
<td>*152 (40.5%)</td>
<td>68</td>
<td>84</td>
</tr>
<tr>
<td>Questions about teaching</td>
<td>37 (9.8%)</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>375</td>
<td>169</td>
<td>206</td>
</tr>
</tbody>
</table>

* indicates that the number of postings increased after the instructors challenged student teachers’ thoughts

During the course, the instructors not only pushed their student teachers to think further but also encouraged them to express more. When two instructors were not satisfied with the level of thinking and critical reflection displayed in blog postings, they asked “what do you think…?” and “what would you do if…?” to foster their student teachers’ thinking. Table 2 shows that initially 22 percent of the total number of reflections (81/375) were critical but for the second topic, the posting numbers dropped. After the first instructor intervention, this percentage steadily rose from 19 percent and eventually reached 40 percent. The two instructors also provided explanations when students received or posted misinformation on the blog and said that a blog was a good platform for two classes from different universities to share ideas with each other and exchange thoughts.

It is not difficult to see that teacher educators did play a vital role in promoting these student teachers’ views. The two instructors gave feedback and developed their student teachers’ own views and a capacity for self-critique and self-reflection. In this study, the roles of these two instructors were like those of directors and counselors. When instructors found the discussion not active enough, they took initiative and wrote reflections on their own teaching for this particular class in order to model the process. Thus, student teachers perceived the instructors as active learners and instructors who continuously reflected critically and made themselves vulnerable to share their thoughts with student teachers. The instructors served as role models to demonstrate to what extent critical reflection can lead to more thinking. They also built a bridge for two classes to communicate with each other by blogging without the restrictions of time and space.

**How can a blog promote critical reflection and community of practice?**

Instructors did promote critical reflection and community of practice on the part of their students using the blog tool, but the extent of this critical reflection and its duration remains problematic. The student teachers expressed many concerns and asked a lot of questions to each other during the semester. Many of them pointed out that due to the use of a blog for discussion, there were no limitations of time and space, and discussions were therefore easier and more inviting. Generally speaking, nearly two thirds of the 43 student teachers expressed very positive attitudes toward the use of blogs as a platform to reflect their learning and teaching. Twenty-one participants said that it was easy and comfortable to post comments and challenge their peers on the blog instead of discussing the issues in a face-to-face context, allowing comments to be more critical. Two classes’ students could exchange thoughts and ideas as well as sharing experiences in the blog.

During the course, members from two classes participated actively. All student teachers appreciated the opportunity to use technology to communicate with people from different classes. Many student teachers wanted to post concerns about EFL issues to other English learning and teaching associations. In addition, the two instructors sometimes made reflections and comments or posed questions in class meetings. They also asked their student teachers to pay attention to others’ writing styles. The two instructors spent a little time giving comments on student teachers’
grammatical errors and basic writing structure. By so doing, the two instructors believed their student teachers could reflect not only what they learned but also practise basic English skills through the course. Furthermore, student teachers could examine their reflections and thinking by dates on blogs. They could see their learning process and perspectives on certain issues at different periods of time. Therefore, the blog was considered a great tool for these student teachers to record their growth and changes as well as build a learning community.

In summary, blogging did provide a more flexible time and space for student teachers to reflect and discuss once instructors encouraged student teachers to participate in blogging. Posting messages did lead to a kind of inquiry that accentuated critical reflection. Community of practice is driven by this process of stepping back, reflecting, and analysing (Freeman, 2000). Johnson and Golombek (2002) further “[renders teachers] increasing control over their thoughts and actions; grants their experiences enriched, deepens meaning; and enables them to be more thoughtful and mindful of their work” (p. 7).

Anonymity on the Internet

According to the survey collected from the student teachers and two class instructors, the issue of revealing identity on the Internet remained a concern during the whole semester. The two instructors reported that students were afraid that their messages would be read and might hurt their friendships among class members even though all students used pseudonyms on the web. Some students even felt that they would jeopardize their final grade by posting some aggressive responses toward this course.

In conclusion, the two instructors believed that anonymity is a big issue when grades and friendships are at stake. Cultural differences and educational background might be the other reasons, as highlighted by the Class A instructor. The instructor indicated that most Taiwanese students are taught to be moderate and gentle, based on Confucian teachings. In a Confucian society, a good student is supposed to be diligent, persevering, well behaved, and obedient to authority (Tamney & Chiang, 2002). Therefore, speaking sharply (questioning and/or challenging) is not considered a good personality trait. Unlike western education, criticizing or questioning in class might not be favored in Taiwan’s education. However, the instructors stated that the education in Taiwan is changing toward a more western style. Traditional thinking and behaving in classrooms are changing through the influence of western culture. Questioning in class is becoming more welcome.

The two instructors in this study gave positive views of using blogs as a medium for reflection; however, encouragement and intervention were needed from time to time to push student teachers to think further and become more comfortable with expressing critical thoughts.

Implications

When student teachers came together to discuss or give feedback on each other’s work and teaching, it was not clear that they would engage in critical reflection. However, by using blogs as a platform for reflection, participants got more opportunities to make comments and challenge each other’s viewpoints. They could still converse about or express what had been left out in the traditional classrooms. Quite a few participants admitted the usefulness and the convenience of using blogs to reflect and give comments. They even suggested voicing the EFL issues and concerns to other educational organizations. The findings displayed the importance of feedback and risk-taking in language education, which can reinforce students’ confidence and motivation in language learning.

This study indicates that the blog studied here demonstrates a community of practice in that it was used by the participants as a discussion space. It was a forum that prepared each of the student teachers to relate theories to practice by discussing beliefs, learning from each other, demonstrating to each other how they would act in their actual classrooms. The blog was a place for these student teachers to voice their doubts, struggles, discomforts, and successful and unhappy teaching and learning experiences because the participants shared very similar experiences of being EFL language teachers and learners. Dewey (1933) claims that if we want to make our experience educative, it is essential to support ongoing growth in a process of continuing new inquiry. Freire (1985) argues that inquirers not only are problem-solvers but also problem-positers. The posing of questions in the process of their inquiry served as the anchors to help these student teachers explore further. As Short, Burke, and Harste (1996)
propose, “As we work through inquiry, we do not usually end with one answer or even a set of answers. Inquiry does not narrow our perspective; it gives us more understandings, questions, and possibilities than when we started. Inquiry isn’t just asking and answering a question. Inquiry involves searching for significant questions and figuring out how to explore those questions from many perspectives” (p. 8–9). Getting together in a group armed with the same concerns for EFL contexts enabled these student teachers to reflect and identify the lingering questions, work together to think through the questions, and push their thinking further as a group.

As a result, learners generated more inquiries that would take their conceptions further. The negotiation phase made it possible for the student teachers to make sense out of the new knowledge and ponder its potential utilization in EFL contexts. Also through group discussions in class meetings, one issue tended to raise another, propelling the inquiry towards a clear comprehension of EFL pedagogy.

**Acknowledgments**

The funding for this research was provided by the National Science Council of Taiwan, ROC under Grant NSC97-2410-H-150-006.

**References**


### Appendix 1: Differences between descriptive and critical reflection

<table>
<thead>
<tr>
<th>Topic category</th>
<th>Descriptive</th>
<th>Critical</th>
</tr>
</thead>
</table>
| Theories of teaching                         | 1. A belief/conviction  
2. How a theory was applied                                                   | 1. A justification  
2. A personal opinion  
3. Contradictory practice between theory and practice |
| Instructional approaches and methods used    | 1. Approaches and methods  
2. The content of the lesson                                                    | 1. The teachers’ knowledge in teaching  
2. Sociopolitical impact                                                          |
| Teaching evaluation methods and criteria     | 1. School context/classroom management  
2. Solutions to problems: from experts                                             | 1. Evaluating lessons  
2. Diagnosing problems  
3. Solutions to problems: alternative ways                                          |
| Self-awareness                               | 1. Perceptions of self as teacher: style and comments on language proficiency | 1. Recognition of personal growth  
2. Setting personal goals                                                            |
| Questions about teaching and requests for advice | 1. Asking for advice                                                        | 1. Asking for reasons  
2. Problematizing                                                                   |

(Adapted from Ho and Richards, 1993)
Automated Inattention and Fatigue Detection System in Distance Education for Elementary School Students

Kuo-An Hwang\(^1\) and Chia-Hao Yang\(^2\)

\(^1\)Department of Computer Science and Information Engineering & Graduate Institute of Networking and Communication Engineering // \(^2\)Graduate Institute of Informatics, Doctoral Program // Chaoyang University of Technology, Taiwan // kahwang@cyut.edu.tw // s9433901@cyut.edu.tw

ABSTRACT

Most courses based on distance learning focus on the cognitive domain of learning. Because students are sometimes inattentive or tired, they may neglect the attention goal of learning. This study proposes an auto-detection and reinforcement mechanism for the distance-education system based on the reinforcement teaching strategy. If a student is detected to be inattentive or fatigued, then the alert reinforcement feedback window is given; if the attentive time in learning has been reached, then encouraging reinforcement feedback is given. In distance education, the teacher could maintain the learning quality of students through this mechanism. Results of the simulative test and class experiment indicate that this mechanism detects the behavior of students with high accuracy. In other words, it can accurately identify learners who are inattentive or fatigued. The proposed mechanism can urge the inattentive learners to achieve the Attending Teaching goal, according to a comparison of results before and after the experimental class obtained from questionnaires and interviews. However, we also found that the mechanism could not effectively alleviate mental fatigue resulting from physiological exhaustion. Thus, students who feel fatigued must be required to learn in a better mental state to avoid wasting time and learning without satisfying effects. This work presents a detection reinforcement mechanism and teaching procedure designed from the perspective of an Affective Domain Teaching Objectives. The proposed mechanism can accurately detect inattentiveness in students, thus helping teachers of distance education courses to handle the students’ learning conditions.

Keywords

Distance education, Image processing, Computer-based assessment, Affective domain teaching objectives

Introduction

Distance education enables learners to learn anytime and anywhere. In an attempt to meet the needs of the public in learning, various distance education websites, including teaching materials, teaching platforms, on-demand videos, etcetera, have been constructed in schools of different levels and related education institutes (Harasim, Hiltz, Teles, & Turoff, 1995). Students can access the websites freely, and decide for themselves how much time they wish to spend learning online. They can earn credits as long as the learning prescribed by the course is completed during a specified period and the related reports are handed in or the assessments are completed.

However, students might log in to an online course when they are fatigued or inattentive. Because the students are far away and out of the teacher’s control, teachers cannot effectively supervise students with poor self-discipline. The circumstances in which students need to learn by distance learning include suspension of classes because of an outbreak of an infectious diseases, absence from school due to illness, or famous teachers (at home or overseas) being invited to give lessons through distance-learning technologies. Some students, especially pupils, might not have strong learning motivation or high self-control when their parents are too busy to supervise them. The students might not learn effectively when they are logged into the system. They might leave their seats to do something else or fall asleep. Teachers and parents cannot easily control the students’ learning environment, especially for younger students who are more easily distracted by environmental factors. The accumulated hours of attending the class cannot guarantee that students are learning attentively. Therefore, even if the student is distracted or feels fatigued after he or she has logged into the system, the teacher has no way of knowing this, and has no way of encouraging the student to learn attentively. Since the student has completed the required number of hours and has completed both the schoolwork and the assessment, the student can still earn the credit. Nevertheless, credits earned this way have little value. This is a common flaw of distance education (Mason, 1997; Rourke, Anderson, Garrison, & Archer, 2001).

If the student has neither left midway during the class, nor had another person attend the class as a substitute, but is learning attentively without feeling fatigued, then the accumulated class hours can be considered to have produced
good learning effects, indicating that the distance education is of high quality. However, most current distance education research bases participation on the number of times a student contributes to a class discussion or the number of hours the student has spent in class (Pena-Shaff, Martin, & Gay, 2001). Satisfactory attendance and class hours are not equivalent to attentive participation (Mason, 1997). Hence, no technique is available to detect conveniently and precisely a distance-education student’s attention to learning.

Few studies have combined image recognition with the reinforcement mechanism in the distance education system. Webcams and image-recognition technology have been adopted to capture facial images in real time in order to detect fatigue in drivers (Eriksson & Papanikotopoulos, 1997; Grauman, Betke, Gips, & Bradski, 2001; Hamada, Ito, Adachi, Nakano, & Yamamoto, 2002). However, these techniques are mostly used to detect fatigue. Because the detection conditions are different from those of distance education, these techniques could not be applied to distance education directly. Furthermore, inattentive learning in distance education includes not only fatigue, but also other inattentive behavior, which must be assessed comprehensively before it can be judged. Effective management of the learning conditions of students in a distance education environment will help solve the problem by detecting students’ inattention and fatigue, and thus maintain good learning quality.

The conditions of students can be detected precisely by integrating these techniques with the development of the detection and reinforcement mechanism for distance education. After conducting Bayesian network assessment, if a student is found to be inattentive or fatigued, then the alert feedback is sent instantly (Grauman et al., 2001; Hamada et al., 2002). Additionally, encouraging feedback is sent if the learning attention of the student has lasted a period of time. The learning conditions of all students are recorded to promote good behavior in distance education students with above reinforcement principle, so that the teacher could also take hold of the attending condition of the students in class in order to achieve the learning attendance goal. Students who want to earn the degree or credit through examination would necessarily study hard. Students with a lower level of self-control would need to be supervised by teachers or parents. However, teachers and parents cannot always supervise. The distance education system needs the assistance of a reinforcement mechanism when students learn passively in distance education.

According to the teaching theory of operant conditioning, the application of a reinforcement principle in class can ensure that the learning behavior of the students is supervised and handled well (Goge & Broiler, 1992; Woalfolk, 1997). However, in distance education, since a teacher cannot easily observe the learning condition of students, teachers encounter difficulty in applying the reinforcement principle to encourage good learning behavior. Because elementary-school students might easily be distracted by the environment, or doze off, distance education cannot yet replace classroom education. Currently, distance education is mostly provided for adults or as remedial teaching, since it requires students to be in a good learning state to achieve optimum learning results. In this study, image processing and detection techniques enable distance-education teachers to manage the students’ learning condition. Remediing the defect of traditional distance education, which hinders teachers’ ability to take hold of the learning condition of students, means that the interaction in distance education courses becomes much more like that in traditional classes. Moreover, detection reinforcement mechanisms and teaching procedures are also designed to cope with the existing defects of distance education from the perspective of the Affective Domain Teaching Objectives. In addition to improving the accuracy of the detection mechanism and the actual learning attending effects of the reinforcement mechanism, detection simulation and class experiment also help distance education courses achieve good learning attendance. This work focuses on detecting inattention and fatigue in students, but not on the reasons for these behaviors. Hence, the detection process does not attempt to discriminate among detailed behavior. Teachers would need to perform such detailed detection.

Section II explores the operant conditioning theory, reinforcement strategy and image recognition technique. Section III discusses the design of the auto-detection and reinforcement mechanism, as well as the detection, recognition and reinforcement procedures for distance education. Section IV analyzes the auto-detection and reinforcement mechanism based on a practical simulative test and experiment class. Section V discusses the evaluation results and studies the effects of the mechanism on students’ attentiveness to learning in class. The final section draws conclusions and makes suggestions for future research.

**Literature review**

Many scholars have asserted that the Affective Domain Teaching Objectives must not be neglected when enhancing learning effects and cultivating a healthy personality (Gagne, Briggs, & Wager, 1985). The following purposes can
be deduced from the Affective Domain Teaching Objectives and the time to achieve goals at each stage: the learning objectives of each class period for achieving the attending and responding stage, the objectives of every unit course for obtaining the evaluation and organizing stage, and the objectives of a semester or academic year for achieving the characterizing stage (Lefrancois, 1982).

The reinforcement teaching strategy consolidates learning behavior by employing positive or negative stimuli in accordance with the behavior of the student (Martin & Briggs, 1986). The two reinforcements proposed in the operant conditioning learning theory are positive reinforcement and negative reinforcement. Positive reinforcement gives the student a positive incentive for desirable behavior, while negative reinforcement withdraws a desirable incentive as a result of undesirable behavior. Positive reinforcement is equal to a reward, while negative reinforcement is equal to punishment (Hilgard, 1962; Holt, B. J., & Hannon, J. C., 2006). Thus, giving a student instant feedback with the reinforcement teaching strategy, according to the response of the student, can achieve the learning attendance goal (Anderson, 1981; Saltz et al., 2007).

In applying the reinforcement strategy of operant conditioning in distance education activities, the distance factor prevents teachers from controlling the learning condition of the remote student and giving suitable incentives. Encouraging distance education learners to achieve the expected learning goal involves first detecting fatigued and inattentive behavior of the students participating in class. Besides blinking, yawning, wrinkling caused by yawning, and failing to look straight at the screen as the result of fatigue, inattentive behaviors such as leaving the seat, turning one’s head for chatting or running other application programs can also be detected by image recognition and detection (Eriksson & Papanikotopoulos, 1997; Gu & Ji, 2004; Wang, P., Tran, L. C., & Ji, Q., 2006).

For instance, a human figure can be easily distinguished from the background through the brain’s cognition of human features. Computer vision can adopt similar logic to the brain to obtain a correct facial region. Results of face image binarizing show that the pixels at the corner of the mouth are darker than those in the surrounding regions of the face. This property can be used to identify the possible location of the mouth (Ito, Mita, Kozuka, Nakano, & Yamamoto, 2002; Lee, Park, & Park, 2005; Smith, Shah, & Lobo, 2004). The following physiological responses are signs that a learner is fatigued: eyes closing; failure to look straight ahead; yawning; head turning right and left, and wrinkling between the eyes and above the nose due to yawning (Gu & Ji, 2004). Image recognition can be utilized to obtain some information quickly from facial expressions. For example, the darkest spots found by grayscale manipulation in the possible eye region, and adjusting the appropriate threshold value can be adopted to locate the pupils (Stiefelhagen, Yang, & Waibel, 1996). The eyes can be located from the location of the pupils (Lee et al., 2005; Miyakawa, Takano, & Nakamura, 2004). The eye-blinking variation can be calculated from the part with larger variation by capturing two successive images, then performing gray scale processing and image difference (Grauman et al., 2001).

A Bayesian network is a graphical decision-making tool. Haisong found that the Bayesian network is the best option to deal with such an issue of mental detection (Gu & Ji, 2004). Besides, the Bayesian network could be applied for evaluating or predicting the learning behavior of the distance education students (Xenos, 2004).

These image recognition and detection techniques could be used to recognize fatigued and inattentive behavior in a learner. Bayesian network assessment can reduce detection misjudgment and enhance accuracy. Such a mechanism in distance education could be applied to detect attentiveness of learners. Providing instant feedback in the reinforcement teaching strategy could help a teacher to supervise the students and prevent the students from being distracted, thus achieving the learning attentiveness goal.

**Design and implementation of inattention/fatigue detection system**

In the detection mechanism of this research, the images (15 images/sec.) captured with the webcam were collected for image detection processing. First, the system extracts the skin-color area, and defines the facial feature locations such as the eyes and lips. The system then employs different algorithms to find the feature variations. For instance, consider the features of the eyes. The system needs to decide whether the image is of a closed or open eye. For the mouth feature, the system must determine whether the image is of a stretched mouth. Next, the system must categorize and recognize these image processing results to detect whether the learner is inattentive, fatigued or dozing in class. These features are then assessed with the Bayesian network model. Based on the assessment results,
if the student is determined to have “inattentive behavior” or “feelings of fatigue,” then an alert is transmitted instantly to notify the teacher. The system then captures the real-time image, and delivers it to the server end. This is to record the student’s image and detected condition, so that the teacher can browse and check online.

In this investigation, “inattentive behavior” indicates something irrelevant to the course that a student is doing when not feeling fatigued. “feelings of fatigue” indicates that feeling mentally fatigued in learning owing to physiological exhaustion; characteristic behaviors include drowsiness and raised eye-blinking frequency.

**Figure 1. Detection and recognition procedure**
Detection procedure

Figure 1 shows the entire operation procedure of the detection and reinforcement mechanism. First, the face image is obtained from successive input images. If the width of the face is determined to be larger than the length of the face, then either the region obtained is probably not a human face or part of the background has a similar color to skin. After obtaining the face region, the image is horizontally segmented into equal parts. The eyes’ features can then be obtained from the top half image, and the mouth’s corner features from the bottom half-image. Next, after obtaining the mouth’s corners features, the middle coordinate of the mouth corners are utilized to segment the top half of the face into left and right parts. This is done for two reasons. The first is to shorten the processing time, and thus increase the accuracy of the eye features obtained. The second reason is that, the left eye and right eye can generally be separated into two independent images if the middle coordinate of the mouth can be obtained correctly, and if the vertical axis of this coordinate is adopted to segment the image.

In tracking eyes, the fast movement of the user, or the change in illumination in the searching space, might lead to gradual deviation in the eye template matching. Hence, the movement bound has to be defined. The coordinates of the horizontal or vertical axis between two successive images are compared when the tracing result is obtained. Tracking deviation is defined as a difference exceed in. The system needs to perform the detection again to prevent further feature misjudgments. In tracking the mouth location, the mouth coordinate location of the input image can be discovered merely from the variation of the movement of the eye features between two successive images.

In the successive image time duration, the user’s inattentive/fatigued behavior in every image can be identified, respectively. However, since the actual condition of the user cannot be inferred from only a single image, a second threshold must be defined. The first threshold indicates that the image time is an interval (T1), and the second threshold defines the number of alert events in this interval (T2). In the interval of T1, if the number of alert events exceeds T2, then the system sends an alert message that the learner might be distracted or fatigued. This study defines more than 15 images showing a user distracted or fatigued detected in the time interval of 30 images as evidence of inattentive behavior or feelings of fatigue. These two thresholds are defined because a fatigued person generally yawns for at least two seconds at a time. Because the system captures and detects 10 to 15 images per second, the time for detecting 30 images is about 2 to 3 seconds. If either inattentive behavior or feelings of fatigue is detected, then the system sends the alarm feedback message automatically to the student. Time with no detected inattentive or fatigued behavior is counted. The system sends a positive feedback message to a student automatically if five minutes pass with no inattentive or fatigued behavior.
Bayesian network assessment

A Bayesian network model was constructed in accordance with the variation probability of the learners’ facial expressions and behavior features for assessing whether a learner is with inattentive behavior or feelings of fatigue. The classified inattentive behavior features are defined as the two modes of learner leaving and inattention. “Learner leaving” indicates a student not sitting in front of the computer. “Inattention” might be the result of the learner using other application programs or gazing around, rather than concentrating on learning. For example, the behavior of leaving one’s seat after starting the system, turning one’s head for chatting or watching other things, or running other application programs at the same time, could directly affect the learning quality of a learner. Feelings of fatigue is defined as “Drowsiness” and “Falling Asleep.” The occurrence of drowsiness is defined by one of three occurrences: head nodding, eyes closing and yawning. Falling asleep is defined as eyes closed or head nodding as a result of dozing. Figure 2 illustrates the node causalities of these mental conditions.

Two thresholds for the eye-opening size are defined. If the detected result is less than the first threshold, then the student is detected to be in the state of drowsiness. If the detected result is lower than the second threshold, then the student is detected as having fallen asleep. “Head Turn” is detected according to the frequency of head turning. If a student turns his head or nods with a fixed frequency, he is detected as being in a state of having fallen asleep or in drowsiness. If the student turns his head to the fixed direction, then he is detected to be in a state of inattention.

System interface

Figure 3 depicts the system implementation interface, which includes the source image window, image processing window, image recording window, and the related control items of computer detection. The learner only needs to install and put the webcam above the computer screen and start the system to begin running the core programs. Figure 4 plots the image detection results. The horizontal axis represents the image time, and the vertical axis indicates the probability of inattentive behavior or feelings of fatigue. The system warns the user with sound effects once the number of inattentive or fatigued behaviors exceeds the threshold. Finally, the images and conditions are recorded in the server database for the teacher to check online. The system does not need to transmit all the images of a student to the server for processing. Therefore, problems with network bandwidth and server performance can be
avoided. Image processing is performed by the client computer. The system captures a face image of a student, and sends it to the server for recording only when the student is detected to have inattentive or fatigued behavior.

![Figure 4. Output curves of detection results](image)

**Methodology**

**Particulars**

In this research, sixty sixth-grade students from elementary schools were sampled randomly and divided into two groups, with thirty students in each group. Because of the limited equipment in the experiment environment, only 30 personal computers were simultaneously available for the experiment. A computer classroom was employed as the test environment to simplify the simulation of possible behavior of students in distance learning. This is because the same learning situations experienced at home, such as chatting with other people, being inattentive, or feeling fatigued, are also found in a computer classroom. Additionally, the environment in a computer classroom can easily be changed to improve the detection accuracy. If the home environment is taken as the test environment, then the home light and background must be considered, and help from parents necessary. Although 100 percent accuracy cannot be achieved, adopting the Bayesian network assessment could reduce the incidence of detection misjudgments. To simulate the conditions of the distance education course and avoid the teacher affecting learning by students in the computer classroom, the teacher was only allowed to sit at the front of the classroom during the course. The teacher was separated from the students, and was not permitted to have direct contact with students or correct students’ behavior directly. The teacher was not visible to the students, and so did not affect their behavior.

**Design**

Research assumption: Teachers of distance learning courses who use the proposed detection and reinforcement mechanism show significant differences in guiding and managing the learning attention of students:

A. Independent variable: the proposed detection and reinforcement mechanism for distance education

B. Dependent variable: learning attention of students during the simulative distance education courses, determined by using the proposed detection and reinforcement mechanism

C. Variable control: Because Round 1 was devised to survey inattentive behavior in students, the teacher was asked to sit in front of the class without interfering in class management. Round 2 was designed to survey the feelings of fatigue caused by spending long hours in class. To prevent the students from being affected by incentives of
other external factors, the courses were given in two continuous classes, and the teacher was required to actively manage the class. A possible cause of fatigue in class might be having played with classmates during break time.

**Procedure**

The investigation was split into two parts, with the simulative test performed in the first part and the teaching experiment performed in the second part.

**Table 1. Experiment of round 1 “Inattentive Behavior” process chart**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Time</th>
<th>Teaching activities</th>
<th>System operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st part</td>
<td>0-5 mins.</td>
<td>Webcam adjustment</td>
<td>Simulative detection</td>
</tr>
<tr>
<td>2nd part</td>
<td>5-35 mins.</td>
<td>Teaching experiment course 1-1</td>
<td>Without starting the detection and reinforcement mechanism</td>
</tr>
<tr>
<td>Stage A</td>
<td>5-35 mins.</td>
<td>Teaching topic: The advantages and disadvantages of nuclear power plant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35-40 mins.</td>
<td>Filling in the questionnaire scale for the 1st time in response to Stage A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-5 mins.</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>Stage B</td>
<td>5-35 mins.</td>
<td>Teaching experiment course 2-1</td>
<td>Starting the detection and reinforcement mechanism</td>
</tr>
<tr>
<td></td>
<td>35-40 mins.</td>
<td>Teaching topic: The choice between the resource use and environmental protection</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-5 mins.</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>Stage C</td>
<td>After class</td>
<td>Interviewing ten students by random choice</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Experiment of Round 2 “Feelings of Fatigue” process chart**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Time</th>
<th>Teaching activities</th>
<th>System Operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st part</td>
<td>0-5 mins.</td>
<td>Webcam adjustment</td>
<td>Simulative detection</td>
</tr>
<tr>
<td>2nd part</td>
<td>5-35 mins.</td>
<td>Teaching experiment course 2-1</td>
<td>Without starting the detection and reinforcement mechanism</td>
</tr>
<tr>
<td>Stage A</td>
<td>5-35 mins.</td>
<td>Teaching topic: The advantages and disadvantages of thermal power plant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35-40 mins.</td>
<td>Filling in the questionnaire scale for the 1st time in response to Stage A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-5 mins.</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>Stage B</td>
<td>5-35 mins.</td>
<td>Teaching experiment course 2-2</td>
<td>Starting the detection and reinforcement mechanism</td>
</tr>
<tr>
<td></td>
<td>35-40 mins.</td>
<td>Garbage classification and resource recycle</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-5 mins.</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>Stage C</td>
<td>After class</td>
<td>Interviewing ten students by random choice</td>
<td></td>
</tr>
</tbody>
</table>

In the first part, the simulative test, the students participating in the experiment were asked to imitate the feature behavior of inattentive behavior (e.g., learner leaving, turning head to chat) or feelings of fatigue (drowsiness or
falling asleep) as defined in this research. Actual detection was then performed by the system. Environment adjustments were made to factors such as light and image view to guard against misjudgment in accordance with the detection results and misjudgment probability, thereby lowering the probability of detection misjudgment.

Distance education allows students to learn anytime and anyplace. Thus, students are able to study with classmates, teachers or parents, or on their own. To simulate different situations, the second part was divided into two rounds, each lasting 80 minutes and comprising two class periods. Each round was divided into three stages: A, B, and C. Every student was required to complete Rounds 1 and 2. The detection and reinforcement mechanism were not run in stage A, but were run in stage B. Ten students, five in each group, were interviewed during stage C. Tables 1 and 2 present results of the interviews. Round 1 was designed to survey inattentive behavior among students when they were learning with classmates. In order to simulate an actual distance education situation in which no teacher is around to maintain class order, the teacher was required not to play any part in class discipline. Round 2 aimed at surveying feelings of fatigue and simulated studying in the presence of a teacher in a classroom or around parents at home. The teacher was required to actively maintain class order to prevent the students from interacting with each other and affecting the appearance of fatigue during the actual teaching procedure. Two continuous classes were given in each round.

Tools

The Likert scale is the most common adopted approach for evaluating the learning attention of students (Kubiszyn & Borich, 1996). Therefore, this work performed a survey based on a Likert scale to measure inattention and fatigue among students. Table 3 shows the questions and statements utilized in the survey. The students filled in the questionnaire after completing the two stages of teaching courses. The results of the questionnaires indicated the levels of inattention and fatigue of students in class.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inattentive behavior (Round 1)</td>
<td>1. I was attentive during class.</td>
</tr>
<tr>
<td></td>
<td>2. I have chatted with classmates or left my seat during class.</td>
</tr>
<tr>
<td></td>
<td>3. I have done something irrelevant during class.</td>
</tr>
<tr>
<td>Feelings of fatigue (Round 2)</td>
<td>4. I was energetic in class.</td>
</tr>
<tr>
<td></td>
<td>5. I felt fatigued.</td>
</tr>
<tr>
<td></td>
<td>6. I was dozing during class.</td>
</tr>
</tbody>
</table>

After, the experimental class of ten students was sampled randomly and interviewed with the following questions related to this research:
A-1. Under what circumstances were you inattentive during class?
A-2. Did the teacher’s warning feedback make you more attentive in class?
A-3. Did the teacher’s encouraging feedback make you more attentive in class?
B-1. When did you feel sleepy in class?
B-2. Did the teacher’s warning feedback refresh you in class?
B-3. Did the teacher’s encouraging feedback refresh you in class?

Reliability test

Thirty more students were invited to test the reliability of the scale of this research. A pre-test was administered after these thirty students had finished the course of stage A, part 1. The pre-test was a reliability analysis by Cronbach α. The Cronbach α values were 0.8310 and 0.7790, showing that the results reached a satisfactory reliability standard.
Results and discussion

Results of simulative detection

The first part was simulative detection. The students were required to simulate the following actions: normal learning behavior of looking straight at the computer screen; inattentive behavior of learner leaving or turning head to chat, and feelings of fatigue. Feelings of fatigue included behavior of dozing, such as eyes closing, nodding, and yawning, (eyes narrowing and mouth opening). Table 4 shows the detection results of the sixty students.

Table 4. The result of the simulative detection

<table>
<thead>
<tr>
<th>Category</th>
<th>Simulative behavior</th>
<th>Before adjustment</th>
<th>After adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of students</td>
<td>Correct detection</td>
</tr>
<tr>
<td>Normal learning behavior</td>
<td>Normal learning behavior</td>
<td>60</td>
<td>10</td>
</tr>
<tr>
<td>Inattentive behavior</td>
<td>Learner leaving</td>
<td>60</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Inattention</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(turning head to chat)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feelings of fatigue</td>
<td>Fall asleep</td>
<td>60</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Drowsiness</td>
<td>60</td>
<td>48</td>
</tr>
</tbody>
</table>

The partial misjudgments in the first simulative behavior detection were assumed to be mainly caused by the classroom light, eyeglass reflected rays, and webcam shooting angle. However, some misjudgments were found even after the teacher adjusted for these environmental factors, probably because the interval of the students’ simulative behavior was not long enough to reach the warning threshold defined by the system. Accurate judgment could be obtained after the students performed further specific behavior simulation. The simulative detection results demonstrate the inattention and fatigue of the students could be precisely detected through facial detection after adjusting the mechanism of this research.

After adjustment was made to precisely detect students’ inattentive behavior and feelings of fatigue, the teaching experiment was undertaken with the adjusted mechanism to avoid detection errors in the course experiment.

Teaching experiment and interview results:

Table 5 presents the frequency and paired t-test results for inattention in stages A and B.

Table 5. Table of inattention frequency and t-test, Round 1

<table>
<thead>
<tr>
<th>No.</th>
<th>Stage A</th>
<th>Stage B</th>
<th>t value</th>
<th>significance (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>21</td>
<td>23</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>8</td>
<td>21</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>6</td>
<td>32</td>
<td>18</td>
</tr>
</tbody>
</table>

The experiment of stage A, Round 1 was performed without using the proposed detection reinforcement mechanism. In response to proposition 1, “I was attentive during class,” only 18 percent of the students considered themselves to be attentive, while 43 percent considered themselves inattentive during class. In answer to question 2, “I have chatted with classmates or left my seat during class,” 51.7 percent of the students admitted that they had chatted with
classmates or left their seats during class. Only 13 percent of the students stated that they did not chat with classmates or leave their seats during class. In answer to question 3, “I have done something irrelevant during class,” 36.7 percent of the students admitted that they had done something irrelevant during class, while only 10 percent stated that they had done nothing irrelevant during the class. The questionnaire results reveal that in the distance education, without any detection reinforcement mechanism, most of the students were inattentive and would easily chat, leave their seats, or do something irrelevant.

An in-depth interview was held with ten students on question A-1, “Under what circumstances were you inattentive during class?” All ten students held the same opinion: that in a learning procedure like this, the teacher could not see the condition of the students and therefore would not be able to detect students chatting or leaving their seat for a while. Because being inattentive to the course would not influence the participation grade, the students could easily play truant.

In stage B, the experiment was conducted using the proposed detection reinforcement mechanism. In response to question 1, “I was attentive during class,” the percentage of students who still considered themselves to be inattentive fell to 10 percent, representing a reduction of 33 percent in comparison with the percentage in stage A, while 48.3 percent of the students considered themselves to be attentive during class, a rise of 30 percent in comparison with the percentage in stage A. The comparison results show that the proposed detection reinforcement mechanism increases the attention of the students in learning. In answer to question 2, “I have chatted with classmates or left my seat during class,” only 3 percent of the students admitted chatting or leaving their seats, a reduction of 48.7 percent from the percentage in stage A, while 63 percent considered that they did not do so, an increase of 50 percent in comparison with the percentage in stage A. In response to question 3, “I have done something irrelevant during the class,” only 3 percent of the students still admitted doing other things, a reduction of 33 percent from the percentage in stage A, while 65 percent of the students considered that they did not do so, a rise of 55 percent in comparison with the percentage in stage A. The comparison between stages A and B confirm that the proposed detection reinforcement mechanism discourages students from chatting, leaving their seats, or doing irrelevant activities during learning.

From the in-depth interview on question A-2, “Did the warning feedback of the teacher make you more attentive in class?” nine of the ten students agreed that the warning feedback from the teacher made them more attentive to the course. However, one student still showed no concern for the teacher’s warning. In response to A-3, “Did the encouraging feedback of the teacher make you more attentive in class?” Six students agreed that the encouraging feedback from the teacher helped them to become more attentive in class. The answers to these two interview questions indicate that teachers can increase the learning attention of their students by giving them timely feedback in class.

In distance education, students who are not self-disciplined enough easily exhibit inattentive behavior, thus reducing the learning results of distance education. Consequently, most current distance education programs are designed for adults, undergraduate learners or above, and do not yet replace the at-school curriculum designed for younger students. This is because such distance education focuses only on knowledge learning and provides no effective guidance. The effect of this experiment is particularly significant for elementary school students with little self-discipline. All ten interviewed students agreed, after a further interview, that they should not let themselves become distracted in class, because they would be given warning messages by the teacher if they were inattentive. Comparison results reveal that the proposed mechanism made students feel that the teacher was observing their learning, thus encouraging them to be attentive in class.

Table 6 lists the fatigue frequency and paired t-test results of stages A and B of Round 2.

<table>
<thead>
<tr>
<th>No.</th>
<th>Stage A</th>
<th>Stage B</th>
<th>t-value</th>
<th>significance (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 standard deviation</td>
<td>1 2 3 4 5 standard deviation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3 22 23 10 2</td>
<td>0.91 2 10 32 15 1</td>
<td>0.72</td>
<td>-1.83 0.072</td>
</tr>
<tr>
<td>5</td>
<td>0 4 21 26 9</td>
<td>0.82 0 4 23 25 8</td>
<td>0.80</td>
<td>-1.76 0.083</td>
</tr>
<tr>
<td>6</td>
<td>8 19 26 7 0</td>
<td>0.87 9 23 23 5 0</td>
<td>0.85</td>
<td>-2.05 0.045</td>
</tr>
</tbody>
</table>

Table 6. Table of fatigue frequency and t-test, Round 2
The experiment in stage A was performed without applying the proposed detection reinforcement mechanism. In response to statement 4 in stage A, “I was energetic in class,” twelve students considered themselves to be energetic, while 25 considered themselves to be unenergetic. For statement 5, “I felt fatigued,” 35 students considered themselves to be fatigued, while only four students considered themselves not feeling fatigued at all. In answer to statement 6, “I was dozing during class,” seven students admitted that they had been dozing.

All ten students who participated in an in-depth interview on question B-1, “Under what circumstances did you feel sleepy in class?”, felt that interesting course content would make them less sleepy, while boring course content would make them sleepy. Eight students felt that staying long hours in class would affect their mental states. In particular, looking at the computer screen for a period of time would make the eyes fatigued. Seven students stated that they might not feel sleepy if the class could be given with high levels of interaction between the teacher and students or among the students, e.g., discussion or teamwork.

In stage B, the experiment was undertaken with the proposed detection reinforcement mechanism. In response to the questionnaire statement 4 in stage B, “I was energetic in class,” 16 students considered themselves to be energetic, while 12 students considered themselves unenergetic. Regarding statement 5, “I felt fatigued,” 33 students considered themselves fatigued, while only four students said that they did not feel fatigued at all. In response to statement 6, “I was dozing during class,” five students admitted that they had dozed off in class.

A comparison of data from stages A and B shows that the proposed detection reinforcement mechanism did not significantly decrease the mental fatigue of students during learning. Although this mechanism discouraged the students from being intentionally lazy, it made no significant difference to their feelings. Although two students could force themselves not to doze off in class as a result of the reinforcement mechanism, they still felt fatigued. Thus, the learning result for fatigued students was not good.

All ten students in the in-depth interview on questions B-2, “Did the teacher’s warning feedback refresh you in class?”, and B-3, “Did the teacher’s encouraging feedback refresh you in class?”, admitted that receiving encouraging feedback from the teacher would refresh them during class. However, warning feedback resulting from fatigue had no influence on student behavior. Seven students responded that although they felt like refreshing themselves, they still felt fatigued. In particular, only some of the students who were already dozing could still refresh themselves after receiving the warning from the teacher. The remaining students could not refresh themselves and might have dozed off in class again.

Thus, many factors affect the mental states of students for learning. As well as teaching materials, teaching methods must also be considered to improve the mental state of students, especially in distance education without face-to-face interaction. The external reinforcement mechanism did not improve the physiological responses of fatigue and the risk of dozing was not apparently minimized either. Additionally, the interview indicates that many factors influence the mental state of students. Some factors were associated with the lifestyles of the students, while others resulted from the content of teaching materials and the teaching methods of the teacher. The students still significantly affect their mental states even after receiving the warning feedback from the teacher. Therefore, simply warning a fatigued student does not work. Distance education enables learners to learn anytime and anyplace. The student should be asked to learn at another time in a better mental state. Learning reluctantly is a waste of time. Consequently, because external effects cannot minimize mental fatigue caused by physiological exhaustion, students who feel fatigued should be allowed to learn at another time, rather than being requested to refresh themselves.

Conclusion

Aside from recognizing inattention and fatigue in learners by employing image recognition and detection, this study also focuses on improving the accuracy with the Bayesian network assessment. No technology that combines image recognition with reinforcement mechanisms is yet available. This study utilizes the simplest approach to achieve this objective, a webcam. Although the Bayesian network assessment could not achieve 100 percent accuracy, it did reduce detection misjudgment.

Students might need to learn by distance education in different circumstances, for instance, when classes are suspended for infectious diseases, or when children are absent from school for illness. The proposed mechanism can
help teachers to recognize the learning conditions of students, and prevent students from learning inattentively after logging into the distance learning system. The teacher is able to observe the learning conditions of the students, making the interaction in distance education courses much more like that in at-school classes. Because the proposed mechanism cannot ensure 100 percent accuracy, in this research can only be utilized to help teachers and parents to supervise students, and is not appropriate for use in assessment.

This mechanism was implemented and adopted to perform simulative tests and class experiments. Simulative test results reveal that the accuracy in detecting the feature behavior simulated by students is quite high. Thus, the mechanism can precisely detect learners who are inattentive or fatigued. Class experiment results show that the proposed mechanism can encourage inattentive learners to achieve the learning attention goal. However, mental fatigue caused by physiology can only be solved by requiring students to learn at another time in a better mental state to avoid wasting time. The proposed detection reinforcement mechanism and teaching procedure were designed from the perspective of Affective Domain Teaching Objectives, and can detect accurately the students’ inattentive learning. Teachers of distance education courses thus explicitly controlling the learning condition of students avoiding the common difficulty in distance education of students’ exhibiting inattentive or fatigued behavior. The results of the simulative experiment class demonstrate that the mechanism of this research can detect inattentive behavior in students and encourage students to be attentive to learning.

The proposed mechanism can accurately detect the behavior induced by inattention and fatigue. However, although this mechanism can urge students to be attentive to learning, it cannot refresh students with mental fatigue resulting from physiological exhaustion. Other factors must be studied to gain an in-depth understanding of the impact of teaching strategies and content and presentation of teaching materials on learning by students. Distance education allows learners to learn anytime and anyplace, enabling a teacher to take preventive measures to relieve the learning fatigue induced by exhaustion or allow students to learn at times when they are not tired. Although the proposed mechanism is not mature and its outcome may not exactly meet the teaching or learning needs, it is a helpful starting point for this research issue. Accordingly, future research will perform in-depth studies on factors such as teaching materials and teaching methods. A distance education mode that can achieve the Affective Domain Teaching Objectives with the proposed mechanism in this research could hopefully be developed.

The simulative test was performed in a computer classroom to improve the detection accuracy given the limits of the image recognition and detection method, where the appropriate environmental conditions for image recognition could be easily met. Using the home as the test environment would require help from parents. For example, the home light and the angle of the webcam for the detection system would need to be adjusted according to various environmental conditions. This study aimed to detect inattentive learning behavior and did not classify such behavior specifically. Further detection and recognition of all aspects of student behavior would require more flexible and accurate image recognition and detection techniques or the integrated usage of other technologies.

**Acknowledgements**

The authors would like to thank the National Science Council of the Republic of China, Taiwan for financially supporting this research under Contract No. NSC97-2221-E-324-044. The editorial assistance of Ted Knoy is also appreciated.

**References**


The Instructional Effect of Online Reading Strategies and Learning Styles on Student Academic Achievement

Pei-Hsuan Hsieh
Institute of Information Management, National Cheng Kung University, No. 1 University Road, Tainan 701, Taiwan
// pei.peace@gmail.com

Francis Dwyer
Instructional Systems Program, The Pennsylvania State University, University Park, PA 16802, USA

ABSTRACT
The purpose of this study was to examine the instructional effectiveness of different online reading strategies for students identified as possessing different learning styles, either internal or external locus of control styles, on tests measuring different learning objectives. Participants were 169 undergraduate students, randomly assigned to four online reading treatments: none, rereading strategy, keyword strategy, and question and answer strategy. Immediately after interacting with their respective instructions, students received four individual criterion measures. Analyses indicated an insignificant interaction between learning style and treatment type; however, comprehension tests reflect a significant main effect for students receiving the online rereading treatment ($F = 3.09, df = 3/169, p = .03$), with an effect size of .40. The rereading treatment also appears to be significantly more effective than the control for the comprehension test. Results indicate that not all types of reading strategies are equally effective in facilitating different types of learning objectives. The results indicate that, even though different reading strategies may be structurally different, they are functionally identical for raising questions relative to the cost and amount of time required for student interaction.

Keywords
Reading strategy, Learning style, Locus of control, Learning objectives

Introduction
Many studies have found reading strategies useful when implemented before, during or after reading (e.g., Brown, 2002; Ediger, 2005; Fagan, 2003; McGlinchey & Hixson, 2004; Millis & King, 2001; Sorrell, 1996). For example, reading strategies include rereading, scanning, summarizing, keywords, context clues, question-answer relationships, inferring, thinking aloud, activating prior knowledge, setting a purpose, and drawing conclusions. Online learning environments are becoming popular for most teachers and students. However, few studies focus on appropriate online reading strategies for different types of learners, and most studies focus only on the effectiveness of text-based reading strategies.

The online learning environment has become more and more popular for educators and learners, due to its multiple visual and audio representations. Online learning is a trend that has the potential to enhance learning and increases the importance of knowledge of new teaching methods which apply to new learning environments (Jung, 2001; Romero, Berger, Healy & Aberson, 2000). According to previous studies, some learners encounter difficulties learning online, since they have difficulty changing their learning habits to accept reading electronic texts (Aragon, 2004; Steinhauser & Friederici, 2001). Learners feel doubtful about their learning abilities and believe that they are not as skilled as readers who can overcome the changes in the learning process. Some learners, contrarily, always try their hardest to adapt to the current learning environment, since they believe that they will eventually become comfortable with reading electronic materials (Ehrlich, Kurtz-Costes & Loridant, 1993; Ferguson, 1999; Schommer-Aikins & Easte, 2006). The question remains: What is the best way to employ a text-based reading strategy in an online environment, so that the learners can maintain their confidence for learning?

In addition, students with their own learning styles may process information differently in an online environment (Atkinson & Shiffrin, 1969; Burton, Moore & Holmes, 1995; Huitt, 2003). Different online reading strategies may influence the students’ information processing methods (Driscoll, 2005; Schunk, 2004). This study seeks to discover the proper ways of employing online reading strategies and explores the effects of those online reading strategies and different learning styles on academic achievement. Most studies indicated that students regarded as having an internal locus of control learning style used significantly more reading strategies than students having an external...
locus of control learning style (Cappella & Weinstein, 2001; Fehrenbach, 1991; Maguiness, 1999; Rotter, 1966). In order to consider readers of every learning style, this study focuses on during reading process and considers implementation of different during online reading strategies to complement different information processing. During reading strategies are varied, since more and more visual presentations occur in web learning environments for teachers and students. However, the effects of implementing during reading strategies for different learning styles in an online environment have not been investigated. In addition, not all reading strategies are equally effective methods in facilitating reading comprehension (Baumann et al., 2002). For example, rereading strategies ask readers to read a text more than one time to enhance reading comprehension before readers proceed to the next text. Context clue reading strategies, however, use phonic instructions to help readers gain speed and accuracy in identifying words as they read through a text. Readers may not sequentially comprehend the whole content at the same time as they recognize an unknown word under phonic instructions (Baumann et al., 2002; Ediger, 2005; McGlinchey & Hixson, 2004). Not all during reading strategies are appropriate for an online environment. In this study, through tests measuring different learning objectives, students are required to acquire information relative to the physiology of the human heart, specifically circulation of blood during both the diastolic and systolic phases. Jung (2001) stated that only a few studies have attempted to address the theoretical or conceptual framework of web-based instruction. Not many studies of implemented reading strategies investigated instructional effects on different learning objectives. Therefore, this study contributes its findings to the field of distance education and the futures of e-learners. Prior studies substantiated that varied online reading strategies offer different ways to process information; the students’ learning styles influence information processing methods. Therefore, this study focuses on two variables: online reading strategies and locus of control learning styles. The study methods applied in this study were described afterward with a report of the study results, followed by this study’s contributions and conclusions. Proposed recommendations for future research appear as the final comment.

Reading strategies and learning styles

This section reviews and defines reading strategies and learning styles, and is followed by discussion of theoretical justification for exploring learning effects of locus of control and reading strategies on achievement. A literature review of three online reading strategies and learning styles is also included in this section. Finally, a brief summary ends this section.

Online reading strategies incorporate three reading strategies in an online environment. This study selects three online reading strategies: rereading strategy, keyword strategy, and question and answer (QA) strategy. The rereading strategy is a useful pedagogical tool and has potential benefit for enhancing readers’ comprehension as well as enjoyment of literature (Faust & Glenzer, 2000). Faust and Glenzer concluded that the rereading strategy helps students obtain meaning of their favorite reading sections and makes meaning with texts. However, Short, Kane, and Peeling (2000) found that rereading a longer text may be time consuming, although using cues and rereading students’ favorite sentences can help students familiarize themselves with the texts. Millis and King (2001) conducted research with 42 undergraduate psychology students who were recognized as good readers and found that rereading strategically improved their comprehension and retention of ill-structured information. Brown (2002) also found that female Japanese college students’ reading comprehension improved through the use of rereading strategies.

The keyword strategy has been found, through research, to be useful in improving students’ ability to comprehend. De Courcy and Birch (1993) conducted research through open-ended interviews, observation, and think-aloud protocol with four Japanese students and found that the students mainly used keywords and inference as their reading strategies to comprehend the whole text. Fagan (2003) found that English as second language learners need the keyword strategy as a scaffold during the reading process. Some researchers found significant effects on reading comprehension, such as employing keyword strategies before reading. For example, O’Donnell, Weber, and McLaughlin (2003) found that students obtained high comprehension scores when they previewed materials and discussed keywords before reading.

The question and answer strategy, which is usually called question-answer relationship (QAR), can increase readers’ metacognition awareness (Benito et al., 1993; McIntosh & Draper, 1995; Raphael, 1982). Most research indicated
significant comprehension effects on reading while implementing this type of reading strategy. For example, Raphael (1982) reported that students were able to locate the information from the text and could properly respond to questions. Benito et al. (1993) also found that students could comprehend three different types of questions after employing QAR strategy. McIntosh and Draper (1995) found that QAR strategy helped students read, answer questions, and learn from texts.

Locus of control, a learning style variable, is commonly thought to be the degree to which an individual perceives personal responsibility for a specific outcome. Individuals possessing high levels of responsibility are generally labeled “internal control” while individuals processing low levels of responsibility are labeled “external control” (Lefcourt, 1982). In situations that involve skills, internals generally spend more time on decision making than do externals. For tasks that require less skill-demand and are more by chance, internals exhibited carelessness and impulsiveness in their responses (Johnson & Kilmann, 1975; Kukulu, Buldukoglu, Kuladaç & Köksal, 2006; Wheeler & Davis, 1978). Data from a number of studies indicated that readers regarded as having external or internal locus of control learning styles need to apply certain reading strategies to comprehend a text (Cappella & Weinstein, 2001; Coldevin, Tovar, & Brauer, 1993; Coker, Coggins, & Flaitz, 1989; Maguiness, 1999; Whitney, 1992). For example, Coker et al. (1989) reported that internal locus of control learning style readers have better performance in reading and, as good readers, they are able to use many reading strategies during the reading process. This enhances internals’ understanding of the meaning of unknown words, as compared to poorer readers (Arden-Close, 1993). Contrarily, poor readers, regarded as using external locus of control learning styles, might need support before, during, and after their reading process (Coker et al., 1989; Coldevin et al., 1993).

Rotter’s Social Inventory Scale (Rotter, 1966) separates individuals into two perceptual groups based on individuals’ perception of personal, causal role in the outcome of specific events, either internal or external. Internals are individuals who perceive personal responsibility for success or failure, whereas externals attribute luck, chance, or another factor for the reason for success or failure (Jonassen & Grabowski, 1993; Lefcourt, 1982). Prior research indicated differences in situations which facilitate increased achievement among internals and externals. Externals have been shown to perform best in situations of external reinforcement and a controlled instructional environment (Daniels & Stephens, 1976; Wesley, Krockover, & Devito, 1985). Internals interact intensely with whatever instructional treatment they receive, and this interaction results in maximum information acquisition and subsequent performance (Buckley & Dwyer, 1987). Internals score consistently higher than externals on uncued instruction (Brooks & McKelvie, 1986; Bursik & Martin, 2006). Externals prefer teacher-controlled method in a traditional educational setting as opposed to independent instruction, and have higher achievement scores in that preferred setting (Daniels & Stephens, 1976).

A commonly accepted view of internal and external locus of control is that locus of control, as a learning style variable, is the degree with which an individual perceives personal responsibility for a specific outcome. That is, the expectancy of being personally responsible for achievements is the measurement of the degree to which a person feels responsible for the outcome of a behavior. The students may be dichotomized into two learning styles (internal or external locus of control); each style has specific characteristics which influence the potential for profiting from instructional structures. Individuals possessing high levels of responsibility are generally labeled as internal locus of control, while individuals possessing low-levels of responsibility are labeled external locus of control (Lefcourt, 1982). When in a situation that involves skill, internals generally spend more time on decision making when compared to externals. For tasks that require less skill and are more chance-based, internals exhibited carelessness and impulsiveness in their responses (Johnson & Kilmann, 1975; Wheeler & Davis, 1979). Contrarily, poor readers, regarded as having an external locus of control learning style, might need all supports before, during, and after their reading processes (Coker et al., 1989; Coldevin et al., 1993). Those students believe their behaviors to be unaffected by rewards and are generally hard to control. Therefore, they need support from their teachers and even doubt their learning abilities. Maguiness (1999) identified some external locus of control students still needed support, even though they participated in a joint reading problem program at Westerns Springs College in Auckland, New Zealand.

The literature indicates that different types of reading strategies may vary in effectiveness in facilitating student achievement of different types of learning objectives and that students may be dichotomized into levels of learning styles (locus of control) each having specific characteristics which differentially influence the potential to profit from instructional structures. Specifically, the purpose of this study is to examine: a) the instructional effectiveness of different types of reading strategies in facilitating student achievement of different types of learning objectives, b) the effect of different types of reading strategies on students, identified as either internal or external locus of control,
by tests measuring different learning objectives, and c) whether an interaction exists between type of reading strategy and learning style.

In sum, this study explores the relationship between locus of control and reading strategy in an online environment. By reviewing information processing theory, three reading strategies presenting different effects on students’ comprehension are clearly presented. Most researchers have also confirmed that three online reading strategies have significant effects on comprehension. The theoretical justification section presents a number of prior studies which examine the relationship between learning styles and online reading strategies. The next section describes the study methods applied in this study.

Study methods

This section includes three sub-sections: Participants, Instrumentations and Research Design. Beginning in the spring of 2006, the study used an open-source programming language to recruit participants. Detailed information about subject variables appear in the Participants section. The Instrumentations section details learning materials, criteria of achievement measures, and locus of control measures. The Research Design section describes how the study was conducted in the preparation, implementation, and analysis stages.

Participants

Participants were 169 undergraduates from English, statistics, and education psychology programs, recruited at the Pennsylvania State University. They were all college-level undergraduates in their first or second university year. Most of them from the English program were freshmen and pre-service teachers. Although most participants from English and education psychology programs were female, the locus of control measures used in this study to determine learning styles and to randomly assign the students to different experimental group assignments were not gender sensitive. All participants were reminded to read the participation directions before starting the study. Because of their instructors’ high level of support for the study, participants were motivated to participate. All of them signed a consent form ensuring that they were over 18 years old. They received an incentive to participate, which consisted of obtaining one or three credit point(s) for their final course scores, based on their instructors’ grading policies. None of the participants had a medical education background nor did any participant have known psychological, personal, or academic risks that could influence this study.

Instrumentation

Learning materials

This study used a 2,000-word instructional module focusing on the physiology of the human heart and the related processes occurring during systolic and diastolic phases (Dwyer, 1972). The original instructional content was converted to a 20-page, web-based online instructional module. A description of the experimental treatments is:

- Treatment 1 (control group): Students in the control treatment interacted with the online instructional content in the conventional manner (Figure 1). Students were told that they would be tested on the content. They progressed through the content at their own rate and could move forward and backward depending on need.
- Treatment 2 (rereading group): In the rereading treatment, students received the same information more than one time; however, in this treatment, selected sentences related to specific learning objectives and test questions were repeated (Figures 2 & 3). When students finished reading the first web page and then clicked on a “Next” button, the selected sentences or paragraphs from the first web page showed on the second web page word by word in teleprompter fashion until the “CONTINUE” button appeared. The sentence to be reread was not presented all at once; words appeared sequentially and created the sense of animation. After the rereading was complete, participants then clicked on the “CONTINUE” button to proceed to the rest of the web pages. They could have also clicked on the “Previous” button to go back to the first web page. The participants in this online instructional set were forced to view the same sentences or paragraphs twice. In this study, the rereading strategy was added to the original 20 web pages.
Treatment 3 (keyword group): Students receiving this treatment received bolded keywords related to key learning objectives as well as test questions (Figures 4 & 5). Some terminologies referred to the names of the parts of human heart. The keywords were bolded and in larger font than the rest of the text so that they would stand out. In this study, 31 facts, concepts, and procedures were emphasized in this way.

Treatment 4 (question and Answer [QA] group): This treatment comprised 29 questions and answers (Figure 6). The questions’ design focused students’ attention on specific learning objectives and related criterion measures.

Figure 1. Screen shot of the first treatment: control group

Figure 2. Screen shot of the second treatment: rereading

Figure 3. Screen shot of important sentences for rereading
The Human Heart

The heart lies toward the front of the body and is in a slanting position between the lungs, immediately below the breastbone. The wide end points toward the right shoulder. The small end of the heart points downward to the front of the chest and toward the left. The lower portion of the heart is called the apex and is the part that you feel beating.

The human heart is really two pumps combined in a single organ which circulates blood to all parts of the body. The heart is divided longitudinally into two halves by the septum. The two halves may be compared to a block of two houses, which are independent of each other but have a common wall, the septum, between them.

Figure 4. Screen shot of the third treatment: keyword (I)

The Human Heart

Each half of the heart is divided into an upper chamber and a lower chamber. The upper chambers on each side of the septum are called auricles; the lower chambers are called ventricles. Auricles have thin walls and act as receiving rooms for the blood, while the ventricles having thicker walls act as pumps moving the blood away from the heart. Although there is no direct communication between the right and left sides, both sides function simultaneously.

Figure 5. Screen shot of the third treatment: keyword (II)

The Human Heart

Q: Where is the "Apex" located on the heart?
A: The small end of the heart points downward to the front of the chest and toward the left. The lower portion of the heart is called the apex and is the part that you feel beating. The human heart is really two pumps combined in a single organ which circulates blood to all parts of the body.

Q: Where is the "Septum" located on the heart?
A: The heart is divided longitudinally into two halves by the septum. The two halves may be compared to a block of two houses, which are independent of each other but have a common wall, the septum, between them.

Figure 6. Screen shot of the fourth treatment: Q & A
Criteria of achievement measures

The criterion measures correspond to the different knowledge acquisition domains: facts, concepts, rules/principles, and procedures. Each criteria measure contains 20 test items.

- Drawing test (Kuder-Richardson Formula 20 [KR-20] = .84): The objective of the drawing test was to evaluate student ability to construct and/or reproduce items in their appropriate contexts. The drawing test provided the students with a numbered list of terms corresponding to the parts of the heart discussed in the instructional presentation. The students were to draw a representative diagram of the heart and place the numbers of the listed parts in their respective positions. For this test the emphasis was on the correct positioning of the verbal symbols with respect to one another and with respect to their concrete referents.

- Identification test (KR-20 = .80): The objective of the identification test was to evaluate student ability to identify parts or positions of an object. This multiple-choice test required students to identify the numbered parts on a detailed drawing of a heart. Each part of the heart, which had been discussed in the presentation, had a number on a drawing. The objective of this test was to measure the ability of the student to use visual cues to discriminate one structure of the heart from another, and to associate specific parts of the heart with appropriate names.

- Terminology test (KR-20 = .75): This test consisted of items designed to measure knowledge of specific facts, terms, and definitions. The objectives measured by this type of test are appropriate to all content areas which have an understanding of the basic elements as a prerequisite to the learning of concepts, rules/principles, and procedures.

- Comprehension test (KR-20 = .74): Given the location of certain parts of the heart at a particular moment of its functioning, students had to determine the position of other specified parts or positions of other specified parts of the heart at the same time. This test required that the students have a thorough understanding of the heart, its parts, its internal functioning, and the simultaneous processes occurring during the systolic and diastolic phases. The comprehension test design measured a type of understanding in which the individual can use the information being received to explain some other phenomena.

Locus of control measures

Rotter’s internal-external locus of control scale (1966) uses 23 forced-choice items to measure personal belief. The participants can be divided into internal and external locus of control learning styles. The higher the score, the more external a person measures (Lefcourt, 1982). The participants, identified with one of the two styles, enter a randomly assigned process and receive an assignment to one of the four treatment groups. This study helps explain how likely the participants are to follow their learning styles when reacting to the reading strategies provided in the online learning environment. The reliability of Rotter’s internal-external locus of control scale has a range of .6 to .9, depending upon the population (Jonassen & Grabowski, 1993; Lefcourt, 1982). This study’s reliability rated .63 (KR-20).

Research design

According to the study purpose, a two-way multivariate analysis of variance (MANOVA) analyzed the null hypothesis in the research design. The null hypothesis has two independent variables (reading strategy and learning style) and four dependent variables (drawing, identification, terminology, and comprehension learning objective tests). Some extraneous variables, such as gender and education background, had no significant effects on the study results. Overall, this study contains three stages: preparation, implementation, and analysis. In the preparation stage (summer and fall of 2005), we conducted a pilot study, which was purely experimental; its purpose was to identify response difficulty for all test items. In the spring of 2006, after revising the learning materials and adding different online reading strategies, we conducted the actual study and analyzed the study results.

Preparation stage

A pilot study, conducted with 27 students, was subjected to an item analysis to identify response difficulty for all test items (with item difficulty of less than .6). This information identified students who were having difficulty. The
experimental reading strategies were positioned to facilitate the acquisition of the information necessary to perform the specific criteria measures. The test items (80 test items in total) with less than .6 correct-response rates had been given one or more frame numbers identifying where the instructions occur. Overall, 16 out of 20 frames referred to 59 difficult items needing instructional enhancement. Revision of the learning materials resulted in the addition of different online reading strategies for the next stage of the study.

**Implementation stage**

In the actual study, students received a random assignment to treatment groups, and after interacting with their respective treatments, completed the four learning criteria measures. Since larger numbers better represent a normal distribution of learning outcomes, the study used an open-source programming language to create dynamic hypertext preprocessor (PHP) web pages to recruit participants. PHP made it easier to recruit students online. Finally, 169 participants were invited to take part in the study. All participants completed the study within about 70 minutes.

Figure 7 is the concept map for the implementation stage of the study procedure. The letter C represents the control group. The three online reading strategy groups are: rereading (R), keyword (K), and question and answer (QA). The two locus of control learning styles are internal and external (I and E, respectively).

**Analysis stage**

Descriptive analysis occurred first, followed by correlational analysis to determine the relationship among the four criteria measures of achievement (dependent measures) and to verify the appropriateness of conducting MANOVA. MANOVA examines the main and the interaction effects of one or more categorical independent variables which are predictors for multiple interval dependent variables (Ramsey & Schafer, 2002). For this study, the categorical independent variables are the reading strategies, and the multiple interval-dependent variables are the criterion measures of achievement scores. If any significant differences of a dependent measure yielded in favor of a certain online reading strategy over the control, a follow-up analysis, such as Tukey’s honest significant difference (HSD) tests, is available for verification. HSD provides boundary on differences between independent and dependent group averages (Ramsey & Schafer, 2002). Other pair-wise comparison methods, such Fisher’s protected least significant difference (LSD) and the Bonferroni correction, are not appropriate, since the former cannot control experiment-wide error, and the latter can, conservatively, produce a very small coefficient (Ramsey & Schafer, 2002).

**Study results**

This section presents the results of the analyses of the study. The reliability test results reported in the previous section verify the proper test-item designs for the four criteria of achievement measures (i.e., drawing, identification, terminology, and comprehension tests) and locus of control measures. The descriptive statistics are the first analysis
results. A correlational analysis determines the validity of the measurements as well as the relationships among the dependent variables for verifying the appropriateness of conducting MANOVA. Finally, the results of a 2 x 4 MANOVA and an additional 2 x 1 ANOVA end this section.

Descriptive analysis

Descriptive analysis was the first investigation for this study, in which the control group had 43 participants, the rereading group had 42 participants, the keyword group had 41 participants, and the QA group had 43 participants. Overall, all 169 participants obtained, on average, 11.46 points (SD = 3.50) in the locus of control measurement. Therefore, the two categories became: internal (obtaining more than 12 points) and external (obtaining less than 12 points). In the control group, 24 participants were of the internal locus of control type (external N = 19). In the rereading group, 17 participants were the internal locus of control type (external N = 25). In the keyword group, 20 participants were the internal locus of control type (external N = 21). In the QA group, 25 participants were the internal locus of control type (external N = 18). Overall 86 participants were the internal locus of control type, and 83 were the external type (see Table 1). The former, on average, obtained 30.33 points (SD = 14.41) in the composite test. The external locus of control type of student (83) obtained 33.82 points (SD = 16.83).

<table>
<thead>
<tr>
<th>Treatment groups</th>
<th>Numbers of internals</th>
<th>Numbers of externals</th>
<th>Numbers of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: Control</td>
<td>24</td>
<td>19</td>
<td>43</td>
</tr>
<tr>
<td>T2: Rereading</td>
<td>17</td>
<td>25</td>
<td>42</td>
</tr>
<tr>
<td>T3: Keyword</td>
<td>20</td>
<td>21</td>
<td>41</td>
</tr>
<tr>
<td>T4: QA</td>
<td>25</td>
<td>18</td>
<td>43</td>
</tr>
<tr>
<td>Total Numbers</td>
<td>86</td>
<td>83</td>
<td>169</td>
</tr>
<tr>
<td>Total Composite</td>
<td>(M/SD)</td>
<td>(M/SD)</td>
<td>(M/SD)</td>
</tr>
<tr>
<td>Scores</td>
<td>(30.33/14.41)</td>
<td>(33.82/16.83)</td>
<td>(32.04/13.75)</td>
</tr>
</tbody>
</table>

In addition, all participants, on average, obtained 5.98 points (SD = 4.17) in the drawing test; 8.83 points (SD = 3.97) in the terminology test; 9.00 points (SD = 4.30) in the identification test, and 8.23 points (SD = 3.98) in the comprehension test (see Table 2). All participants had an average of 32.04 points (SD = 13.75) for all four tests. Of all participants, students in the rereading strategy groups performed well in every test and equally well with composite test scores (M = 36.26, SD = 15.13). The students assigned in the control group received, on average, 32.02 points (SD = 13.80). In the keyword group, students averaged 31.66 points (SD = 14.36). The QA group performed the least well (M = 28.30, SD = 10.66).

<table>
<thead>
<tr>
<th>Treatments</th>
<th>Drawing (M/SD)</th>
<th>Identification (M/SD)</th>
<th>Terminology (M/SD)</th>
<th>Comprehension (M/SD)</th>
<th>Total composite scores (M/SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: Control</td>
<td>5.81/4.24</td>
<td>9.09/3.59</td>
<td>9.19/3.93</td>
<td>7.93/4.49</td>
<td>32.02/13.80</td>
</tr>
<tr>
<td>N = 43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2: Rereading</td>
<td>7.02/4.61</td>
<td>9.71/4.50</td>
<td>9.79/4.56</td>
<td>9.74/3.88</td>
<td>36.26/15.13</td>
</tr>
<tr>
<td>N = 42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4: Q&amp;A</td>
<td>5.35/3.69</td>
<td>8.02/3.32</td>
<td>7.77/3.79</td>
<td>7.16/3.23</td>
<td>28.30/10.66</td>
</tr>
<tr>
<td>N = 43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5.98/4.17</td>
<td>8.83/3.97</td>
<td>9.00/4.30</td>
<td>8.23/3.97</td>
<td>32.04/13.75</td>
</tr>
</tbody>
</table>

44
Correlational analysis

Correlational analysis establishes the validity of the measurements (Cavana, Delahave, & Sekaran, 2001), since it determines the relationship among the dependent variables. To verify the appropriateness of conducting MANOVA, correlational analysis is necessary. Glass and Hopkins (1996) mentioned that if a correlation coefficient value is less than .3, the dependent variables are less correlated with each other (a weak or low relationship); however, if the value is between .3 and .7, the relationship is moderate (French & Poulsen, 2002). With the same rationale, a strong or high correlation has a value of more than .7. Table 3 illustrates the correlation coefficients among the four tests achieved at the .01 level of significance. The relationship among different measures is moderate and even strong, providing justification for using MANOVA in this study.

### Table 3. Pearson correlations among different measures

<table>
<thead>
<tr>
<th></th>
<th>Drawing</th>
<th>Identification</th>
<th>Terminology</th>
<th>Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drawing</td>
<td>1</td>
<td>.643**</td>
<td>.523**</td>
<td>.487**</td>
</tr>
<tr>
<td>Identification</td>
<td>1</td>
<td>1.667**</td>
<td>.606**</td>
<td></td>
</tr>
<tr>
<td>Terminology</td>
<td>1</td>
<td>1.685**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td></td>
<td></td>
<td></td>
<td>** p &lt; 0.001**</td>
</tr>
</tbody>
</table>

** MANOVA

Consequently, a 2 x 4 MANOVA table (see Table 4) indicates that no significant interaction exists between reading strategy and level of locus of control learning styles ($F = 1.31$, df = 3/169, $p = .27$). A significant main effect ($F = 3.09$, df = 3/169, $p = .03$), with an effect size of .40, exists on the comprehension test; however, insignificant differences exist on the drawing test ($F = 1.37$, df = 3/169, $p = .30$), identification test ($F = 1.69$, df = 3/169, $p = .17$), and terminology test ($F = 1.76$, df = 3/169, $p = .16$) criterion measures.

### Table 4. Reading strategies on learning objective tests

<table>
<thead>
<tr>
<th>Source (Groups)</th>
<th>Type III sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
<th>Mean</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drawing</td>
<td>69.052</td>
<td>3</td>
<td>23.017</td>
<td>1.368</td>
<td>.255</td>
<td>6.041</td>
<td>.319</td>
</tr>
<tr>
<td>Identification</td>
<td>81.527</td>
<td>3</td>
<td>30.844</td>
<td>1.686</td>
<td>.172</td>
<td>8.888</td>
<td>.306</td>
</tr>
<tr>
<td>Terminology</td>
<td>92.533</td>
<td>3</td>
<td>27.176</td>
<td>1.755</td>
<td>.158</td>
<td>9.037</td>
<td>.332</td>
</tr>
<tr>
<td>Comprehension</td>
<td>142.943</td>
<td>3</td>
<td>47.648</td>
<td>3.091</td>
<td>.029*</td>
<td>8.272</td>
<td>.305</td>
</tr>
</tbody>
</table>

*p < 0.05

Table 5 is a follow-up analysis (LSD), conducted on the comprehension test, which yielded a significant difference in favor of T2, the rereading treatment, over T1, the control ($T = 1.99$, df = 2/83, $p = .05$). A significant difference also appears between T2, rereading, and T4, QA treatment, in favor of T2 ($T = 3.33$, df = 2/83, $p = .001$). Analyses also indicate an insignificant interaction between reading strategy and learning style (level of locus of control) on each criterion measure.

### Table 5. LSD analysis on comprehension test

<table>
<thead>
<tr>
<th>Source (Groups)</th>
<th>t-value</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>Mean difference</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rereading (M = 9.74, SD = 3.88) &amp; Control (M = 7.93, SD = 4.49)</td>
<td>1.985</td>
<td>83</td>
<td>.05*</td>
<td>1.81</td>
<td>.85</td>
</tr>
<tr>
<td>Rereading &amp; Keyword (M = 8.12, SD = 3.87)</td>
<td>1.898</td>
<td>81</td>
<td>.061</td>
<td>1.62</td>
<td>.86</td>
</tr>
<tr>
<td>Rereading &amp; QA (M=7.16, SD=3.23)</td>
<td>3.328</td>
<td>83</td>
<td>.001**</td>
<td>2.58</td>
<td>.85</td>
</tr>
</tbody>
</table>

*p < 0.05 **p < 0.001
A 2 x 1 ANOVA analyzed the effect of reading strategy and student learning style for the composite test scores. No interaction appears between the reading strategy and the student learning style. However, the main effect of the reading strategy treatments had an approaching significance ($F = 2.541, p = .058$, see Table 6) on the composite test scores. The student learning style (or the locus of control type) still had no main effect on the composite test scores.

Table 6. Reading strategies on learning objective tests

<table>
<thead>
<tr>
<th>Source (Groups)</th>
<th>Type III sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>$F$</th>
<th>Sig.</th>
<th>Mean</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected model</td>
<td>2407.389</td>
<td>7</td>
<td>343.913</td>
<td>1.887</td>
<td>.075</td>
<td>32.520</td>
<td>2.073</td>
</tr>
<tr>
<td>Intercept</td>
<td>172187.316</td>
<td>1</td>
<td>172187.316</td>
<td>944.881</td>
<td>.000</td>
<td>36.493</td>
<td>2.122</td>
</tr>
<tr>
<td>Group</td>
<td>1389.335</td>
<td>3</td>
<td>463.112</td>
<td>2.541</td>
<td>.058</td>
<td>31.602</td>
<td>2.109</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32.520</td>
<td>2.073</td>
</tr>
<tr>
<td>Rereading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>36.493</td>
<td>2.122</td>
</tr>
<tr>
<td>Keyword</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>31.602</td>
<td>2.109</td>
</tr>
<tr>
<td>QA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.338</td>
<td>2.086</td>
</tr>
<tr>
<td>LCTP</td>
<td>322.058</td>
<td>1</td>
<td>322.058</td>
<td>1.767</td>
<td>.186</td>
<td>33.632</td>
<td>1.493</td>
</tr>
<tr>
<td>External</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33.632</td>
<td>1.493</td>
</tr>
<tr>
<td>Internal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30.844</td>
<td>1.473</td>
</tr>
<tr>
<td>Group * LCTP</td>
<td>717.484</td>
<td>3</td>
<td>239.161</td>
<td>1.312</td>
<td>.272</td>
<td>36.789</td>
<td>3.097</td>
</tr>
<tr>
<td>Control — external</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>36.789</td>
<td>3.097</td>
</tr>
<tr>
<td>Control — internal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.250</td>
<td>2.756</td>
</tr>
<tr>
<td>Rereading — external</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35.280</td>
<td>2.700</td>
</tr>
<tr>
<td>Rereading — internal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37.706</td>
<td>2.734</td>
</tr>
<tr>
<td>Keyword — external</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33.905</td>
<td>2.946</td>
</tr>
<tr>
<td>Keyword — internal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29.300</td>
<td>3.019</td>
</tr>
<tr>
<td>QA — external</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.556</td>
<td>3.182</td>
</tr>
<tr>
<td>QA — internal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.120</td>
<td>2.700</td>
</tr>
<tr>
<td>Error</td>
<td>29339.321</td>
<td>161</td>
<td>182.232</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>205251.000</td>
<td>169</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>31746.710</td>
<td>168</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

This experimental study shows evidence that reading strategies can be applied in web-based or online learning environments to support students with different learning styles for processing information, such as comprehending a text. Specifically, this study explores the effects of three during online reading strategies (rereading, keyword, and QA) and two kinds of locus of control learning styles (internal and external) on students’ academic achievement. Different data analysis methods (reliability, descriptive statistics, correlational analysis and MANOVA) employed in this study support information processing theory that the students with different locus of control learning styles process information differently. The following explains the rationale for the conclusions according to the study methods, analyses, and study results presented in the previous sections.

First, this study examines the interaction between locus of control types of students and online reading strategies. The assumption is that a certain type of locus of control students obtains more benefits from using a certain type of during reading strategy for learning online content. However, no interaction is discernable between the locus of control types and the reading strategies in the four learning objective tests. The students, regarded as internal locus of control type, did not perform significantly well on each test as compared to the external locus of control type of students’ test performance. Encouragingly, those internals assigned to the rereading strategy obtained higher scores for the four learning objective tests than did the students regarded as external locus of control type. Also uncovered is that the students assigned to the rereading strategy obtained significantly higher scores than did others, regardless
of their status as internal or external. Previous studies also found that rereading is a useful pedagogical strategy (Olmstocks, 1999; Nathan & Stanovich, 1991) and yields the most significant effect on student achievement since it helps increase students’ reading fluency and creates a critical connection with reading comprehension. Faust and Glenzer’s (2000) study concluded that the rereading strategy helps students intensify their impression of favorite reading sections and engenders a deeper connection with texts. The study results again upheld an assumption about rereading which allows individuals to employ it differently for comprehending a text. Faust and Glenzer (2000) and Millis and King (2001) conducted two experimental studies with undergraduate psychology students who were asked to read short passages twice from computer screens. Both study results showed a significant main effect on readers’ memory scores, especially when they, to some extent, strategically reread texts.

Conclusions and implications

Overall, the conclusion of this study is that different reading strategies have different instructional structures and functions in facilitating student achievement of different types of learning objectives. Rereading strategy, implemented on the web in teleprompter fashion, resulted in students having greatly improved comprehension. In addition, practical implications of implementing rereading strategy are like using an online supervisory system in counseling centers. For example, Klitzke and Lombardo (1991) suggested using “bug-in-the-eye” (BITE) technology to assist clinical training. BITE is computer-assisted supervision with a one-way communication process. It allows supervisors to effectively provide prompt, on-screen, with a few words, during counseling sessions for their trainees (Miller, Miller, & Evans, 2002). In the past, the supervisors used telephone call-in or knock-on-the-door approaches, which easily interrupted the counseling process.

Today, increased computer-assisted supervision has proven effectiveness for providing feedback to trainees. Smith, Mead, and Kinsella (1998) reviewed several live supervision techniques and found that direct supervision with computer monitors is the most effective method. A prompt shown on the computer monitor can include words and/or visual icons. Scherl and Haley’s (2000) clinical notes also had positive comments for placing two 14-inch color computer monitors separately, one in the therapy room and the other in the supervisor’s room. They conducted a qualitative investigation with six master’s-level students during a 10-week practicum training. The students’ clients knew of the purpose of using the monitors and eight live-supervision sessions were videotaped. The students found computer-assisted supervision less disruptive. They received, on average, six to seventeen messages in the one-hour sessions. In addition, only one client with very wide peripheral vision indicated awareness of the prompts. Computer-assisted reading strategies may be best during instructional support for students. This study encourages investigation into ways of strengthening online, rereading strategy effects on students’ learning comprehension. For example, animated effects can be incorporated into online rereading strategy utilized by language learners. Students with different learning styles also need further consideration. Investigating their other learning styles or reading habits, conducting interviews, and distributing questionnaires are ways to obtain participatory information as well as assistance for interpreting a study’s statistical results.

Three limitations of this study exist. College-level undergraduates perceive completing the study to be more important than trying their best to perform well in the study. Hence, the participants, having an internal locus of control learning style, can demonstrate the same performance as those with an external locus of control. The amount of time required for reading should be less than 30 minutes to prevent the participants assigned to the control, keyword, and QA groups from reading the materials more than twice. Some factors may also influence study results that need further investigation, such as drawing skills and ability to view texts on a computer screen.

Finally, future research should include online reading strategies for before and after student learning processes. For example, the teleprompter-type online rereading strategy can effectively enhance students’ comprehension during their learning processes. Some databases, which are full-text web pages, might need other types of online reading strategies before students download too many resources to read. In addition, even though the study setting is the same, a new finding results from dual- or multi-reading strategies and learning styles on reading. Perhaps varied measurements can help make appropriate instructional decisions for both study and instructional designs. In sum, educators can encourage students to be aware of their own strengths in varied learning situations. A continued focus on exploiting varied ways of using online reading strategies for different e-learners will enhance reading performance results.
References


MEAT: An Authoring Tool for Generating Adaptable Learning Resources

Yen-Hung Kuo and Yueh-Min Huang
Department of Engineering Science, National Cheng Kung University, No.1, University Road, Tainan City, Taiwan
// Tel: +886-6-2757575 ext. 63336 // Fax: +886-6-2766549 // keh@easylearn.org // huang@mail.ncku.edu.tw

ABSTRACT
Mobile learning (m-learning) is a new trend in the e-learning field. The learning services in m-learning environments are supported by fundamental functions, especially the content and assessment services, which need an authoring tool to rapidly generate adaptable learning resources. To fulfill the imperious demand, this study proposes an authoring tool named Mobile E-learning Authoring Tool (MEAT) to produce adaptable learning contents and test items. In addition, the visualized course organization tool has also been provided to teachers to organize their teaching courses. All functionalities of the MEAT are designed according to the teachers’ feedback and their technological learning needs. To evaluate the MEAT, we have conducted an extensive comparison between the MEAT and other (adaptation) content authoring tools. The result indicates the MEAT is the only tool, which can produce adaptable contents and test items while supporting learning standard. In addition to technical comparison, the qualitative feedbacks from teachers and students are also shown in the evaluation section, and result points out the advantages and shortcomings of the MEAT. According to our findings, we have summarized some design principles for readers who are interested in designing e-learning applications. It hopes our precious experiences can inspire readers to develop more valuable learning tools.

Keywords
Adaptable learning materials, Adaptable test items, Authoring tool, Learning map design, Mobile learning

Introduction
Owing to the proliferation of e-learning, the content production becomes a great deal, which aims at satisfying the learning demands of distance learners. In addition, with advances of mobile technology, the learning demands may also come from mobile devices, and that increases the complexity and difficulty of the content production. To deal with the diversified demands, an authoring tool is needed by content creators to produce adaptable learning resources. To fulfill the imperious demand, an authoring tool named Mobile E-learning Authoring Tool (MEAT) was conceived in this work; for it contains adaptable content and test item generation abilities, user-friendly learning design functions, and intuitional GUIs. Except for advantages hereinbefore, the produced contents of MEAT are Sharable Content Object Reference Model (SCORM) conformable learning objects, which have high interoperability and reusability, and a Learning Map Design Tool has also been provided by MEAT to assist teachers with systematically organizing of their course structures in the SCORM Sequencing and Navigation (SN) form (ADL, 2006).

In 2005, Simon et al. summarized three authoring types of adaptable content generation approach, which are introduced in following (Simon et al., 2005).

• Multiple Authoring: The Multiple Authoring generates multiple versions of content to fit each specific electronic device used by learners to access course content. Contents which are generated by the Multiple Authoring approach have better exchangeability among learning platforms as well as higher management costs.

• Single Authoring: The Single Authoring provides a single source code implementation of the user interface that is valid for all devices. In addition, it is a feasible way to generate adaptable learning content, but its functionalities and design usually depend on the supported markup language interpretation.

• Flexible Authoring: The Flexible Authoring freely combines the Multiple Authoring and the Single Authoring technologies to produce adaptable learning materials. With the Flexible Authoring, content creators can create truly adaptable content by taking into account a wide variety of different factors and circumstances.

In this work, authors have implemented the Flexible Authoring approach in the MEAT, including the Multiple Authoring approach in the content authoring functionalities and the test item generation functionalities handled through the use of the Single Authoring. Based on such an implementation, the content creators can design fully interoperable custom creations without markup language restrictions. Since the format of test items is fixed, the Single Authoring method is used for test item generation functionalities, allowing items to be easily reorganized and
reducing maintenance costs. The contents and test items produced by MEAT are the adaptable learning materials, which can be learned by learners with any device at anytime and anywhere, and that is why the MEAT could contribute to m-learning field.

In this study, authors intend to share the experience of the MEAT development, which consists of technical aspect issues and encountered problems of engagement. An extensive comparison with other authoring tools and the academic evaluations were also conducted in this article to reveal the usability of the MEAT. Based on our findings, we have summarized a series of authoring tool design guideline, and we hope these valuable experiences can help other developers in mobile learning area.

Mobile E-learning Authoring Tool

There are three major roles that must play in the authoring environment to produce the learning content. These are learning designer, programmer, and artist. In the authoring environment, the learning designer bases on the goal of instruction to select a proper way to design the learning scenario and content, and the productions of this phase of development are usually limited to descriptive text. The programmer is then providing technical support according to the requirements of the content and scenario, and several software components and tools are produced along with this production phase. Meanwhile, the artist focuses on embellishing the learning interface by using appropriate images and layout designs. Subsequently, all productions in previous phases are well organized as the learning contents (organization of the content structure), which also further be arranged in a fitting manner based on the learning designer’s previous experiences (organization of the course structure, such as SN). While the learning contents being conducted on Learning Management System (LMS), the formative assessments are also utilized to determine the current state of the learner’s understanding. The assessment result then can be treated as the useful feedback for content evolution and refinement. The entire instructional content development process can also refer to rapid prototyping supported by ADDIE (Analyze, Design, Develop, Implement, and Evaluate) model (Molenda, 2003). Most often, teachers fill all three roles in the authoring environment. The great advantage of the MEAT is taking care of the most complicated technological functions for teachers. Consequently, the MEAT frees up teachers’ precious time, which can then be devoted to content improvement or other tasks. The functionalities of the MEAT are shown in Figure 1, and details will be introduced in following sub-sections.

<table>
<thead>
<tr>
<th>Support Formats</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video</td>
<td>Video record</td>
</tr>
<tr>
<td>Audio</td>
<td>Audio record</td>
</tr>
<tr>
<td>Slide (Power Point)</td>
<td></td>
</tr>
<tr>
<td>Web page</td>
<td>Screen Motion Capturing</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Transformation functions

From the content authoring aspect, material transformation is an essential preprocessing work that aligns different formatted materials to predefined formats. In our case, the accepted image format is JPEG, and the multimedia audio
and video files are transformed to streaming. In this work, we have expended the most effort on transforming slide-typed lectures, which happens to be the most popular electronic teaching medium in universities. We transformed each slide to a JPEG image without preserving the animations on the slide. This not only captures the snapshots of slide but also extracts all slide texts, which enables the system to automatically generate a preview of information of the slide for students. Using the preview information, students can get an abstract of the slide, and they can then decide whether or not it is worth studying according to their demands.

Figure 2: Synchronized adaptable learning content on different devices. Notice that the images in mobile devices support the zoom in/out function for students to check details of images (see the leftmost two images of Fig. 2c)
Synchronization functions

According to Dual Coding Theory (Tan et al., 2003), learning efficiency is enhanced by engaging both visual and verbal learning, and therefore, to synchronize multimedia and static learning materials is an important task. The Synchronization functions consist of a tool set to assist teachers to synchronize streaming (audio and video) and static learning materials (slides, web pages, FLASH), and for any synchronized learning object, its static materials would automatically play according to the associated time stamps of streaming. Figure 2 shows the synchronized adaptable learning content in different devices.

Figure 3: The snapshot of Post Synchronization in MEAT

Figure 4: The conceptual and virtual mapping of Extensible Time Stamp Indexing Model
Based on the synchronization types, the synchronization functions provided by MEAT can be classified into two categories: the Post Synchronization and the Real-time Synchronization. The two types of synchronization are similar and their main difference is the streaming sources. In Post Synchronization, streaming is sourced from pre-recorded audio/video, and during the synchronization process the teacher can search the stream to find a specific scene for synchronizing. This capacity to undertake a direct search in audio or video stream speeds up the synchronizing progress because there is not a 1:1 ratio between streaming time and the time needed for synchronization. In our experience, to synchronize a one hour lecture stream takes only about 10 minutes. Figure 3 shows the snapshot of Post Synchronization. Differing from Post Synchronization, Real-time Synchronization records the teacher’s lecture as a video stream when synchronizing. Since Real-time Synchronization course is continuous, it takes 1:1 time to produce the lecture video, which is the limitation of Real-time Synchronization. Nevertheless, the Real-time Synchronization still advances in teaching in classroom where teacher can teach while doing the Real-time Synchronization. After the class, the learning content is produced and ready for students. This efficient process indirectly improves students’ learning efficiency, thereby fulfilling the purpose of “Rapid e-Learning” (Archibald, 2005).

To precisely associate static materials with the correct time stamps of streaming is one of the major tasks in Synchronization phase. This study adopts the XML (Extensible Markup Language) (W3C, 2006) and XSLT (Extensible Stylesheet Language Transformations) (W3C, 1999) to implement an Extensible Time Stamp Indexing Model for instantly associating the static material with the correct streaming time stamp, in which users can add additional information with respect to static materials. Next paragraph defines the Extensible Time Stamp Indexing Model and explains how it works from an abstract conceptual perspective.

Let \( M \) be an Extensible Time Stamp Indexing Model, which is characterized by a set of triplet \( T = \{ t_1, t_2, \ldots, t_n \} \), where \( n \) is associated with the number of registered time stamps and \( t \) is a set of meta-information of a specific time duration of given time flow \( F \) (\(|F| = \) length of streaming or time flow). For any triplet \( t = \{id, s, e\} \), where the \( id \) is an unique ordinal identification in \( T \), the time stamp \( s \) is a time index associated to a beginning of time duration, and the extension \( e \) is a profile object, which collects a set of user-defined information to describe the corresponding static learning material, e.g. title and file location. It denotes the time duration of \( t_i \) as \( d_i \), where \( d_i \subseteq F \). If \(|T| = n\), the time duration \( d_i (i = \{1, 2, \ldots, n-1\}) \) can be distinguished by \( (t_i[s] - t_{i+1}[s]) \) and the last time duration of \( d_n \) can be obtained by \( (|F| - t_n[s]) \). Given two time durations \( d_i \) and \( d_j \), where \( i \neq j \), then \( d_i \cap d_j = \phi \). Moreover, \(|\bigcup_{i=1}^{n} d_i| \leq |F|\), because there is no restriction on \( (t_1[s] = \text{beginning of } F) \). Figure 4 shows how to map the conceptual Extensible Time Stamp Indexing Model to virtual one. Considering Figure 4, the left hand side shows the conceptual Extensible Time Stamp Indexing Model, where each block presents a triplet \( t \) with respect to certain time duration of \( F \). Subsequently, the right hand side of Figure 4 is a virtual xml file, which presents the mapping result of conceptual view. Notice that the real implementation of Extensible Time Stamp Indexing Model is similar to the Virtual view of Figure 4, but some differences exist between them, e.g., tag names are different and the model structure is not exactly equal.

**Figure 5: The Seek algorithm**

Generally speaking, when playing the learning streaming, the static materials would be displayed accordingly. However, in online asynchronous learning, the students usually seek the streaming for a specific purpose (seeking
inconsistency) and navigate directly to a certain material by skipping previous lectures (skipping inconsistency). These behaviors usually create inconsistency between the streaming and the static materials, and therefore a remedial operation is required to correct this. To correct the skipping inconsistency is relatively easy, as it can be corrected by setting the navigated material as the present one. Furthermore, the seeking inconsistency can be corrected by Seek algorithm (see Figure 5). In Seek algorithm, the input \( p \) is the new seeking time point, \( T \) is the set of triplet of \( M \), and the output is the correct triplet \( t \), which satisfies \( t[s] \leq p < t[s] + d \), where \( d \) is the duration of \( t \) (its duration includes \( p \)). The only non-trivial step is steps 5-7, which consist of an iteration to locate correct triplet \( t \).

**Packaging functions**

The Packaging function is composed of three sub-functions: The Mobile Content Translation Module, The Metadata Editing Tool, and the SCORM Packaging Module. Following paragraphs will introduce how these three sub-functions cooperate to perform packaging task.

The Mobile Content Translation Module is responsible for generating mobile content according to the imported learning materials. In particular, the Mobile Content Translation Module analyzes the structure of learning materials and reconstructs a similar structure in mobile web pages. In order to deal with the image resizing work, the Mobile Content Translation Module also invokes Transformation Function to size down the images for rebuilding learning materials in a mobile manner. Following that, the Metadata Editing Tool provides a logical view of metadata for teachers to describe the (mobile) learning content, and the SCORM Packing Module then packaging the metadata file for producing a SCORM-compliant learning object. Hereinafter we summarize the selected parameters used to generate content for mobile devices.

- **Descriptive text**: The title and main text are extracted from slides and they are then presented on mobile devices by clear text-based form. It enables learners to read them by their preferred devices without taking into account the resolution issue.
- **Layout structure**: In addition to descriptive text, the original text layout structure (the nest structure, such as items and sub-items) is also reproduced into mobile-based content. Text items usually have the group relationships among themselves, and this parameter keeps the relationships in mobile content for illustrating the original meanings designed by creators.
- **Memorandum**: Memoranda are often added into slides as the detailed annotations or supplements for reminding instructors of important concepts in slides. Since the descriptions of memorandum are relevant to slide, they are also extracted from slides to mobile content.
- **Sequence of slide**: Based on the sequence of slide, the MEAT generates one index page (see the leftmost pictures of Figures 2b and 2c) to organize the mobile content. In addition, the page head and foot of each mobile content page has been added the hyperlink for navigating to next, index, and previous pages (see Figures 2b and 2c). Consequently, mobile learner can traverse among mobile contents via minimum clicks.
- **Snapshots of slide**: All snapshots of slides are generated and downsized to mobile manner. Owning to the main texts of slides has been extracted, the purpose of these snapshots are to provide the figured concepts (such as trend of a curve and so on) instead of the descriptive concepts.

The interoperability and reusability of learning object that conforms to the SCORM standard would be improved, and following the standard substantially reduces the cost of learning content reproduction. In order to conform to the standard, teachers have to fill in the fields of content metadata with the correct information. If the fields are not filled in correctly, the advantages of the standard are replaced with the need to go through a lengthy authoring step. In actuality, with regards to early versions of MEAT, which only provided the complete SCORM metadata filling form (see Figure 6a), there was a feedback from teachers indicating that the metadata filling work was too lengthy, and consequently there were few teachers who filled it in. It has been noted that people tend to perceive what they are motivated to perceive – that which has either intrinsic or extrinsic value (Bruner, 1966). If Bruner is correct, we can assume that teachers did not complete the filling task because of costs associated with filling the enormous metadata fields exceeding the benefits of that, and the scale of the filling work obviously has to be reduced. To this end, we summarized the key fields from complete SCORM metadata according the necessary fields defined in SCORM (see Figure 6b). Henceforth, if a teacher does not want to fill in the complete metadata, she can have the alternative of filling in the summary one, and the change result in teachers’ appreciations.
Assistant functions

There are two assistant functions in the MEAT to help teachers accomplish further instructional tasks. The first is the Test Item Design Tool, which allows teachers to design test items for evaluating students' statuses, and the second is the Learning Map Design Tool, which provides an intuitive interface for teachers to organize their teaching courses and test items. Details of the two functions are shown in following paragraphs.

The primary purpose of assessment is to provide teachers with an objective view of students’ status, and to enable teachers to determine whether or not students have sufficient proficiency to move on to the next level of learning (Tan et al., 2003). In addition, assessment results can also provide the valuable feedback to the teacher for indicating
the suitability of the instructional style. Figure 7 shows the interface of the Test Item Design Tool, which supports the generation of types of adaptable (single authoring) multimedia test items such as true-false questions, multiple-choice, multi-select, blank-filling, and short-answer questions. These adaptable multimedia test items could be imported into the item bank of a learning platform, and the teacher could choose a set of items to be used as a test sheet for performing online and mobile based formative/summative assessments over the Internet. Figure 8 shows an adaptable multimedia test item on some different learning devices.

![Figure 7: The snapshot of the Test Item Design Tool](image1)

![Figure 8: The snapshots of adaptable multimedia test item](image2)

The organization of a meaningful learning map is an important task, one which can be achieved by SN. The SN defines learning activities and corresponding navigations by predefined conditional rules. Moreover, the teachers can design adaptive learning sequencing for students by manipulating the rules, and the students can access appropriate learning materials according to their actual learning level at any given time (ADL, 2006). However, teachers who are not familiar with the SN specification could not perform such task efficiently. In 2005, Chen et al. (Chen et al., 2005) incorporated Petri net into a fuzzy environment to model the learning map visually and transfer the visualized learning map to SN form. Their result indicates that Petri net is a feasible approach to bridge the gap between SN
specification and teachers’ level of expertise. In this work, authors have utilized the Petri net to implement a Learning Map Design Tool to assist instructors with course organizing tasks. Teachers who are interested in designing learning map can easily manipulate the user-friendly drag-and-drop interface of Learning Map Design Tool to organize their courses’ structures (see Figure 9).

Initially, the Petri net was developed by Petri (1962) as a mathematical representation to model the discrete distributed systems. Based on the high-level structure, Petri net can be utilized to present knowledge and complex systems. Moreover, the ease of visualization, one characteristic of Petri net, also facilitates the arrangement of learning objects in learning maps. With concrete perspective, the components of the Petri net are bipartite directed graphs, with two types of nodes: places and transitions. The places formed as circles are utilized to represent learning materials, the transitions formed as bars are registered as assessments with predefined thresholds, and the arcs formed as arrow lines are used to represent the interrelationships among the places and the transitions. In addition, the places in Petri net may associate with a token value to represent current Petri net status, and the token value can also be treated as learner token to trace and record learner’s status, portfolio, and personal profile (Chen et al., 2005). Based on learner token, the Petri net based learning map can then support adaptive learning to learners. Figure 10 shows an example of Petri net which developed by Chen et al. (2005). In Figure 10, a learner is now learning Content A. After completing Content A, if the learner passes the Assessment A, the Content B will then be taught to the learner (solid arrow line represents the static learning path) otherwise the Supplement to Content A will be activated (dotted arrow line registers the alternative learning path). As the result, the learner can acquire adaptive learning materials through the formative assessments and the well-organized learning maps.

Figure 9: The snapshot of the Learning Map Design Tool

Figure 10: An example of Petri nets
Design of manipulation process

The MEAT includes 16 different content authoring processes, 5 different test item generation processes, and 1 learning map design tool, for a total of 22 similar functions on the screen, which overwhelmed most of our pilot group of teachers. The pilot group indicated that about 8 items on a screen at a time is the upper limit. Members of the group also stated that they did not like irrelevant information or components on screen while authoring. A similar idea can be found in cognitive psychology theory, which states that the capability of people’s working memory (short term memory) is limited to between 5 and 9 pieces of information (chunks) being held in the working memory at a time (Tan et al., 2003). To overcome this problem, we have moved in the direction of developing a process, which can locate the system function efficiently and meet the teachers’ technological needs. Our main idea is to group similar functions and to adopt functions’ features as the keys to efficiently locate the proper function. In the first phase, we cluster similar functions into groups. After clustering, the content authoring group includes 16 functions, the test item generation group includes 5 functions, and the course management group has 1 function in it (the Learning Map Design Tool). At the end of the first phase, the test item generation group and the course management group were satisfactory and did not need further modification. In the second phase, we used a two layer learning material selection to locate the appropriate authoring function in the content authoring group. Figure 11 shows the grouped two layer selection process, in which teachers can follow the up to two steps of guidance to easily launch their desired authoring functions.

Evaluations

This section consists of two sub-evaluations, which are utilized to appraise the usability of the MEAT from both technical and academic perspectives. In the first experiment, we have conducted an extensive comparison between our authoring tool and other works in the first evaluation. In the second experiment, we have conducted a case study of nurse education in National Cheng Kung University Hospital in which we have surveyed the teachers and the nurses who involved in the online courses by using questionnaires. In addition, we have also summarized some potential weaknesses of the MEAT in the end of section.

Comparative analysis of authoring tools

To fairly evaluate the proposed tool, we have conducted the critical comparative analysis of authoring tools in this subsection. The qualitative comparison of authoring tools is quite difficult to be conducted in an objective manner. Therefore, only the objective technical criteria are chosen to analyze the comparison. Following part introduces the adopted criteria.
- **WYSIWYG**: The WYSIWYG stands for “What you see is what you get”. This design feature furnishes content creators with an intuitive view to develop their contents. Furthermore, the feature also saves lots of time from troublesome production checking, and the saved time can be further used to enrich learning contents.

- **Multimedia**: This criterion indicates the supported multimedia types of the authoring tool.

- **Content/Test Item Production**: This variable reveals the authoring tool’s ability of learning resource creation. In addition to the production resource types (content and test item), the resource execution environment is also taken into consideration (web/client based resource).

- **Adaptation**: This factor indicates whether the production (content/test item) of the authoring tool is adaptable to accommodate to different learning devices, such desktop, PDA, cellular phone, and so on.

- **Course Organization**: The course organization function furnishes teachers with organization and management functions to arrange learning resources in a systematic manner. Particularly, suchlike function can assist teachers with producing learning maps or sequencings to learners, and these productions can be further applied to LMS for achieving adaptive learning.

- **Standard**: This variable shows the supported standards of the authoring tool, such as QTI, SCORM, and so on.

- **Client/Web**: This criterion indicates the execution environment of the authoring tool.

- **Purpose**: The design purpose of the authoring tool is determined by this criterion.

To guarantee the quality of the evaluation, eight authoring tools have been selected into this comparison. Since the information technology advances rapidly, only the relevant research works published within the recent five years could be considered as our comparative candidates. As the MEAT is introduced in this article, the other authoring tools and their features are briefly introduced below. Note that not each candidate has a name, and it assigns alphabetical ID to the unnamed candidates for identification.

- **IWiLL** (Kuo et al., 2004): This web-based authoring tool is integrated in the IWiLL learning system, and the aim is to create interactive multimedia contents for English learning. To shorten the content production process, movie clips and subtitles are transformed from DVD to online multimedia database, which allows teachers and content creators to insert multimedia into English learning contents and test items directly. With the supports from the tool, teachers can create content with their preferred design styles, and students can searching relevant contents by importing vocabularies.

- **A** (Simon et al., 2005): The authoring tool supports single authoring for producing adaptable web-based contents. Creators who are interested in authoring by the tool should understand XHTML (Extensible Hyper Text Markup Language). To visualize the XHTML, three device simulators (PDA, smart phone, and WAP phone) have been integrated in the authoring environment for preview.

- **Test Editor** (Romero et al., 2006): The Test Editor is designed for generating adaptable test items by single authoring approach. The generated test items are archived in a XML file, which are reusable and can be interpreted by test engines for assessment. In addition, both the adaptive and the classic linear tests are supported by the test engines.

- **FAÇADE** (Zhan & Kurz, 2005): This tool supports single authoring for adaptable web-based content creation. The FAÇADE framework is a context-aware adaptable tool for single-source web pages. Users with the authoring tool can author with a WYSIWYG graphical user interface, and then the corresponding XHTML for devices will be automatically generated.

- **B** (Wang et al., 2007): This study utilizes web 2.0 technology to design the rich-client authoring environment for creating learning contents which compatible with various e-learning standards without redundant efforts. When authoring with the tool, teachers can directly utilize available learning resources from backend server without pre-downloading into local host, and the index will be automatically made for connecting present authoring and external learning resources.

- **Mobile Author** (Virvou & Alepis, 2005): Mobile Author allows instructors to design mobile Intelligent Tutoring Systems (ITS), which contains mobile contents and test items by both mobile phone and desktop. In addition, each student’s learning behaviors are stored in the long term student model of ITS for adaptive learning.

- **HyCo** (Garcia & Garcia, 2005): The HyCo is an authoring tool introduced to facilitate the composition of hypertext, which are stored as semantic learning objects in backend database. The aim of the HyCo is authoring through a simple and extremely intuitive interface and interaction model. Following that, any teacher with a minimum computer background has the possibility of transforming his/her experience and knowledge into useful and quality hypermedia educational resources.
<table>
<thead>
<tr>
<th>Criteria/Tools</th>
<th>MEAT</th>
<th>IWILL</th>
<th>A</th>
<th>Test Editor</th>
</tr>
</thead>
<tbody>
<tr>
<td>WYSIWYG</td>
<td>Supported</td>
<td>Supported</td>
<td>Not Supported</td>
<td>Not supported</td>
</tr>
<tr>
<td>Multimedia</td>
<td>PowerPoint</td>
<td>Multimedia Database</td>
<td>Image</td>
<td>Text</td>
</tr>
<tr>
<td></td>
<td>Web Page</td>
<td></td>
<td>XHTML</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FLASH</td>
<td></td>
<td>VoiceXML</td>
<td></td>
</tr>
<tr>
<td>Audio files</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video files</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audio recording</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video recording</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content / Test Item Production</td>
<td>Web-based content production / Web-based test item production</td>
<td>Web-based content production / Web-based test item production</td>
<td>Web-based content production / No</td>
<td>No / Client and Web-based test item production</td>
</tr>
<tr>
<td>Adaptation</td>
<td>Flexible Authoring Content (Multiple Authoring) Test Item (Single Authoring)</td>
<td>No</td>
<td>Single Authoring</td>
<td>Single Authoring</td>
</tr>
<tr>
<td>Course Organization</td>
<td>Learning Map Design Tool</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Standard</td>
<td>SCORM</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Client/Web</td>
<td>Client</td>
<td>Web</td>
<td>Client</td>
<td>Client</td>
</tr>
<tr>
<td>Purpose</td>
<td>Generation of adaptable learning contents and test items</td>
<td>Generation of English learning contents</td>
<td>Generation of adaptable web-based interfaces</td>
<td>Generation of adaptable test items</td>
</tr>
<tr>
<td>Criteria/Tools</td>
<td>FAÇADE</td>
<td>B</td>
<td>Mobile Author</td>
<td>HyCo</td>
</tr>
<tr>
<td>WYSIWYG</td>
<td>Supported</td>
<td>Not supported</td>
<td>Not supported</td>
<td>Supported</td>
</tr>
<tr>
<td>Multimedia</td>
<td>Images</td>
<td>Learning Object Repository</td>
<td>Text</td>
<td>Multimedia database</td>
</tr>
<tr>
<td>Content / Test Item Production</td>
<td>Web-based content production / No</td>
<td>Web-based content production / No</td>
<td>Web-based content production / Web-based test item production</td>
<td>Web-based content production / No</td>
</tr>
<tr>
<td>Adaptation</td>
<td>Single Authoring</td>
<td>No</td>
<td>Single Authoring</td>
<td>No</td>
</tr>
<tr>
<td>Course Organization</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Standard</td>
<td>No</td>
<td>SCORM IEEE LOM TW LOM</td>
<td>No</td>
<td>LOM</td>
</tr>
<tr>
<td>Client/Web</td>
<td>Client</td>
<td>Web</td>
<td>Web</td>
<td>Web</td>
</tr>
<tr>
<td>Purpose</td>
<td>Creation of adaptable web pages</td>
<td>Generation of general learning contents</td>
<td>Generation of adaptive mobile contents and test items.</td>
<td>Generation of general learning contents</td>
</tr>
</tbody>
</table>

Based on the selected criteria and authoring tools, Table 1 summarizes the comparative analysis result. Among the comparative candidates, only our tool can produce both adaptable learning contents and test items while supporting content exchange standard. In addition, the MEAT is the only tool which provides visualized user-friendly course organization function.

From implementation perspective, the authoring tool can be implemented as client-based or web-based tool. With the client-based tool, it has richer and friendlier user interface, and the generation process can be done at local host.
without Internet connection. However, authors with client-based authoring tools usually have to manage the duplicate learning resources both at local host and on LMS, and the external learning material support is also relative weak due to its lack of good integration with LORs (Learning Object Repository) or external (multimedia) backend databases. Opposite to the client-based authoring tool, the web-based authoring tools are usually implemented as a portion or function of entire LMS or LOR. Suchlike integral design allows authoring tool to directly upload its production to databases and also to make links between present design and existing learning resources. Moreover, thanks to advances of Internet and Web 2.0 technologies, web-based authoring tools provide friendlier user interface than before, however some sophisticated functions such like real-time recording are still difficult to implement. The major weakness of web-based authoring tool is that the requisite of Internet accessing. With that, two problems are derived. One is teachers cannot do authoring without Internet connection, and the other is the authoring process would be interrupted by an accidental Internet disconnection. Especially for the latter problem, the tool should provide recovery function to avoid data and status losing. The MEAT is a kind of the client-based authoring tool. While designing the MEAT, we believed that a decent tool must be platform-independent, giving more design flexibility and reliability when teachers authoring with it.

Case study: Nurse education

The MEAT was introduced to facilitate teachers who do not have computer science background to create learning contents and test items with both web and mobile based settings. To reveal the usability of the MEAT, in this subsection we introduce the case study of online nurse education in National Cheng Kung University Hospital (NCKUH).

Since the nurses in NCKUH work with the three-shift workday system, the regular training course is not applicable to this case. However, the nurse training programs and assessment are necessary in hospital, and the assessment result is one of the important criteria of promotion. In this situation, Nurse Department recorded all lectures as VCDs as the asynchronous learning contents for nurses who cannot attend at regular classes. In addition, the Nurse Department has to arrange a proper time to perform nurse assessment. In particular, it is quite difficult to practice assessments since nurses who took the same course may not have the same available time to participate in the examination event. Until two years ago, the Nurse Department of NCKUH decided to choose e-learning as the solution to train and assess nurses. The selected LMS is ANTS (Agent-based navigational training system), which was developed by Jeng et al. (2005), and the MEAT was chosen by Nurse Department as the authoring tool. In the training program, each nurse can choose the courses she needs, and then she can obtain the certifications from Nurse Department if she passed the online assessments. In the nurse training site, all contents and test items are provided by instructors and produced by MEAT. Until now, there are 19 courses and 483 test items available on the ANTS for nurse education.

To reveal the usability of the MEAT, we have surveyed 13 instructors and 583 nurses (students) involved in the training program by questionnaires. Table 2 shows the surveyed results from the instructors. All instructors thought they can easily find out the wanted functions without strenuous training and manual by their side. 61.5% of the instructors believed that synchronization function is easy to manipulate. However, some instructors indicated that there are too many controllable components on the synchronization panel, which let the instructors who first engaged in authoring feel disoriented and fall into the difficulty of information overload (Eppler & Mengis, 2004). All instructors appreciated the summarization of complex SCORM standard and agreed that it is easy to operate the authoring tool through the step-by-step authoring process. In terms of the produced learning contents, teachers thought the web-based contents are better than the mobile-based contents. Nevertheless, they appreciate that the MEAT can produce the acceptable mobile-based contents without paying extra efforts and the productions can be learnt by students in mobile manner. More than half teachers (53.85%) thought the Test Item Design Tool is easy to manipulate. However, some teachers indicated that after several times of engagement, they finally understood the tool can produce different types of test items at a time. We believe there are two reasons to cause this situation: 1) the lack of hints on the test item switch panel and 2) teachers do not need (frequent) switch panel among functions since teachers usually design the same type of test items at a time. Based on the evaluation result of test item, most teachers thought the web-based one is satisfactory (61.55%) but they did not agree that the overall condition of the mobile assessment is good enough for assessing nurses. According to the instructors’ opinions, they mentioned that they worried about that the instable mobile network may interrupt the assessment progress and the delayed transmission of test items may harm students’ interests. Nevertheless, in terms of rehearsal, instructors still approved
the convenience of mobile-based assessment and they encouraged us to pay more efforts to complete the insufficient parts. In our survey, all instructors did not use the Learning Map Design Tool in their teaching and therefore they cannot judge the academic value in either manipulation or production of Learning Map Design Tool. Their opinions revealed that they did not know what Learning Map Design Tool is and they thought it is not necessary to know when engaging in distance instruction. Finally, even though there is no instructor familiar with all functions of the MEAT, instructors still approved that the MEAT is easy to manipulate and has helpful advantages in e-learning resource production.

Table 2: Results of the survey questionnaire from instructors

<table>
<thead>
<tr>
<th>#</th>
<th>Question</th>
<th>Very approvable</th>
<th>Approvable</th>
<th>Neutral</th>
<th>Opposing</th>
<th>Very opposing</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>You can find out correct authoring function without hard practice and manual help.</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.15</td>
</tr>
<tr>
<td>2</td>
<td>The engagement of learning content synchronization is easy.</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3.54</td>
</tr>
<tr>
<td>3</td>
<td>Comparing to complete SCORM metadata, you prefer to fill in summarized SCORM metadata.</td>
<td>6</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.46</td>
</tr>
<tr>
<td>4</td>
<td>The step-by-step learning content production process is smooth and easy to follow.</td>
<td>1</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.08</td>
</tr>
<tr>
<td>5</td>
<td>The presentation of produced web-based learning content is satisfactory.</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.38</td>
</tr>
<tr>
<td>7</td>
<td>The presentation of produced mobile-based learning content is satisfactory.</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>3.77</td>
</tr>
<tr>
<td>8</td>
<td>The Test Item Design Tool is easy to manipulate.</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>3.54</td>
</tr>
<tr>
<td>9</td>
<td>The presentation of produced web-based test item is satisfactory.</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.85</td>
</tr>
<tr>
<td>10</td>
<td>The presentation of produced mobile-based test item is satisfactory.</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>1.92</td>
</tr>
<tr>
<td>11</td>
<td>The Learning Map Design Tool is easy to manipulate.</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>The production of Learning Map Design Tool is beneficial to students.</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>From overall point of view, you think that the MEAT is easy to manipulate.</td>
<td>4</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.31</td>
</tr>
<tr>
<td>14</td>
<td>From overall point of view, you think that the MEAT can help you with producing learning resources.</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.38</td>
</tr>
</tbody>
</table>

Table 3 shows the survey results from nurses with regard to the satisfaction of the system. According to the result, most nurses (about 87.23%) approved that the presentation and manipulation of web-based learning and assessment were satisfactory. However, only small portion of nurses (about 6.7%) knew the existence of mobile-based learning site. Based on the small portion of nurses’ opinions, they mentioned that either to browse online learning contents or to engage in online assessment with mobile setting was too expensive even though it is convenient. From overall
point of view, nurses agreed that learning and assessing in e-learning manner were beneficial to them since it frees them from time and space restrictions. Finally, respondents also suggested us allow them to download the mobile version of learning contents from web-based training site. With the downloadable mobile-based contents, the nurses can install them into their cell phone and the nurses can then read the mobile contents without extra Internet connection fee.

Table 3: Results of the survey questionnaire from nurses

<table>
<thead>
<tr>
<th>#</th>
<th>Question</th>
<th>Very approvable</th>
<th>Approvable</th>
<th>Neutral</th>
<th>Opposing</th>
<th>Very opposing</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The web-based learning interface, which includes presentation and manipulation, is satisfactory.</td>
<td>63</td>
<td>488</td>
<td>31</td>
<td>1</td>
<td>0</td>
<td>4.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.8%</td>
<td>83.7%</td>
<td>5.3%</td>
<td>0.2%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>The mobile-based learning interface, which includes presentation and manipulation, is satisfactory.</td>
<td>0</td>
<td>11</td>
<td>547</td>
<td>23</td>
<td>2</td>
<td>2.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0%</td>
<td>1.9%</td>
<td>93.8%</td>
<td>4.0%</td>
<td>0.3%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The web-based assessment interface, which includes presentation and manipulation, is satisfactory.</td>
<td>27</td>
<td>439</td>
<td>111</td>
<td>6</td>
<td>0</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.625%</td>
<td>75.325%</td>
<td>19.025%</td>
<td>1.025%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>The mobile-based assessment interface, which includes presentation and manipulation, is satisfactory.</td>
<td>0</td>
<td>2</td>
<td>541</td>
<td>33</td>
<td>7</td>
<td>2.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0%</td>
<td>0.3%</td>
<td>92.8%</td>
<td>5.7%</td>
<td>1.2%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>From overall point of view, you think that learning and examining on the e-learning system benefited you.</td>
<td>40</td>
<td>519</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>4.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.9%</td>
<td>89.0%</td>
<td>4.1%</td>
<td>0%</td>
<td>0%</td>
<td></td>
</tr>
</tbody>
</table>

According to the opinions from both teachers and nurses, it reveals the advantages and potential weaknesses of our system. In addition to above results, we summarize the potential drawbacks of the MEAT in the following items, which provide us an explicit direction for further investigation.

- Since the formats of multimedia are arbitrary, the Transformation function cannot properly deal with special cases of file format. For instance, we have encountered a video material with embedded hyperlinks. When playing these types of video, the embedded hyperlinks would automatically open a new browsing window and then navigate to a specific web page.
- The Mobile Content Translation Module supports only slide-typed media translation.
- The MEAT does not integrate with LOR and cannot support direct file upload function.
- Twofold packaged content size is necessary for supporting content re-authoring function. Owing to the materials of the produced learning content are transformed, they cannot be directly used in the re-authoring process. It causes original learning materials have to put into the content package, which increases the size of file and the content preservation cost of the learning platform. We believe the shortcoming would be mitigated when all adopted materials use advanced data formats (e.g. JPEG2000) to allow them to do self-adaptation.
- Teachers cannot customize their content template.
- The language in the complete SCORM metadata is not be translated to Chinese.
- Since the generated adaptable test items are produced by single authoring technology, these adaptable test items can only be executed on learning platforms that support specific markup language interpreting. This restriction would not let us provide a test item preview function to the teacher client.

Conclusions

To sum up this study, we have presented an authoring tool named the Mobile E-learning Authoring Tool (MEAT), which can produce adaptable learning content and test items to support both e-learning and m-learning. In addition, all produced learning contents are conforming to the SCORM standard, which enhances the contents interoperability and reusability. In order to conform to the requirements of the SN, the MEAT also furnishes teachers with a graphical learning map design tool for organizing teaching courses in an appropriate manner.
According to evaluation result, the MEAT is the only tool, which can produce adaptable contents and test items while supporting SCORM standard. Besides, the WYSIWYG feature and intuitive user interface also shorten the creator’s learning curve. Although MEAT does not have external backend databases to support direct resource upload and usage, it still provides a stable and reliable authoring environment for creators to produce satisfactory adaptable learning contents and test items. In the MEAT, the designs of functions and the manipulation process reflect feedbacks received from teachers and in accordance with their technological learning needs. Our evaluation indicates that the teachers are satisfied with the MEAT although there are still some shortcomings to its use. The further direction of this work is to look for ways to mitigate or entirely eliminate these shortcomings and continuously improve our system in accordance with the needs of teachers and learners in an e-learning environment.

Finally, with user’s feedbacks and our experiences of engaging in developing the MEAT, we organize some recommendations for readers who would like to pay efforts to develop authoring applications. We also hope to inspire readers through sharing our investigation experiences.

- **Based on users’ preferences to decide the support multimedia types:** In the past, teachers were based on available tools and technologies to develop their contents and media. Since teachers usually do not like to reproduce old contents with new technologies, authoring tool developers nowadays should take teachers’ preferences of media into consideration to decide the support multimedia types in their developments, especially when the media have been created by teachers with previous technologies.

- **Based on users’ demands to develop functionalities:** Many novel research projects or applications are developed based on positive assumptions, however some of them finally become useless, such like our Learning Map Design Tool. To void suchlike circumstance, it recommends the target tool user should be invited into the developing group as the advisor through entire developing process.

- **To simplify the manipulation interface:** According to the experience of the evolution of the MEAT, we discovered that teachers do not like irrelevant controllable components or information on authoring interface. In addition to redundant information, placing too many controllable objects on interface at a time is not suitable either. With our survey, up to eight objects showing at a time is acceptable.

- **To embellish the manipulation interface:** In the very early version of the MEAT, we adopted the default form of windows components in our interface design. Many opinions from teachers especially who do not have computer science background, indicate that the interface is icy and they do not like to use it. Thereby, there is a recommend to pay more efforts on embellishment, which would bring you unexpected benefits.

- **To simplify the manipulation process:** Based on our experiences, users prefer the simple manipulative action at each authoring stage. In addition, to divide one complex action into several smaller parts and then to serialize them are better ways for teachers to accomplish the original complex task.

- **To build the scaffolding instruction in authoring tools:** The concept of scaffolding was introduced by Bruner (1983), and which provides an explicit strategy for teachers to offer interactive support to learners to bridge the gap between present skill level and next level of skill. Based on the concept, scaffolding instruction can be offered to users when he/she uses certain functionality in the beginning. Besides, the scaffolding instruction can lead users to perform a certain task by a step-by-step guidance, and the guidance would fade when the users familiar with the operation.

- **Component layout principle - from top to bottom and from left to right:** For the case that several manipulative actions have to be done by a specific order on one panel or interface, the better layout principle on interface object is from top to bottom for primary order of objects and then from left to right for secondary order of objects. With such kind of layout, users can follow it to complete a systematic movement.

- **To localize the language toward to target users:** The selection of language on system interface should take account of users from different countries, and it would be the best to set up the language system with which users can choose their native language.

- **To integrate with LOR or backend multimedia database:** The authoring tool integrates with LOR can support direct upload learning object to database and also can utilize available learning resources on server. Such feature can not only simplify the authoring process but also shorten the required time of learning content production.

- **To support e-learning standards:** For reusability and interoperability, the productions of authoring tools should follow e-learning standards such as LOM, QTI, and SCORM.

- **To decrease the control complexity while increasing interactivity of the produced learning resources:** It seems to have conflict between decreasing the control complexity and increasing the interactivity on learning contents at a time. Although it is difficult to take two dimensions into account at a time, there is a successful example such
as Apple iPod (Apple Inc., 2007) utilizing one controllable button to handle all functions. Therefore this is worthy of paying efforts to achieve the goal.

• To utilize new adaptation technologies to produce learning contents: In this article, we have mentioned three content adaptation approaches: single authoring, multiple authoring, and flexible authoring. These adaptation methods only focus on producing adaptable web pages, excluding the adaptations of multimedia in the web pages. In order to reach the full adaptation of learning contents (both adaptations of web pages and multimedia are taken into account), the multimedia adaptation technologies should be involved into the content production mechanism. For example, the BSDL (Bitstream Syntax Description Language) was introduced to perform multimedia resource adaptation (Panis et al., 2003).

Acknowledgements

This work was supported in part by the National Science Council (NSC), Taiwan, ROC, under Grant NSC 95-2221-E-006-307-MY3, NSC 95-2221-E-006-306-MY3, and NSC 96-2524-S-032-001.

References


Qualitative assessment across language barriers: An action research study

Johannes C. Cronjé
Faculty of Informatics and Design, Cape Peninsula University of Technology, PO Box 652, Cape Town, South Africa, 8000 // Tel: +27 21 460 3877 // Fax +27 21 460 3553 // johannes.cronje@gmail.com

ABSTRACT
If students cannot express themselves in the language of the assessor, and if the assessor is not familiar with the cultural constraints within which students operate, it is difficult for the assessor to collect evidence of adequate performance. This article describes the assessment of three digital artefacts where the assessor strove to understand, rather than rate the students and their achievement. Three distinct approaches were followed. The first was positivist, rubric-based. The second was post-structural/interpretive, with groupwork and peer assessment. The third was feminist/holistic relying principally on students' own presentation of video material of what they did. Data sources include artifacts produced by students, assessment rubrics and assessment scores. Results indicate that qualitative assessment with criteria negotiated between the instructor/assessor, the twelve learners and their peers over a period of two years, allowed for a 'thick description' of the evaluation experience that adds to student development. Further research needs to be conducted into standardizing the procedures of auditing assessments, rather than standardizing procedures of assessing.

Keywords
Language barriers; qualitative assessment; blended learning; computers in education

Introduction and background

This article is a sequel to the story of how we taught a two-year Master’s programme in Information and Communication Technology (ICT) for Education in Khartoum, Sudan. (Cronje, 2006). It considers the continuous assessment strategies used in three of the courses. The article is built upon Brenda Johnston’s “examination of different approaches to agreement over outcomes” (Johnston, 2004) and is presented here as contribution towards beginning to fill some of the “many gaps in the research on innovative assessment methods such as portfolios” (Johnston, 2004, p. 396). The article also enters in dialogue with Wilkerson & Lang, (2003) who present very useful caveats when using portfolios in “high stakes” assessments – such as for graduation or promotion reasons. The specific application of the research lies in the domain of assessment across cultural and linguistic barriers, for as is pointed out by MacKinnon & Manathunga, “If our assessment continues to be based upon a Western template of knowledge that only values Western ways of knowing and learning, all our lip service to developing interculturally competent students is meaningless” (p. 132).

The purpose of this article is not to “prove” or “reject” any cross-cultural assessment theories. Neither is it an attempt at reporting quantitative, transferable results but rather to share the story of how we dealt with a practical manifestation of the above problem that arose when two colleagues from Pretoria, South Africa, and I presented a Masters’ course in Information and Communication Technology (ICT) for Education in Khartoum, Sudan. A comprehensive account of the instructional design and presentation of the course has been published in the Journal of Educational Technology and Society (Cronje, 2006). One table and two figures from that article are repeated here. In the previous article they were used to explain the instructional design, whereas they are used here to explain the assessment strategies and outcomes.

The key research question addressed in this article is: “How could we assess fairly a group of students with whom we did not share a common first language, or a common culture?” Or more succinctly, how did the instructors and students create common understanding across language and cultural barriers?

Linguistically the problem was that the students could understand us reasonably well. They could read English, and they could, in their groups, support one another in the reading and interpretation of English articles, but they were shy to speak English and their writing was poor. Thus, even if they knew and understood the content, and even if they were able to apply it, their ability to display that insight was marred by their inability to write clear, comprehensible English. Culturally one of the first obstacles was not behavioural, but aesthetic. What they regarded as beautiful in terms of screen design we regarded as gaudy. What we regarded as elegant, they considered bland.
At a deeper level, though, there was also the suspicion that we were foreign colonialists coming to impose our value systems upon them, and that we would use assessment as a “technology” (Delandshere, 2001) to enforce our ways upon them.

The initiative came from the Vice-chancellor of their university who, after a world-wide search, had selected our programme because it was developed in Africa, based on an American model, and delivered in English. Thus, although their institution chose us, they were suspicious of us. We had to establish culturally sensitive interpretations that would lead to a commonality of expectations.

The conceptual model we devised to understand our situation is shown in figure 1.

![Figure 1. The cultural divide](image)

At the top left of the figure is local relevance and application. At the top right is the global situation. The more global something becomes, the harder it may be for locals to see its relevance. The more local things become the less globally competitive or useful they might be. Any emphasis on increasing either local relevance or global competitiveness could increase the divide. Providing too much localization could disempower the community by reducing their international competitiveness. Too much global emphasis may distract from local culture. However, the two sides come together at the bottom. Thus we argued that, in order for us to reach commonality we needed to descend into the canyon to find what is fundamental to what we were doing and work from there. Instead of assessing computer screen design for the “beauty” of colour combinations, for instance, we assessed for adherence to acceptable norms such as allowing for colour-blindness, placing background colours such as blue and green in the background and foreground colours such as red and yellow in the foreground, to avoid such colours from “bleeding” because of the physical construction of the human eye.

It was towards the end of the two-year program that the third member of our teaching team mentioned to me how impressed she was with the standard of the students. She was the last instructor to visit them and had only been brought onto the programme towards the end. Thus she was unaware of any of their previous work, and, in a sense, acted as an external evaluator. She was a regular presenter of the South African version of the course with at least six years’ experience of presenting coursework to Master students in South Africa.

The question then arose, “how can she say that?” They write poorly. They are shy. They don’t speak English well. Yet, all three of us knew that they were producing excellent work; but without written proof, how did we know? What accounted for our gut feeling that these students were equal to or better than their South African counterparts on whom we have based our development work, and who formed the baseline of our comparative gut-feel assessment? Or, more precisely, what evidence did we have to substantiate our claim that these students were masters of ICT in Education?

Three sub-questions drive the “scenes” that will be presented later. The first question is: “How detailed should we make the briefs and assessment criteria for portfolios?” The second is: “What is the value of portfolios that are created by teams?” The third question is: “How can the assessor of a portfolio gain insight into the process by which the portfolio was created?”
Literature survey

Delandshere, (2001) points out that often assessment is primarily used as a “technology” to exclude those whom the assessing authority deems unsuitable. She continues that “while there have been important technical developments (e.g. item response theory, item-biases) and technological changes (e.g. computer-adaptive testing) the procedures used have remained fundamentally the same with tests being the primary method of educational assessment” (Delandshere, 2001, p.115). Cross-culturally test bias has been identified as a major contributor to unfair exclusion. This bias is further complicated by the relationship between the individual and the group. Skiba, Knesting & Bush, identify “construct validity, content or item bias, inappropriate standardization samples, examiner and language bias, and differential predictive validity” (2002, p.63) as important sources of test bias. They caution further that: “If test results are treated solely as a representation of inherent individual contribution to aptitude, ignoring cultural and educational factors … test scores may yield inaccurate interpretations, even if the tests themselves can be shown to be unbiased” (2002, p. 75).

For us the challenge lay in assessment that was built upon a common understanding between students and instructors, despite cultural differences, of what would be assessed. In his search for such commonality in assessment at a bilingual (English/French) university in Cameroon, Amin (2002) identified six “factors which students consider important in the evaluation of courses” (p. 281). These factors were identified as Rapport with students; Evaluation of Assignments and feedback; Workload; Availability of Help; Course Organisation; and Language or Communications (Amin, 2002, p. 290).

The diverse nature of qualitative assessment makes it attractive when assessing in a cross-cultural situation, since its use of longitudinal evidence derived from multiple sources provides a rich picture of the teaching and learning reality. It also allows for a multiplicity of learning contexts that are more likely to offer a fair and authentic accounts of the potential and performance of different kinds of students. (Tigelaar, Dolmans, Wolfhagen, & van der Vleuten, 2005, p. 596).

In a richly nuanced environment the problem now becomes one of structure and standardisation so that students and assessors work from a common base. Carraccio, (2004, p386) cautions that unstructured portfolios make it difficult to achieve acceptable standards of reliability and validity in educational measurement. MacKinnon & Manathunga’s (2003) observe that “students all stated that what was required of them needs to be stated clearly and unambiguously” (p. 138). Even more seriously Wikerson & Lang (2003) argue that, should portfolios be used for certification purposes, there could be sever legal implications if the assessors cannot demonstrate that their standards of assessment are comparable to psychometric standards. Here Driessen, van der Vleuten, Schuwirth, van Tartwijk, & Vermunt, (2005, p.216) provide valuable pointers: Their strategies for realising dependability – the pendant of reliability – involve establishing an audit trail (i.e. documentation of the assessment process to enable external checks) and carrying out a dependability audit (i.e. quality assessment procedures with an external auditor).

Standardisation presents its own problems. Johnston, (2004) warns against a tendency by independent evaluators to self-moderate when assessing with a rubric, thus leading to higher inter-rater reliability, but lower validity, while Driessen, et al. claim that “detailed checklists can easily trivialise assessment” (2005, p.215). Thus the question driving the first “scene” asks exactly how clear should the instructions be – should we, for instance supply an assessment rubric together with the project brief? In this way students will know exactly what is expected. But, on the other hand, the assessment may be reduced to an almost mechanical checking off of items on the list.

In our second “scene” we tried to reduce the rigidity of “checklist” style assessment by adding team support. Driessen, et al. (2005, p.216) describe a multi-faceted approach using qualitative methods, where “the following three strategies are important for reaching credibility: triangulation (combining different information sources); prolonged engagement (sufficient time investment by the researcher), and member checking (testing the data with the members of the group from which they were collected).”

The third “scene” approaches another problem with portfolios, which is that while they give a rich description of the product that a student is able to create, they might tell us nothing about the process that the student followed, or about what the student understands of the underlying assumptions. At worst we may have no evidence that the portfolio is indeed the work of the student who submits it. The solution may lie in the way in which a digital portfolio is composed. Instead of having a single focus the portfolio could be based on a number of underlying
theoretical assumptions. Two theoretical perspectives informed this research. The first is an integrative approach to objectivist and constructivist learning (Cronje, 2006) and the second is Johnston’s (2004) four approaches to assessment of portfolios: positivist, post-structural, interpretive and feminist.

In an attempt to break the polarisation between objectivist and constructivist learning outcomes Cronje (2006) presents a four quadrant model in which the two main approaches to learning, generative and supplantive, are plotted at right angles, instead of being opposites along a straight line. Four quadrants emerge. The injection quadrant is high in supplantive, objectivist or positivist elements and relies heavily upon direct instruction. The construction quadrant, on the other hand, is high in generative, constructivist activity and learners learn by doing. Learning such as the kind described by Papert (1993), where learners engage in various construction projects in order to learn the underlying mathematics, would occur in this quadrant. The immersion quadrant explains what happens when learners are “thrown into the deep end” without any overt planning, and in the integration quadrant learners are taught directly, and allowed to explore, depending on the nature of the situation.

Followers of positivist approaches to assessment according to Johnston (2004) believe in the possibility of objectivity, scientific measurement and certainty (p 397). They strive for reliability of assessments that are done without any knowledge of the student being assessed, or of the ratings of fellow assessors. This approach closely matches Cronje’s (2006) injection quadrant. For Johnston “A central feature of interpretivist approaches to assessment is that realities, especially social realities, are perceived as mental constructions or interpretations, rather than absolute, objective truths” (Johnston 2004, p.399). This approach corresponds with Cronje’s (2006) construction quadrant.

Post-structuralist approaches, on the other hand “challenge the notion of unified, universal, master narratives, together with their embedded power relationships favouring particular groups” (Johnston 2004, p.398). She continues that “For many post-structuralists, the notion of summative judgements of individuals is so problematic and alien that they tend to focus more on formative, developmental and collaborative assessment” (Johnston, 2004, p. 399). This may match the immersion quadrant. Johnston points out that a feminist assessment system “should strive to be fair and sensitive to these differences and their implications, in order to avoid the systematic advantaging or disadvantaging of particular groups” (2004, p. 400). She points out that the key lies in development rather than accountability. Feminists are more likely to be concerned with an assessment that is sensitive to differences in race, gender class and age “in order to avoid the systematic advantaging or disadvantage of particular groups” (Johnston, 2004, p. 400). This is likely to be the quadrant of Integration (Cronje 2006).
Methodology

For this research an action research methodology (Cohen, Manion & Morrissens, 2000) was followed in the form of a disciplined self-reflective inquiry into a small scale, real-world intervention in an attempt to reform and improve practice. It is reflective in that it strives to bridge the gap between theory and practice, and is done by a practitioner focusing on his own work, combining diagnosis with reflection and focusing on practical issues. Action research was selected as a methodology because of the position of the researcher as a participant in the process. In this way the research was conducted from an internal, rather than an external perspective. On the down side, of course, this means that the researcher can make no claim of impartiality, and the extent to which the results reported here can be transferred elsewhere is limited. Nevertheless, as has been stated before the purpose of this research is not to “prove” but rather to share, and in this respect it is hoped that the results of this action research study will resonate with results obtained by other researchers.

The three “scenes” presented here are drawn from courses in a Masters’ Degree programme in Information and Communication Technologies (ICT) for Education, presented by a professor from South Africa, to twelve students of the Sudan University of Science and Technology (SUST), Khartoum over a period of two years. Pretoria, South Africa is a historically Christian city with English as its lingua franca. Khartoum, Sudan is a predominantly Muslim city with Arabic as its dominant (if not only) language.

Twelve promising graduates of the education faculty of SUST were selected to follow the program, after having attended six months of intensive English language and computer literacy training. I have been a professor of computers in education since 1994, with my main function being the co-ordination and presentation of a two-year coursework Masters degree with a mini-dissertation. In 1994 I read, and met with Seymour Papert and started working in a “constructionist” fashion. Instead of lectures I would usually give students learning tasks involving some development of technology – from which I hoped that they would construct the knowledge for themselves. In 1996 I took this practice onto the Internet in that many of the work done by my students would be web-based (Cronje 1997). The work we did in Sudan was an extension of this approach.

Assessment strategies varied throughout the course but students were required to construct a digital portfolio containing all work produced during the course. The principal data sources are artefacts produced by students during selected courses of the programme as well as the evaluation instruments and evaluation results produced during the assessment. Interpretation of data was qualitative and involved a close scrutiny to identify trends and draw conclusions.

Findings

The following section will discuss the three evaluation scenes that were selected. The first illustrates what happened when a positivist approach was followed, functioning in the injection quadrant. The second shows the implications of a structuralist/interpretivist approach in the construction quadrant, and the third tends towards a holistic/feminist approach in the integration quadrant. The scenes depict the creation and evaluation of digital artefacts that were included in the portfolios of the students during the action research cycles in which we attempted to gather and refine the evidence that allowed us to report on the progress of the twelve students on the programme. The cyclic nature of action research makes it hard to separate the description of the findings from the actual activities from which those findings were derived. Thus a narrative reporting style will be followed, and the rhythm of each scene will be “we found… so we did… so we found…”

Scene one: A mark out of 200

The first graded assignment given to the students was to conduct a dynamic evaluation of a piece of educational software of their own choice and to comment on its suitability for a given target population. The resultant essay was the main element of the portfolio for a course on “The evaluation of educational software and its effect on learning” the course served partly to introduce students to educational software, but mainly to sharpen their skills in empirical research methodology.
The students were not going to have face-to-face contact with me again for at least eight weeks. They had never handed a project to a stranger before. They were uncertain of their English writing skills. Their resultant high levels of anxiety made it necessary to provide more scaffolding than I would have given my South African students. It was clear that my usual approach of setting a task and letting students learn from their mistakes would be more than they could bear.

Discussions with the class indicated that a positivist approach with very clear boundaries would be more acceptable to them. As a result I decided to do three things. I created an article template that contained typical headings and a few useful phrases, as well as one or two citations in the text and in the bibliography that would serve as examples of how to cite and refer. I then created an assessment rubric in the form of a spreadsheet, and finally divided the students into peer support groups of three.

The assessment rubric contained 54 assessment statements with rating scales ranging from 0-1 to 0-5, depending on the criterion being assessed. For instance, the first statement was “Introduction provides clear program information (0-1)”. The template contained an example stating the name of the program, the author, the publisher and the technical specifications for the computer on which it would run. If the student supplies this information, then one mark is given, else nothing. Item 10 and 11, on the other hand were given a 0 to 5 rating (“Literature review logically arranged in sub-sections” and “Supporting and contradicting literature cited”). Rubric statements covered all aspects of a typical assessment of research papers, such as introduction and problem statement, (seven items, 14 points), literature survey, (six items, 25 points) description and defence of method (17 items, 62 points), reporting of findings (six items, 24 points), quality of interpretation and conclusions (six items, 30 points), technical aspects such as tables, and referencing style (eight items, 25 points) grammar, spelling and general impression, though these were deliberately given a much lower weighting so that students would not be penalized for poor English (four items, 20 points). The rubric added up to 200 points, which could be divided by two to make a percentage. The unusually high number of points awarded to the description and defence of the method is accounted for by the fact that the main aim of the course was to teach evaluation methods, and thus I was more interested in the methods followed than in the findings obtained.

Students were expected to do static evaluations of educational software in groups of three working on one computer and going through the software, rating it according to a checklist. Each student would each select one specific title for a dynamic evaluation on an actual target population. They would discuss the appropriate research methodology in their teams, but conduct the evaluation individually and write the evaluation report individually, using the template. Each student would submit the report to the two other members of the team who would rate it according to the rubric and give feedback. The student could then improve on the essay, self-evaluate it according to the rubric and submit to me the essay, together with the three assessments.

Table 1: Student self-assessment compared to my assessment (Cronje 2006, p. 281)

<table>
<thead>
<tr>
<th>Student</th>
<th>Own grade</th>
<th>My grade</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Md</td>
<td>83</td>
<td>86</td>
<td>-3</td>
</tr>
<tr>
<td>Yo</td>
<td>91</td>
<td>86</td>
<td>5</td>
</tr>
<tr>
<td>Et</td>
<td>89</td>
<td>85</td>
<td>4</td>
</tr>
<tr>
<td>Ta</td>
<td>68</td>
<td>82</td>
<td>-14</td>
</tr>
<tr>
<td>Om</td>
<td>70</td>
<td>82</td>
<td>-12</td>
</tr>
<tr>
<td>Id</td>
<td>76</td>
<td>76</td>
<td>0</td>
</tr>
<tr>
<td>Ha</td>
<td>75</td>
<td>75</td>
<td>0</td>
</tr>
<tr>
<td>Ib</td>
<td>70</td>
<td>73</td>
<td>-3</td>
</tr>
<tr>
<td>Ma</td>
<td>74</td>
<td>69</td>
<td>5</td>
</tr>
<tr>
<td>Ab</td>
<td>70</td>
<td>68</td>
<td>2</td>
</tr>
<tr>
<td>Ia</td>
<td>70</td>
<td>67</td>
<td>3</td>
</tr>
<tr>
<td>Ka</td>
<td>53</td>
<td>51</td>
<td>2</td>
</tr>
</tbody>
</table>

I then assessed each essay and compared my assessment with that of the student. Should the two sets of grades be within 5% of each other; I would award the student the average of the two. If the student’s grade was too low I
would moderate. If the student’s grade was too high, I would suggest improvements until the grades came within 5%. The peer grade was disregarded, as the students were able to work on the essay after having seen the peer grading. The result of the grading can be seen in table 1 (Also reported in Cronje, 2006). The abbreviations relate to the names of the students.

As it turned out it was not necessary to moderate. The two students who underrated themselves were two of the better students, and their underrating can be ascribed to a lack of self-confidence, as I determined during interviews with them afterwards. I believe this to be a product of high uncertainty avoidance, coupled with high power distance. The students were not sure that I would assign so high a grade and thus allocated themselves a lower grade which they considered more “likely”. Also they thought it would be arrogant to give themselves such a high mark.

I was initially worried that the grades were too high. In the South African system according to which I worked a pass is 50%, 75% is a distinction, and 80% is usually as good as it gets. Here all the students passed, two thirds obtained distinctions, and almost half got more than “as good as it gets”. However, I kept the grades as they were, because I knew the students had been selected based on their academic prowess. Ka’s very low grade can be ascribed to the fact that he had been ill and had missed the initial set of lectures and the discussion of the template and the rubric.

What was disconcerting, however, was the ‘mechanical’ nature of the essays. They were all close adaptations of the template and very clearly written to the rubric. Thus I had removed much creativity from the process and made the students create “instant” essays that were good, but dull. The assessment could possibly be regarded as “trivial” (Driessen, et al. 2005). Nevertheless, despite what I considered to be too much “hand holding” already, the students still emailed me frequently with questions and requests for comment on their pieces as they were working.

Scene two: Spreading the sheets

Based on my disappointment with the essays despite their high grades, I decided to move closer to my usual problem-based or “constructionist” (Papert, 1993) way of facilitating learning by letting students create a digital artefact, and letting them learn by doing. Another reason for this approach was that the work to be covered was on “Tutorials, Drills, Simulations and Games” taken directly from the text book (Alessi & Trollip 2001). I also knew that letting students construct digital artefacts as the foundation of a learning task would not work well in my absence, and thus arranged for them to engage in a cooperative learning event designed according to the “Jigsaw” model of Johnson, Johnson & Holubec (1994). Students were divided into home groups and expert groups. Each home group had to produce a spreadsheet that would contain a number of statements and a yes/no checkbox. In response to the checkboxes the spreadsheet would suggest the extent to which a given situation would best be served either by a tutorial, drill, simulation or game. They also had to design the aesthetic aspects of the spreadsheet, and program the checkboxes. The expert groups were responsible for determining the criteria for tutorials, drills, simulations and games respectively. Leaders of the home groups were assigned based on their spreadsheet skills, while leaders of the expert groups were selected based on their language skills. In this way they could better assist their team mates in terms of their zones of proximal development.

Once the expert groups had brought together the statements for each checklist the home groups negotiated to determine the values that the spreadsheet would assign to each category (tutorial, drill, simulation or game). Finally the home groups worked through one another’s efforts by way of summative and peer evaluation. The final grade was binary: successfully completed or more work required.

The screenshots of the various sheets, presented as figures 3 to 5 show that, as was the case in the essay with the rubric and template, the degree of similarity was very high. This could be because of the cooperative learning groups, and because the outcome was mastery of the same textbook content and spreadsheet skills rather than originality. As it was a group project the contribution to the final grade was minimal. Students were given a range of other individual tasks that fall beyond the scope of this article.

The first two spreadsheets (Figure 3 – Also shown in Cronje, 2006) fulfilled all the requirements. They contained 39 and 40 questions respectively, with tick boxes for a yes response. We ran a few scenarios through them, and they provided useful suggestions that seemed adequate. The two sheets contained essentially the same questions as but in different sequences as these had been compiled in the expert groups. The weighting of the responses was also
slightly different. The group who called themselves “Arnouba” separated the graph from the numeric information, using a table as well as a graph. The background colour of the spreadsheet was pastel green, but the graph had a bright blue background, while the table with the numeric values had a grey background. The colours of the letters spelling tutorials, drills, simulations, games corresponded to the colours of the bar graph. It is of note that, while the spreadsheet itself is justified to the left, the table and the graph both read from right to left instead of the conventional Western way. This is probably due to the fact that when set to Arabic, MSOffice defaults to right justification and, while it is easy to change the justification of the spreadsheet, it is harder to set the justification of the graphs. The table giving the total scores, however, was done by hand but also from right to left – in the Arabic way.

Figure 3. Methods.xls and Arnouba.xls
The spreadsheets produced by Ka’s team and Yo’s team (Figure 4) used pie charts rather than bar graphs or histograms for the graphic representation. They also chose to reveal the weighting that they assigned to the four modes of computer-based education for each statement.

The spreadsheet of Ma’s group (Figure 5) had a number of problems. Firstly the check boxes were not neatly aligned, the questions were not shuffled, but placed in groups according to the mode, and the pie chart measured
every item, not giving any result. It was clear that more work was required on the spreadsheet, but their knowledge of tutorials, drills, simulations and games appeared adequate.

The spreadsheet exercise showed how the home groups and expert groups negotiated meaning. The fact that the groups had to know enough to be able to do a given task placed the primary responsibility for the assessment on the learner. Instead of the instructor asking “Is there enough evidence that the students have mastered the content?” the students were asking “Do we have enough knowledge to complete the task?” Interpretive assessment worked well in this instance because the instructor was present to give feedback at all times and the assessment was for formative, rather than summative purposes – there being other tasks for final assessment.

Scene three: At the movies

A course on “Learning theory for computers in education” required students to follow Gagne’s (1987) (essentially objectivist) events of instruction to produce a constructivist learning event. They were then to present the event and report on what they had learnt. Students had to go through the entire “analysis, design, development, implementation and evaluation” cycle. Not only did they have to prepare and present a “lesson”, they also had to prepare computer-based support materials so that the emphasis would be on the learners’ construction of their own meaning.

From previous experience with my South African students I knew that this was a difficult task, easily misinterpreted. Students tended to concentrate too much on the development of the technology and too little on the activities that would be required of their learners. Moreover, despite their best intentions the students ended up indulging in a significant proportion of lecturing. I therefore decided that the analysis and design phase would take place during the contact session with the development and implementation being done while I am away, and the evaluation during my subsequent contact visit.

At all times the students would be involved in peer assessment and cooperative learning, while the final evaluation and assessment would take the form of a 10 minute video presentation by the students explaining to me and to their peers what they did and what they learnt.
We started with the students working in four groups of three, each presenting the background to their chosen field, and engaging in discussion until the lesson theme had been established. We then worked in the whole group of 12, during which students explained the concept of their lessons to each other and to me, and we reached a consensus on the basic outline of the lesson. For homework students then filled out a lesson plan called a “learning event protocol” according to a template that I supplied. They also had some time to design and present a prototype of the spreadsheet, database, slideshow or word processing document that they would use as the technological component of their learning events. These were also discussed and students were given another opportunity to refine them. A deadline was set for the completion of the preparation. Students would email me a complete package consisting of the learning event protocol, the supporting material and the evaluation instrument with which they would finally evaluate their learning events. I supplied feedback on a mailing list, so that every student could read the feedback that I had given to each individual. This was both in the interest of transparency and in the hope that they would learn from the feedback to their peers.

The groups of three were involved in the presentation and recording of the learning events. One facilitated the event. Another acted as an assistant to help with logistical and technological problems and the third would record the entire event on video; thus each student was involved in three presentations in different roles.

Learning Event Protocol

**Subject area:** Maths, logical functions.

**Expected outcomes:**

*After this learning event, learner should be able to:*

1. **Know** the logical functions "AND" and "OR".

2. **Using** "AND" and "OR" to linking between two statements.

3. **Distinguish** between outcome to the arguments when the link is "AND".

4. **Distinguish** between outcome to the arguments when the link is "OR".

5. **Distinguish** between symbol of "AND" and symbol of "OR".

<table>
<thead>
<tr>
<th>Time</th>
<th>Learner activity</th>
<th>Facilitator activity</th>
<th>Resource</th>
<th>Rationale</th>
</tr>
</thead>
</table>
| 20 min| Navigation through the lesson.  
- Evaluate the statements.  
- Observe the events.  
- The color of statement box help learner to know the | Encourage as many different types of complex statements.  
- Mouse.  
- Click to open hyperlink. | Slideshow.  
- Mouse.  
- Click to open hyperlink. | Make learners aware that there are different kinds of complex statements. |
edited and presented to the class upon my subsequent visit. In this way we could all see firsthand what had occurred as the student took us through step-by-step.

Here follows the discussion of the progress of one student. She was not the best student, nor was she the worst, but the case is typical of what happened in the course.

Figure 6 shows the original learning event protocol presented by the student. After some discussion the first learning outcome “Know the logical functions ‘and’ and ‘or’” was retained, as the group felt that, although “know” could not be measured, the other outcomes explained the term. We also felt that the lesson protocol should break down the time more carefully – a 20-minute interval was too long. In the refined version, however, the student went to the other extreme and tried to plan for every 2 minutes and I had to suggest longer intervals. Another critique of the initial lesson plan was that the student gave a highly theoretical overview of complex statements without referring to the students’ real life experience. To improve upon the design and reduce lecturing the student constructed a true/false quiz (Figure 7) for the students to work through by way of introduction.

I found the use of police handcuffs a strange symbol, but the student explained the metaphor of two binary devices bound together and the rest of the group liked it, so the student went ahead.
The lesson was to be conducted in Arabic, but, because the student knew that I do not read or understand the language, she built a copy in English (Figure 8).

Once the learners had gone through the quiz and the concept had been established, they worked through a spreadsheet that allowed them to test various combinations, by way of self-assessment (Figure 9).

![Figure 9. Assessing learning with a spreadsheet](image)

As in the evaluation course, students had to write a formal essay. It was assessed with a much less complex rubric than the essay described in the first scene, as it formed a much smaller component of the final grade. A much more intimate assessment was possible when we viewed the edited video produced by the student. Students used a small digital still and video camera with editing software. The camera was designed and marketed as a product to enable school children to produce their own video materials. It produced relatively low quality images, but created relatively small file sizes. The editing software that came with the camera was versatile yet intuitive. The use of this camera served two purposes, firstly it allowed for the assessment of the lessons, but possibly more importantly it was hoped that it would encourage the students to get their own learners to use the camera for projects later.

Images captured from the video (figure 10) show that the student was able to recognise and interpret various significant moments in the learning experience of the learners.

The student presented a 13 minute edited video complete with special visual and sound effects, transitions and subtitles. The subtitles included the name of the school, a description of the area being shown, such as the language laboratory and the computer laboratory, the activities taking place, such as initial installation of the software, briefing the students, and most importantly an interpretation of the events during the lesson.

The video showed that the student had been able to present a successful lesson in a computer-rich environment. In her talk accompanying the video the student explained what she had done during each phase of the lesson, and what happened. The student found it much less stressful to speak about the video playing behind her, than if she had been asked simply to talk about her lesson without any visual aids.
At a deeper level the video showed that the student was able to observe and pick out significant events during the presentation of the lesson. There was also evidence that she was able to tell a story visually, and had mastered the elements of video making. It was clear that she knew about establishing shots, mid shots and close-ups and when to use which. She had also mastered the elementary computer-based video editor and special effects generator and used those cleverly to emphasize certain events.

Over all, the case of the constructivist lesson shows a number of important aspects of portfolio assessment. Firstly, the assessment was initially formative. No grade was assigned and students were allowed to refine their work all the time to the final presentation. The aim of the assessment was to improve the quality of the students’ work, not to assign a grade. The assessment also attempted to be holistic in that the whole design cycle was assessed, and multiple instruments were used, ranging from authentic documentation in the form of the learning event protocol, through
analysis of artefacts produced for the lesson, to a formal essay assessment and an intimate self-evaluation of the video by the student.

Conclusions

This article set out to share ways in which a group of students and the author tried to reach commonality in assessing the quality of their learning across linguistic and cultural boundaries. The first scene showed that a positivist assessment had a high level of correlation between student and assessor grades, but tended to be mechanical and if students know exactly every point of the assessment, their products lack creativity. The second scene showed that the value of team productions were that students learnt from one another, while the downside still lay in a high level of “identical” work. Finally the third scene showed that a continuous, holistic assessment strategy, in which the entire process is monitored, allows us to obtain a rich, thick description of the learning journey, and a common understanding of what the student had learnt. Amin’s (2002) six factors formed the basis of developing commonality. By way of conclusion it would thus make sense to group the conclusions in terms of Amin’s six factors.

Rapport with students was crucial in obtaining their buy-in to the formative evaluation process. Negotiating the nature of the task and the criteria for assessment was a three-way process involving the instructor, the student and their peers. It was only possible to reach agreement when everybody was familiar with the situation of the person being assessed. Much of the rapport building took place outside class time (Cronje 2006) and falls outside the scope of this article, but a significant proportion of the rapport was created by way of evaluation of assignments and feedback.

Initially high anxiety levels among students meant that it was necessary to give regular feedback on small tasks, as is also suggested by MacKinnon & Manathunga (2003), who mentioned that students needed unambiguous instructions.

As time progressed it became necessary to withhold feedback to prevent students from becoming too dependent on the instructor and to encourage them to work independently.

Providing regular feedback on shorter tasks and negotiating tasks and assessment criteria with students on an almost individual basis helped regulate the workload of the students. The instructor came to know not only how competent students were, but how long it took them to become competent and how much work it took. Students were less likely to be swamped with work, as they had input into the process. However the workload of the instructor increased dramatically, as facilitating the course involved much more than preparing and presenting a traditional series of lectures and following it up with a test. However, the fact that lectures were almost excluded from the presentation mode freed up some instructor time. Nevertheless one should stress here that the scalability of this type of assessment is doubtful.

Availability of Help played an important part in ensuring the students’ success. Both the instructor and local facilitators provided help, but the process became unstuck on a number of times when institutional support was lacking during periods when I was not in Khartoum and we had to re-negotiate some tasks and criteria.

Working across radically different cultures means that course organization has to be done dynamically as both the instructor and the students become familiar with their altered situation. Much design flexibility is required. Once again a qualitative approach is useful here as the assignments provide evidence of the growth of the student, not just proof of meeting an external set of criteria.

I find it interesting that Amin speaks of language or communications rather than language and communications. In the course of this research it became very clear that communication was much more important than language. Often during discussions with students it became clear, from the animated nature of their discussion, that all was not well – although I could not understand a word of the Arabic, I could deduce from their body language, facial expressions and vocal tone that things were amiss. As time progressed it was even possible to determine the source of the problem, be it the nature of the task, university matters, interpersonal conflict, or whatever. Eventually the portfolio becomes evidence of a common understanding between all the participants involved in its creation.
Some of the concerns that emerge from the research relate to the mechanical nature of rubric-based assessment of portfolio items, which resonates with Johnston’s (2004) comment regarding the tendency of raters to self-moderate – thus detracting from the reliability of the results.

Another important aspect in qualitative assessment is the philosophical position of the assessor. It is my belief, and that of my two co-presenters, that these are good students. It was also the belief of their home institution who selected them. The purpose of our assessment was interpretive – to understand the students. The purpose of SUST’s assessment was positivist – to make the best possible and fairest (exclusive) selection. And we agreed. But, what would a completely external third party say, and would that differ if that third party were from the West or from the Arab-speaking region?

From the discussion and conclusions a number of recommendations arise.

Recommendations

Research is required to determine heuristics for the composition of credible portfolios, as suggested by Driessen, *et al.* (2005). As the heuristics develop, training should be developed for instructors and learners in the composition of portfolios. In the same way in which the skill of taking a test needs to be developed alongside the subject knowledge, the skill of composing a portfolio that reveals your skills and attitudes is a complementary skill that needs to be developed. In the same way that teachers prepare learners for an examination, they need to prepare learners for the assessment of their portfolios. In a cross-cultural context this becomes even more important as learners need to develop skills to let their true ability shine through the various cultural filters.

In this respect the role of off-task activities in building the rapport required for designing credible portfolios across cultures becomes important. Learners need to be evaluated in context, and that context extends beyond the walls of the classroom and beyond the traditional activities related to the subject matter.

Finally we need to investigate ways of making of qualitative assessment scalable. The cases presented here are drawn from a population of twelve. It would be impossible to give so much attention to a large group, yet, despite any group size each individual will be required to perform in a predictable way. I support Wilkerson & Lang’s (2003) caveat about using portfolios for purposes of exclusion, and agree with Driessen, *et al.* (2005) who believes the key lies in standardizing the procedures of establishing the audit trail and the dependability audit, rather than standardizing the procedures of the assessment itself, as the assessment should assess the learner, and the standardization should assess the assessment.

Acknowledgement

The contributions of Prof Seugnet Blignaut, Dr Dolf Steyn and the students who participated in this study are gratefully acknowledged.

References


An Exploration of Pre-Service Teachers’ Perceptions of Learning to Teach while Using Asynchronous Discussion Board

Lasisi Ajayi
Language/Literacy Education, San Diego State University, 720 Heber Avenue, Calexico, CA 92231, USA // +1 760 768–5504 // lajayi@mail.sdsu.edu

ABSTRACT
Pre-service teachers’ perceptions of their own learning are important to the ways they integrate technology into their practices. Therefore, the purpose of this study was to examine the pre-service teachers’ perspectives of asynchronous discussion board (ADB) as a tool of learning to teach. ADB was integrated into two literacy courses for 33 pre-service teachers over 16 weeks. Data were collected through oral interviews, written reflections, and participants’ postings on the discussion board. The data were analyzed using Chi’s (1997) framework of verbal analysis method. The findings indicated that the participants perceived ADB as an important tool of learning to teach because it promoted situated learning, facilitated a social construction of knowledge, and afforded customized learning experiences.

Keywords
Asynchronous discussion board, Distributed knowledge, Educational technology, Social learning, and Sociocultural theory

Introduction
The material and social affordances of new technologies of communication are transforming different aspects of people’s lives, including the ways they think, work, learn, and communicate (Jewitt, 2005; Gura & Percy, 2005). Obviously, university students’ lives – their intellectual and everyday social activities are increasingly dependent on, expanded and supported by new communications technologies. Thus, the shifting social and technological landscape in the 21st century suggests that policymakers, faculty, and administrators should develop systematic plans to harness the potential of technology-mediated instruction to support and improve how pre-service teachers are prepared to teach the different school curricula. The Academic Senate of the California State University (2003) defines technology-mediated instruction as “all forms of instruction that are enhanced by or utilize electronic and/or computer-based technology. It specifically includes distance education, instructional modules delivered via mass media, and computer assisted instruction” (AS-2321-96).

The assumption that educational technologies have the potential to revolutionize how pre-service teachers are trained to teach is an important one. For example, Johnson (2006) and Simpson (2006) observe that asynchronous discussion board (ADB) is collaborative and interactive and thus opens new opportunities for pre-service teachers to learn how to teach in innovative ways. The increasing availability of new technologies of teaching and learning seems to suggest that teachers’ preparation to teach this millennial generation may be fundamentally different from previous approaches. For one, the material affordances of new technologies of communications challenge the conventional conception of pedagogical practices, social space, social practices, and schedules (Brewer & Klein, 2006). For example, instruction in different subject areas may no longer be restricted to face-to-face meetings within classrooms or the few hours officially allotted for instruction in university schedules.

However, despite the promise of technology as “powerful pedagogical tools” (National Research Council, 1999, p. 218), many faculty members keep to the face-to-face, instructor-dominated, and knowledge transmitted approach to teaching (Bryant, 2006; Alghazo, 2006). Bryant (2006), Otero, Peressini, Meymaris, Ford, Garvin, Harlow, Reidel, Waite and Mears (2005) and Rogers (2000) observe that many faculty members are reluctant to adopt new technologies as a curricula tool and improved pedagogy because of the doubt that it will improve student learning. Gura and Percy (2005) argue that those who run schools have little understanding of “what is possible with the technology and how to make it happen” (p. iv). Indeed, the National Research Council (1999) pleads with school administrators and instructors for a better understanding and appreciation of technology as a significant tool of instruction:
What has not yet been fully understood is that computer-based technologies can be powerful pedagogical tools – not just rich sources of information, but also extensions of human capabilities and contexts for social interactions supporting learning (p. 218).

This quotation suggests a need to integrate technology as a tool for preparing pre-service teachers. Johnson (2007) and Doering and Beach (2002) contend that technology helps students to construct knowledge. More importantly, Johnson (2007) argues that understanding pre-service teachers’ perceptions of their own learning while using technology will help researchers and teacher educators to gain insights into the connection they make between the theory of using technology for learning to teach and what they actually do in real-life situations. In this regard, Molebash (2004) blames teacher education for “the existing gap between how teachers are expected to use technology and how they are actually using it” (p. 412).

Therefore, the purpose of this study was to investigate the pre-service teachers’ perceptions of asynchronous discussion board as a tool for learning. The following research questions provided the framework for the study:

- What are the pre-service teachers’ views of ADB as a tool for learning to teach?
- What are the participants’ perceptions of the kinds of learning opportunities afforded by asynchronous discussion board instruction?
- What are the participants’ views of using ADB for independent learning needs?
- What are the participants’ views of the contribution of other students to the learning process in ADB?

These are important questions because how students perceive their learning experiences and how they conceptualize their roles in teaching and learning (from students’ perspectives) are important for designing pedagogies and tasks that meet the learning needs and interests of learners (Shellens, van Keer & Valcke, 2005; Ashworth & Lucas, 1998). Shellens, et al. (2005) argue that learners’ perception of a learning task “defines the challenge, its difficulty, and the balance of motivation necessary to address it” (p. 734). Molebash (2004) argues that the perceptions of pre-service teachers influence their teaching of specific subjects and their decisions on whether or not to integrate technologies into them. It seems, therefore, that studies of pre-service teachers’ perceptions will advance researchers’ understanding of how technology can best enhance teaching and learning for teacher educators and pre-service teachers. In addition, there is a need to understand how the use of ADB mediates faculty members’ pedagogical practices, their motivation to use it, and the kind of problems they face when using it.

In the rest of this study, I will address the research question by first discussing the relevance of discussion board in pre-service teacher education. I will also develop a theoretical framework, review related literature, present results and discuss them, discuss summary of findings and implications.

Asynchronous Discussion Board and Pre-service Teachers

Ben-Jacob, Levin and Ben-Jacob (2000) and Rogers (2000) observe that a combination of many factors is producing fundamental and far-reaching transformations in teacher preparation across universities. For example, most universities are faced with dwindling financial resources, competition from other institutions, and the demand for technology integration – all pointing to the need to reform the way they do business. Also, pre-service teachers have other commitments in addition to being students (Ben-Jacob et al. 2000). For instance, all pre-service teachers in this study indicated in their pre-treatment survey that they lived off campus, 87% indicated they had family commitments (as parents, wives/husbands), 94% noted they supported themselves by working full-time or part-time while only 6% indicated they were full-time students. The real question and concern for teacher educators becomes how to better understand “the qualities, habits, skills, attitudes, naturalized and habituated practices” (Kress, 1997, p. xv) that are required of individuals for productive engagements with the world.

Such a new understanding and conception of pre-service teachers should start from an appreciation that technologies such as cell phones, television, computers, video, ipods, personal data assistant, (i.e. blackberry), etc. play a major role in their everyday social interactions. As such, it is reasonable to assume that they have been apprenticed to the use of technology through communities of practice (home, peer groups, work) and can apply technological skills in academic contexts. In this regard, Buckingham (2006) and Ben-Jacob et al. (2000) argue that today’s students are inquisitive, digitally savvy, critical, explorative, manipulative, and nonconforming – challenging and questioning established authorities. Again, one can reasonably assume that students of the digital generation learn differently.
As a result of the confluence of diverse factors discussed in the preceding paragraphs, most universities are adopting course management systems (CMS) software such as Blackboard and Taskstream, and brainstorming software such as Inspiration to create new possibilities and provide innovative pedagogical and learning supports for faculty and students (de Smet, van Keer & Valcke, 2008; Mabrito, 2006). Such technologies offer instructors new opportunities to examine their own pedagogical practices in relation to the fundamental challenge to traditional pre-service teacher preparation courses, the traditional conception of knowledge, and the kind of dispositions that today’s students bring into learning (Buckingham, 2003; Kress, 1997). This suggests a need for researchers and teacher educators to understand how students perceive learning to teach while using technologies.

Fortunately, ADB facilitates new models of teaching and learning where knowledge is networked in the affordances of people, tools, and technologies. In this regard, Gee (2007, 2003) argues that rather than teacher-authority, teacher-domination, and knowledge transmission that characterize lecture-based approach to teaching, the use of technology facilitates critical learning, distributed knowledge, inquisition, discovery, and creativity for learners. In particular, ADB has the potential to build learning conditions that facilitate equitable participation for all students irrespective of their gender, ethnicity, language, and social-cultural backgrounds (Gee, 2007, 2003). This is one of the reasons why de Smet et al. (2008), Brewer and Klein (2006), van Aalst (2006) and Lim and Cheah (2003) suggest that there are clear differences between traditional teaching and ADB instruction. Table 1 provides a brief summary of the differences between lecture-based and ADB instruction.

<table>
<thead>
<tr>
<th>Traditional Lecture-based Instruction</th>
<th>Asynchronous Discussion Board Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teacher Domination</strong></td>
<td><strong>Learner-centeredness</strong></td>
</tr>
<tr>
<td>- authoritative</td>
<td>- self-directed</td>
</tr>
<tr>
<td>- knowledge transmission</td>
<td>- self-motivated</td>
</tr>
<tr>
<td>- linear</td>
<td>- inquisitive</td>
</tr>
<tr>
<td>- more analytical and critical</td>
<td>- innovative</td>
</tr>
<tr>
<td><strong>Teacher’s Role</strong></td>
<td><strong>Teacher’s Role</strong></td>
</tr>
<tr>
<td>- deliverer of content</td>
<td>- facilitator of learning</td>
</tr>
<tr>
<td><strong>Knowledge Construction</strong></td>
<td><strong>Knowledge Construction</strong></td>
</tr>
<tr>
<td>- hierarchical knowledge (from instructor to students)</td>
<td>- Social &amp; interactive learning</td>
</tr>
<tr>
<td>- facts and information-based</td>
<td>- Nonhierarchical</td>
</tr>
<tr>
<td>- tests, practices, &amp; learning are detached from social contexts</td>
<td>- Demands the use of individuals’ skills, knowledge &amp; perspectives</td>
</tr>
<tr>
<td>- teacher role-modeling appropriate learning behavior</td>
<td>- Different types of knowledge: individual, distributed, dispersed</td>
</tr>
<tr>
<td>- quiet, concentrated, reflective analytic activity</td>
<td>- Learning is an active process</td>
</tr>
<tr>
<td>- passive learning is encouraged</td>
<td>- Students learn skills, strategies, knowledge, values, and practices within contexts of learning</td>
</tr>
<tr>
<td><strong>Kinds of Social Relation</strong></td>
<td><strong>Kinds of Social Relation</strong></td>
</tr>
<tr>
<td>- competitive</td>
<td>- dialogue</td>
</tr>
<tr>
<td>- face-to-face</td>
<td>- anonymity</td>
</tr>
<tr>
<td>- hierarchical (instructor first, then bright students, then others)</td>
<td>- facilitates taking multiple identities</td>
</tr>
<tr>
<td>- reliance on visual, proxemic, and verbal cues</td>
<td>- equality and democracy</td>
</tr>
<tr>
<td>- more analytical and critical</td>
<td>- multiple ways of engaging with others</td>
</tr>
<tr>
<td>- passive participation (students may choose not to ask and/or answer questions)</td>
<td>- power sharing</td>
</tr>
<tr>
<td><strong>Participatory Structure</strong></td>
<td><strong>Participatory Structure</strong></td>
</tr>
<tr>
<td>- instructors transfer knowledge</td>
<td>- flexibility (about how and when to participate)</td>
</tr>
<tr>
<td>- students learn from instructors</td>
<td>- creativity</td>
</tr>
<tr>
<td>- less social interaction</td>
<td>- mobility across spaces</td>
</tr>
<tr>
<td>- passive participation (students may choose not to ask and/or answer questions)</td>
<td>- social interaction is required</td>
</tr>
<tr>
<td>- share the same space, may engage in different roles</td>
<td>- share the same space, may engage in different roles</td>
</tr>
</tbody>
</table>
As shown in Table 1, traditional instructional practices tend to be authoritative, linear and knowledge transmitted while ADB instruction seems to promote active learning, inquisitive, self-directed, self-reflective and analytical learning (Schellens & Valcke, 2006). The table further shows the difference in social relations and participatory structures between traditional instruction and ADB. For example, while traditional instruction facilitates competitive, face-to-face, and hierarchical kinds of social relation, ADB promotes dialogue, anonymity, multiple identities, and equality (Gee, 2003). More importantly, the potential of discussion board to support self-reflection, collaboration, and learning anytime and anywhere (Lim & Cheah, 2003) for pre-service teachers is very important in contemporary conception of teaching and learning.

Technology and Learning to Teach: A Theory

Learning afforded by new technologies is increasingly becoming important in pre-service teacher education (Schellens & Valcke, 2006; Lee-Baldwin, 2005; Schellens, van Keer & Valcke, 2005). This is because the sociocultural theory of learning posits that learning is interactive, discursive, technology-mediated, and situated (Schellens et al. 2005; Gee, 2004, 2003). Gee (2003) extends this perspective to suggest that human learning is “fully embedded in (situated within) a material, social, and cultural world” (p. 8) and that the affordances of tools and technologies (e.g. computer and the Internet) can enhance learning.

Gee (2003) further argues that when people learn with human and technological resources, such individuals extend their knowledge and social connections. Gee (2003) suggests a theory of distributed knowledge to explain that learning is not only inherently social, in addition, it is “... distributed, and part and parcel of a network composed of people, tools, technologies, and companies all interconnected together” (p. 177). This implies that learning to teach literacy is social as well as networked in the affordances of what people can do with modern technologies.

Implicit in Gee’s view of literacy as social learning is the pivotal role of knowledge, people, context, and technology – all continuously interacting as central elements to the pragmatics of learning to teach. The implication of Gee’s theory is that these four elements constitute an integrated view of pedagogical practices. While the “knowledge” element emphasizes criticality, i.e. the notion that learning to teach entails reflective, critical thinking and extension of knowledge; the “people” aspect explains social connections and social interactions inherent to learning to teach. Similarly, the “context” explains learning to teach as activities situated in embodied experiences resulting from learners’ social-cultural interactions with the world, while the “technology” element explains how learning to teach entails learning of social practices and networking capabilities associated with specific communications technologies and tools. Gee’s (2003) theory seems to suggest that the issue of preparing teachers to teach in this digital age entails knowing “what people can think and do with others and with various tools and technologies” (p. 184).

Gee’s theory of distributed knowledge applies to learning to teach as conceived in this study. Gee (2007, 2003) provides detailed description of distributed knowledge. Specific aspects of the theory as applied to learning to teach through the resources of ADB are briefly paraphrased below:

- Individual students’ knowledge is not only social; it is distributed across their peers, contexts, and technologies.
- Knowledge and skill acquisition resides in the importance of networking the affordances of interconnectedness of people, tools and technologies.
- Learning is extending knowledge, social connections and interactions.
- Learning is a change in identity as knowledge is learning the process of participation, and it is a process in which participants and practices change.
- Learning is situated practices as participants constantly learn new roles in different configurations and contexts.
- Learning takes place as students participate in scaffolded joint activities with other people and their associated tools and technologies.
- Learning promotes formation of groups of learners with common interests as members combine all their resources – sociotechnical expertise, sociocultural affiliations, tools, and technologies to accomplish common goals.

In summary, learning to teach requires a new way of learning by pre-service teachers. This new way requires teacher educators to use technology to engage pre-service teachers in learning (Schellens et al. 2005). It also requires pre-service teachers to construct knowledge through active participation in discussions and knowledge sharing with their
peers (Schellens et al. 2005). Schellens et al. (2005) aptly argue that ADB facilitates active engagement in learning processes as students collaboratively work on a learning task, and by “mutually explaining the learning content, giving feedback to other group members, asking and answering questions” (715). In this way, the dialogue they engage in (through the use of discussion board) provides an understanding of “the nexus between how students understand a phenomenon belonging to a learning task and what they actually do in undertaking the task” (Johnson, 2007: p. 1).

**Review of Related Literature**

ADB is increasingly becoming a powerful tool of creating new possibilities and providing innovative pedagogical and learning supports for teacher educators and pre-service teachers (Simpson, 2006; Lim & Cheah, 2003; Doering, Johnson & Dexter, 2003). Lim and Cheah (2003) explore pre-service teachers’ perception of the role of their tutors in using ADB to teach education courses. Lim’s and Cheah’s (2003) findings suggest that there is a significant difference between the students’ “experiences and perceptions of the roles of the tutors” (p. 43). They recommend that tutors should provide specific guidelines for their roles in ADB.

Johnson (2007) explores students’ perceptions of how they construct knowledge while using ADB. The study is premised on the argument that teacher’s understanding of students’ perceptions of their learning while using ADB provides an understanding of the link between how students understand learning tasks and what they actually do in undertaking the task. Johnson (2006) argues that students experience learning events independently when they use ADB because each learner can work at his or her own pace. Johnson (2007) observes that the variation in learning is important for teacher educators to take into consideration when designing learning tasks for their students in ADB. He contends that since students learn differently, assessment criteria should be built around learning processes (the construction of knowledge) instead of factual knowledge. Johnson (2007) emphasizes the need to “stress the context as one that necessarily requires careful consideration with regards to assessment” (p. 11) in ADB learning environments.

Brewer and Klein (2006) investigate the effect of rewards, roles and affiliation motives in asynchronous learning for undergraduate students. Their findings show that students learn better when the design of ADB provides an opportunity for students to work collaboratively and interact through asking and answering questions within groups. Similarly, Schellens and Valcke (2006) examine whether collaborative learning in ADB enhances learning academic discourse and knowledge construction. Schellens and Valcke (2006) conclude that there are “significant increases in the cognitive interaction, task-orientation and higher phases of knowledge construction” (p. 349) when students use ADB. The two studies suggest a need for educators and instructional designers to provide structures that enhance cognitive interaction among participants.

In their exploration of the use of ADB to support pre-service teachers, Doering, Johnson, and Dexter (2003) argue that participants gain teaching experience and a better opportunity to interact with their students. They further posit that the participants develop important skills in writing of different genres and communicating through multimedia. Simpson (2006) examines the use of ADB to provide graduate students access to a conventional course. The findings suggest that his students consider the technology as an important instrument of learning. He calls for further studies to confirm the benefits of using ADB for course delivery. Schellens, van Keer and Valcke (2005) investigate the impact of learning using ADB on students’ degree of knowledge construction. The study focuses on the effect of students, groups, and task characteristics on one hand and on the other, the assignment of roles to group members. It concludes that there is no conclusive evidence that role assignments do in general result in a positive effect on knowledge construction.

Similarly, Biesenbach-Lucas (2003) investigates students’ perceptions of the efficacy of ADB as a teaching tool in teacher training courses. His findings are also mixed. On one hand, students in this study suggest that discussion board offers “greater social interaction with other class members” (Biesenbach-Lucas, 2003, p. 24) and promote understanding of course content. Conversely, non-native speakers of English note that ADB fails to provide opportunities for additional language practice. Biesenbach-Lucas (2003) summarizes his findings: “For all students, the two main issues perceived as negative related to their perceptions of forced, unnatural interaction promoted by the asynchronous discussions and lack of topic prompts, the requirement to make connections to prior postings, and the frequency of required contributions to discussions” (p. 24).
In his contribution to the study of ADB, Lee-Baldwin (2005) explores the potential of the technology to facilitate reflective thinking among pre-service teachers. The findings suggest that discussion board tends to lend itself to peer-scaffolding and thus promotes social dynamics within groups as an important tool of facilitating cognitively higher levels of reflective thinking. However, the study notes that while ADB has the potential to facilitate reflective thinking among pre-service teachers, tapping into the potential depends on “an array of explicit and implicit factors involved with the complexities of teaching and learning” (Lee-Baldwin, 2005, p. 110).

The literature review above seems to suggest that ADB has promises and challenges as a tool of improving how pre-service teacher preparation courses are taught and learned. This seems to suggest further studies to better understand pre-service teachers’ perspectives of how they use ADB to mediate learning to teach.

Methodology

Participants: a total of 33 pre-service teachers participated in this study. While 21 of the participants were in the elementary education program and enrolled in Skills in Teaching Reading to Bilingual Elementary Students, the remaining 12 were in the secondary education program and enrolled in Skills in Teaching Reading in Secondary School. The elementary education candidates consisted of 19 (90.48%) females and two (9.52%) males, while secondary school education candidates had five (41.67%) females and seven (58.33%) males. The age range of the students varied from 22 to 47. In addition, fourteen (42%) worked as substitute teachers, three (9%) as full-time teachers and two (6%) as teacher aides. Also, two (6%) were full-time students and 12 (36%) worked in offices.

The Courses: Skills in Teaching Reading to Bilingual Elementary Students is the first of two literacy courses that the students are required to take. The course was designed to provide a variety of language arts experiences for the candidates to learn including such concepts as multiliteracies/multimodality; different approaches to reading instruction; early (home) literacy practices and transition to school literacy: phonemic awareness and activities, instruction and assessment; word recognition, word decoding; reading fluency; vocabulary and concept development; and reading comprehension. Others included spelling, reading-writing connections, and identifying literacy needs of diverse learners.

Similarly, Skills in Teaching Reading in Secondary School was the first of two literacy courses the secondary education students need to take. The primary focus of the course is to provide the students’ experiences regarding theoretical foundation and practical application of literacy instruction in content areas. Specific instructional strategies and concepts that were discussed included the following:

- the notion of literacy: traditional, multiliteracies/multimodality;
- literacy across curriculum: definition, the role of teachers, theories of content reading, etc.;
- creating a positive learning environment: activating and/or building background experiences, understanding the context of students lives, linking literacy content with students’ out-of-class social and cultural experiences;
- meaning-making strategies, vocabulary teaching strategies, vocabulary learning activities;
- comprehension strategies: summarizing, clarifying, questioning, predicting;
- literacy assessment: types of assessment, learning about students, assessing textbooks; and
- issues in literacy teaching: language learning, diversity, culture.

The university, the site of this study, provides faculty members and students access to a course management system (CMS). The system consists of (a) Blackboard with features such as voice annotation, virtual classroom, and discussion board and (b) Taskstream with rubric capabilities. ADB is a web application for holding discussions and user-generated content (Wikipedia, n.d.). Discussions are grouped in threads that contain a main posting and all related replies. For example, when a student posts a question, it appears in the main thread and subsequent responses will be indented under the thread. Typically, each posting may have multiple indented threads as responses to the original question. Students can post questions and responses at any time. There is, therefore, no time constraint on users. Also, students can navigate the postings in nonlinear order (Mabrito, 2006). In this way, it allows students to deliberate, reflect, and simultaneously utilize other resources (texts, tools, people, and other technologies – websites), thus promoting active and critical learning (Biesenbach-Lucas, 2004). Also, it allows students, when used regularly to develop a sense of virtual community (Wikipedia, n.d.).
The integration of discussion board into the two courses was designed to facilitate an out-of-class engagement with course content as well as provide the participants a space for a social relation and collaboration. In both courses, 100% of teaching was done in lecture rooms during the university assigned time. The course syllabus states: “This assignment requires each student to reflect on the teaching/learning in this course on a weekly basis. Identify one issue that you will like to explore further or one area you need further explanation. Then develop this into a question or comment and post it on the discussion board for comments from your course mates. Read other students’ postings and make sure you post a response to one question.”

Instruction was delivered through a combination of different strategies – including hand on activities, in-class discussions, PowerPoint presentations, and web discussions. Learning activities involved the use of different technologies such as Taskstream (for students to prepare lesson plans for microteaching), discussion board (for reflection, and posting questions and answers – as describe above), and websites (listed in the syllabus, for reading specific articles to supplement classroom instruction). The range of functions used in the board is tabulated in Table 2 below:

<table>
<thead>
<tr>
<th>Discussion Board Facility</th>
<th>Utilization</th>
<th>Student Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Announcements</td>
<td>• Provide information</td>
<td>• Students generate questions</td>
</tr>
<tr>
<td>• Whole-class conference</td>
<td>• Provides instructions</td>
<td>• Students respond to questions</td>
</tr>
<tr>
<td>• Record of students’ work</td>
<td>• Post course materials</td>
<td>• Students read other’s responses</td>
</tr>
<tr>
<td></td>
<td>• Post assignments</td>
<td>• Students work collaboratively</td>
</tr>
<tr>
<td></td>
<td>• Write comments</td>
<td>• Students work independently</td>
</tr>
<tr>
<td></td>
<td>• Student-student discussion</td>
<td>• Students upload documents</td>
</tr>
<tr>
<td></td>
<td>• Teacher-student discussion</td>
<td>• Provide links to websites</td>
</tr>
<tr>
<td></td>
<td>• Group or independent work</td>
<td>• Source information from texts</td>
</tr>
<tr>
<td></td>
<td>• Shared material (content</td>
<td></td>
</tr>
<tr>
<td></td>
<td>available to all students)</td>
<td></td>
</tr>
</tbody>
</table>

As indicated in Table 2, ADB allowed the participants to generate questions and responses to questions, read other students’ responses, work collaboratively and independently, and provided links to different websites and sourced information from peers, lecture notes and textbooks. In addition, it allowed the instructor to post announcements, syllabi, course assignments, and materials; write comments, and engage in student-student and student-instructor discussions. In particular, it allowed for shared material as all comments, questions and responses posted were available to all the participants.

Classroom Procedure: This study ran for a semester – 16 weeks. The course instructor (and the author) prepared the students to use discussion board by first introducing them to the software, its capabilities and how its range of functions were specifically adapted for the purpose of the classes (see Table 2). During the subsequent weeks, all students logged onto ADB to reflect and discuss their perceptions of the topics covered during lectures. Each student (a) posted a question about topics covered in the class, (b) posted questions about topics of interest not sufficiently covered or not covered at all, (c) posted a response (five sentences) to another student’s question, (d) postings addressed course related topics, and (e) postings discussed and reflected critically on their perceptions of learning. The structured nature of the assignment was to facilitate a productive social interaction by ensuring that the participants engage in in-depth discussions rather than random postings. On weekly basis, the course instructor read the postings and provided comments. The instructor’s comments were designed to provide additional ideas and concepts on different topics. All the threads were available to the students to view throughout the semester. The first 15 minutes of each lecture were devoted to addressing such questions and related issues of working on such a space. Students also printed out and submitted hard copies of their weekly entries (questions posed and questions answered) for grading and recording of their grades.

Instrument for Data Collection: Three instruments were used to collect data for the study. They were: (a) students’ postings on the discussion board: this included the weekly posted questions, answers and comments for the semester,
(b) oral interview: students were interviewed regarding their experiences using discussion board, and (c) written reflection: each student wrote a two-page essay of their perceptions regarding the impact of discussion board as a tool of learning to teach. Assignments pertaining to discussion board carried 40 points out of 200.

Interviews with the participants took place in lecture rooms during the last week of the semester. The interview questions were based on the theoretical framework proposed by Gee (2007, 2004, 2003); Davies (2006); Karppinen (2005); Hewson and Hughes (2005); and Rogers (2000) that tools and technologies mediate learning and that technology is a “product and process of socially dynamic relations” (Davies, 2006, p. 219) that promotes:

- active learning by promoting reflection;
- learning as a social practice involving shared practices, interaction, and collaboration;
- dialogue between students and between students and faculty;
- freedom, autonomy, and flexibility;
- diverse types of learning styles; and
- different types of knowledge.

During the interview sessions, I asked the following specific questions: (a) How did the discussion board help you learn in this course? (b) What are your views of using ADB as a tool of learning to teach? (c) How did the discussion board allow you to tap into other resources to help you learn in this course? (d) What kind of unexpected skills and knowledge did you acquire by using the discussion board? (e) What kind of social practices did you develop by using this technology? (f) How did discussion board motivate you to learn in this class? (g) In what ways have you changed regarding how you think of teaching/learning in relation to technologies?

To establish content validity of the questions used for oral interview, the draft of the questions was given to two professors for their comments. Their comments and suggestions were used to refine the instrument.

Data Analysis: The content of the participants’ postings and responses to interview questions were analyzed within the framework of verbal analysis method developed by Chi (1997). The goal of the verbal analysis method is to analyze students’ verbal utterances to capture the knowledge underlying them, and in the process, to understand what a student knows and how that knowledge shapes the way she or he reasons and solves problems (Chi 1997). In this approach, students’ utterances are segmented based on semantic features such as ideas, concepts, argument chains, and topics of discussion.

The participants’ responses were segmented and then coded according to the meanings they expressed. Broadly, the meanings were coded as (a) Knowledge construction – such as when the participants asked for clarification of ideas, define or explain ideas, concepts, and so on. (b) Interactivity – when the participants’ responses confirmed, appreciated, challenged or learned from others’ ideas or positions. (c) Development and growth – in cases when they expressed learning diverse skills such as linking to websites, uploading, and downloading texts, printing, etc. (d) Independent learning – such as when the participants used the technology to design learning experiences suitable to their own conditions, including scheduling and selection of questions to respond.

Results

This study examined the pre-service teachers’ perceptions of ADB as a tool for learning to teach. Four specific questions were raised for investigation at the beginning of this study. Therefore, results were first presented according to the issues and this was followed by the discussion of the general themes that emerged from data analysis. The verbal data that were quoted for illustrating the specific themes were “typical” of the participants’ responses to such categories. The participants posted a total of 491 questions and 491 answers in addition to the interviews.

What are the pre-service teachers’ views of ADB as a tool of learning to teach?

ADB allowed the participants to learn different ideas and create different types of knowledge. In particular, the discussion board allowed the pre-service teachers to participate and contribute to discussions of different topics unlike what typically happened in traditional face-to-face lectures where vocal students tended to dominate. This is
because the technology allowed the participants to freely reflect, critique, and extend knowledge. Some of the participants, in their reflective essays, wrote:

- “What are some effective strategies for teaching reading fluency, particularly when you have English language learners in your class?”
- “What are some resources (extra materials), other than the ones provided by the school sites, that we can use to improve students’ literacy practices?”
- “Last week we discussed paired reading. What is paired reading? Is this strategy limited to language arts or can we use it to teach other school subjects?”
- “Do you think standardized testing scores should be the only source of criteria for judging students’ success or failure in literacy?”

In their reflective essays, the participants identified specific kinds of learning activities they viewed as useful. For example, 27 students (82%) noted that using the discussion board “pushed” them to read over their lecture notes, textbooks and reflect on class discussions before posting their questions and answers. In addition, 25 participants (76%) wrote that they learned from other students who had practical experiences of teaching because of their backgrounds as full-time or substitute teachers. Also, 25 (76%) suggested that the use of the discussion board provided them a chance to ask questions about topics they considered important but were not covered in class. Furthermore, 23 students (70%) noted the activity forced them to think of how to use technology in their own classrooms when they become teachers.

During the follow-up interviews, I asked the participants to explain how the discussion board helped them to learning the course materials. The following are some of their responses:

- “When I saw questions that were interesting, I usually went back to read my notes and even went back to the chapter [of the topic] in my textbook [and read] in order to post a good answer.”
- “The links to different websites provided by my course mates helped me to access very important information online. Clicking, surfing, and connecting with other sites are fun for me; it makes learning interesting.”

**What are the participants’ perceptions of the kinds of learning opportunities afforded by ADB instruction?**

Buckingham (2006) argues that technology-mediated instruction affords learners innovative, immediate, interactive, and investigative approaches to learning. This is because learning activities are geared toward helping students to access, explore, process, and apply information in ways that are radically different from lecture-based pedagogies. For example, the students in this study understood the need to learn the features of the software, how to manipulate it, how to extend their knowledge, and how the interactions associated with the discussion board as a social space. Here are some selected samples of the students’ written reflections to the kinds of learning opportunities that they perceived the technology afforded them.

- “Students are not only sharing with their peers their questions and comments, but they are also learning about topics that they never thought about before.”
- “You can print a particularly interesting or relevant discussion, because threads of discussions are generally archived.”
- “I had forgotten how to use the discussion board . . . I [later] became an expert in using it. I also learned to use several other features having to do with discussion board, as I had to explore the web page I was using.”

Furthermore, their written reflections indicated that 31 (94%) students learned to manipulate the functions of the discussion board to print, upload and save documents, make connections or links with other websites, integrate visual images, use different font sizes, and request help from other students or the course professor when they ran into problems. During the follow-up interviews, the students expressed ideas such as:

- “what I really like in discussion board is that one is learning and doing many things all at the same time. You can use your notes, textbooks, other postings or even link with other websites as you prepare your questions or responses”

**What are the participants’ views of using the discussion board for independent learning needs?**
Learning in ADB promotes freedom and autonomy. This is important as students learn differently. Learning in discussion board therefore, means different experiences for different people. The pre-service teachers in this study used ADB to meet their own unique learning needs as shown in their reflective essays:

- “I like to wait till the last minute to post my questions and answers as I am a big procrastinator. The good thing is that the discussion board still gave the chance to complete my assignments.”
- “I never had the time to read all the questions and answers posted each week by each student, but it was up to me to go back and read the questions and answers some other day.”
- “Many times, I used the links posted by peers to connect with other sites to source additional materials that helped me to post my own answers to questions posted by my classmates.”

During the follow-up interview, the participants expressed positive views regarding the use of ADB:

- “Since a discussion on the Blackboard is done behind a computer and not face-to-face, there is less intimidation of expressing myself compared to sharing my concerns during a whole class discussion.”
- “Discussion board was a great way to read people’s different perspectives or ideas about assigned topics. Many times . . . I noticed that my ideas were different from others which was great because it helped me think of things differently which made me grow as a student and an educator.”

What are the participants’ views of the contribution of other students to the learning process in ADB?

Students’ perceptions of the value of their interactions with others in ADB significantly influenced their interests in using the technology as a learning tool. This means that students need to perceive ADB as a technology that does not only provide quality postings but serves as a social space for mutual respect and trust. The participants had positive views of the contributions of their peers because despite the variety of activities, they shared a common endeavor in the sense that all activities related in one way or the other to learning to teach. The students wrote following in their reflective essays:

- “I like discussion board because it allowed students to work with, contribute to one another and share thoughts and ideas freely.”
- “Another thing I learned from the discussion board was to respect others’ thinking. When I read questions from classmates, I understand them. I appreciated their beliefs and feelings towards education.”
- “I feel that the different topics addressed in the discussion board showed the maturity and responsibility of some of the classmates, as well as, their interests in learning more strategies and other ways of helping the students we work with.”

Data from the participants written reflections suggested that they had positive views of the contributions of their peers. For example, while 28 (85%) respondents considered the postings as helpful for them; 26 (79%) noted that their peers’ postings helped them to clarify what they learned during previous lessons. Also, 23 (70%) of the participants wrote that the dialogues helped them to expand their perspectives, and 22 (67%) noted that the postings encouraged them to be more self-reflective before postings their questions and answers. During the follow-up interviews, the participants expressed positive views regarding the contribution of their peers:

- “I used to think that the teacher was the only one who has all the right answers, but now I know that I can also learn from my classmates. Even though most of us are working in the school setting, we have different experiences that make us know or learn differently.”
- “Usually, when you attend classes at a university, you do not get a chance to talk to many people from your class. Even though I did not talk to some of my classmates face-to-face, I talk to them through the discussion board. . . . I connected with many of the peoples’ questions and answers. It helped me think I was not alone in my ideas.”

Discussion

To summarize the thematic discussion in this study, pre-service teachers perceived ADB as a tool for fostering active participation in learning to teach processes and a means of engagement in different kinds of learning opportunities. In addition, they viewed the technology as a resource that helped them to meet their specific learning needs. They also perceived the contributions of their peers as positive while they use ADB to learn. These themes will be further explored.
The pre-service teachers perceived ADB as a tool of learning

The pre-service teachers in this study tended to perceive ADB as a learning tool with the potential to foster knowledge construction. For example, data from their postings, essays and interviews showed that they focused on issues of immediate concern to their interests – how to solve real-life classroom problems. They asked questions about supplementary reading materials and strategies for teaching literacy. Furthermore, their answers provided varied discussions and perspectives about different topics. In many instances, they provided links to related websites for participants who needed additional readings on different topics. For example, a student, in her written reflective essay noted that the discussion board did not only afford her “extended class sessions,” in addition, it provided “very enriching activities” because the course mates posted “varied and useful” questions and answers.

The participants’ reflective essays and interviews seemed to indicate that they understood that knowledge construction resided in the networks of relationships created by members of a learning community (Gee, 2003). First, individual members brought with them some specialist knowledge as students of teacher education to the discussion board learning endeavor. In addition, some brought other kinds of knowledge as full-time teachers, substitute teachers, teacher aides, etc. that contributed to the groups’ pool of knowledge. Also, the participants’ reflective essays, interviews and postings showed that they enhanced their understandings of concepts, ideas and topics by asking questions of one another, articulating their own views and experiences, and gaining new insights from the experiences and perspectives of others.

Johnson (2007) argues that the wide range of questions and answers posted by students while using ADB can help them to integrate different learning activities such as analyzing and synthesizing ideas from their peers’ postings, textbooks, and additional sources such as websites. This kind of shared, varied and rich knowledge built around peers, texts, and technology afforded by ADB is difficult to attain in traditional face-to-face lectures settings. Wijekumar and Spielvogel (2006) conclude that ADB is a powerful resource that “scaffolds learners into contributing new ideas and managing their learning by providing feedback on their paraphrasing and focusing on the relevance of the posting” (p. 231). Also, Biesenbach-Lucas (2003) argues that ADB provides students with a realistic audience (their peers) and a real purpose for writing.

Pre-service teachers use ADB to foster situated learning

Gee (2003) observes that the principle of self-knowledge requires that students learn not only about course content and context of learning, but also “about themselves and their current and potential capacities” (p. 208). The participants in this study perceived ADB as a technology that promoted learning as situated in their embodied experiences (e.g. experiences of learning to use ADB to learn) and a space for growth. For example, some students initially had problems – while some made mistakes, others had forgotten how the different features of the discussion board related to each other and how the different components interactively and independently functioned to afford learning. But through practice and persistence, they learned new skills – printing, accessing and exploring webs, uploading/downloading texts, etc.

In particular, the students’ responses seemed to suggest that they “link new information with existing and future-oriented knowledge in uniquely meaningful ways” (The American Psychological Association, cited in Kayler & Weller, 2007, p. 141). The participants learned to take risks where real-life consequences were little. They could start, stop, or continue at any time; they were not discouraged by the initial “failures.” In addition, discussion board appeared to afford the participants ample opportunities to practice, experiment, make mistakes, learn, and make discoveries. Also, the students learned the basic skills to manipulate the software by actually using it. Lastly, they learned to use the discussion board by thinking of it at a “meta” level as a complex system of interrelated and interconnected parts that could be used to read others’ views, connect to websites, upload, and print. This showed that the participants perceived ADB as a technology they could use to foster self-knowledge – the kind of understanding that learners gained about the technology they are using while at the same time they acquire new knowledge about themselves, their present and potential capacities (Gee 2003). This kind of knowledge is hardly made available to students by the old-style lecture-based instruction.
The participants used ADB to approach learning in ways that suited their needs and conditions

An important finding in this study was that ADB allowed pre-service teachers to approach learning in unique ways; to customize learning experiences to align with their unique conditions and expertise. This finding is important as available literature (Brewer & Klein, 2006; van Aalst, 2006; Lee-Baldwin, 2005) tends to stress only the social aspect of learning in ADB. However, the pre-service teachers learned independently by sourcing additional materials on their own, making their own schedules to use ADB, deciding what questions were important to ask and respond to. More importantly, the participants perceived that what they learned from their peers’ perspectives and experiences created an incentive for them to reflect on their own ideas and, in the process, gained important insights into teaching and learning.

Kayler and Weller (2007) argue that when students use ADB, they write about their unique experiences and develop their own voices and perspectives. This suggests that freedom and autonomy are critical in using ADB as Schellens et al. (2005) observe that learners do better when there is “freedom for students to define their own problems to be solved, rather than a teacher presenting the students with rigidly designed problems” (p. 734). In a crucial way, this study shows that ADB afforded pre-service teachers multiple ways of learning that will encourage them to “make choices, rely on their own strengths and styles of learning and problem solving, while also exploring alternative styles” (Gee, 2003, p. 209).

The pre-service teacher viewed the contributions of their peers positively

Pre-service teachers’ perceptions of the kinds of social interaction surrounding the use of ADB have the potential to shape the principles by which knowledge, values, practices, skills, and teaching strategies are negotiated, learned and applied to teaching/learning activities in real-life situations (Johnson, 2006). Data from the participants suggested that they created a learning community built on mutual trust and respect, and in addition, valued and appreciated each others’ contributions. For instance, the students’ written reflections and interviews suggested that they respected their peers’ perspectives, experiences, maturity, and sense of responsibility. Johnson (2006) argues that ADB facilitates student learning due to the cognitive processing required in writing postings, time to reflect on their peers’ postings, and think over their own responses. Similarly, Biesenbach-Lucas (2004) observes that ADB provides students not only the opportunity to discuss content and teaching and learning experiences outside the classroom; also it allows them to expand their perspectives on course readings and class discussions.

Brewer and Klein (2006) argue that meaningful interaction between students is an important component of ADB. From a constructive perspective, Kayler and Weller (2007) argue that students construct knowledge as they work in the company of their peers. Wijekumar and Spielvogel (2006) posit that ADB affords the opportunity for students to actively engage in the construction and use of knowledge. Kayler and Weller (2007) put it aptly: “The social construction of knowledge embedded in dialogue creates new opportunities for self-reflection, growth, and intrinsic motivation for belonging . . .” (p. 141). More importantly, such dialogues and interactions happen in a “free” social space (ADB) where learners can afford to take risks with lowered real-life consequences as they share ideas, perspectives, agree and disagree (Biesenbach-Lucas, 2004; Gee, 2003).

Implications and Conclusion

The purpose of this study was to investigate the pre-service teachers’ perceptions of how ADB mediated learning of teaching. Data from this study indicated that the participants perceived ADB as important tool for sharing ideas, experiences, and perspectives in ways that facilitated a social construction of knowledge. This is an important step in learning to teach. Kayler and Weller (2007) argue that the kind of social construction of knowledge afforded by ADB creates new opportunities for pre-service teachers to engage in self-reflective practices, development and self-motivation.

This study, also, indicated that the participants perceived ADB as a technology that afforded learning as situated experiences because they had the opportunity to manipulate the technology and apply knowledge and skills in contexts of using it. They learned to navigate their peers’ postings, upload, download, print, and link with websites – thus extending their knowledge and social interactions associated with the discussion board as a social space.
Furthermore, the study demonstrated that the participants perceived ADB as a technology that afforded them the opportunity to design their own learning experiences to suit their unique situations. They seemed to appreciate the freedom and autonomy associated with ADB. Lastly, the pre-service teachers perceived the contributions of their peers as essential to learning.

The findings in this study have some implications. Universities and faculty should realize that technology has become “an external extension of human intelligence” (Gura & Percy, 2005, p. v) with a potential to support pre-service teachers’ active exploration and manipulation of learning materials, and create contexts for social interactions that support knowledge construction. In particular, after classroom lecture online interaction provides learners a community of practice with the potential for “cognitive and affective development and opportunities for growth as independent learners” (Kayler & Weller, 2007, p. 145). In this regard, this paper calls for approaches that integrate ADB into pre-service teacher preparation in ways that increase the quality of instruction and learning. Such approaches should raise some fundamental questions, including: (a) What is the impact of ADB on the way pre-service teachers learn to teach? (b) What are the students’ perceptions of their own learning in ADB? (c) How do they want to experience learning? (d) How do teacher educators want them to experience learning? (e) What is the nature of ADB in relation to the learning needs of students? (f) How can assignments and reading materials be most appropriately prepared to meet the learning needs of pre-service teachers working with ADB?

These questions will help teacher educators to see the world from their students’ perspectives and prepare themselves to understand variations in ways pre-service teachers perceive their own learning in the context of ADB (Johnson, 2007). In this way, individual teacher educators will be better prepared to develop ADB as a social space for learning as well as nurture students’ own individual learning pathways (Johnson, 2007; Gee, 2003). This is important as Ashworth and Lucas (1998) contend that research and pedagogy must “be sensitive to the individuality of conceptions of the world – it must be grounded in the lived experiences of its research participants” (p. 417).

References


A Capstone Experience for Preservice Teachers: Building a Web-Based Portfolio

Leah M. Herner-Patnode and Hea-Jin Lee
The Ohio State University at Lima, Lima, OH 45804, USA // herner-patnode.1@osu.edu // lee.1129@osu.edu

ABSTRACT
This study presents the results of the examination of the use of technology to enhance the capstone portfolio process for teacher candidates completing a master's degree at a small regional campus in Northwest Ohio, United States. Students at this institution complete a fifth year program that results in a master's degree and a recommendation for an Early or Middle Childhood license. The study focuses on determining the effectiveness of using a web-based portfolio. Teacher candidates' perceived knowledge, skills and dispositions about technology literacy and usage in developing an e-portfolio are measured by comparing a pre and post survey. Results indicate a stronger knowledge and skill base concerning technology use and state standards, as well as a better disposition towards utilizing technology for creating and maintaining portfolios.

Keywords
E-portfolio, web-based portfolio, technology, teacher preparation

Technology That Improves Teaching and Learning
Technology use in higher education has increased as instructors see the opportunity to increase the effectiveness of subject matter delivery along with increased flexibility for teacher candidates and instructors (Blicblau, 2006). Technology can be an effective tool to help the learner more fully understand the target knowledge, and develop higher order thinking skills and problem solving strategies (Fletcher, 2001; Jonassen, 2001). Recent technologies such as Google docs and the concept of Web 2.0 can help to facilitate greater communication of candidates' skills and talents (Henke, 2007; McPherson, 2007). The integration of technology specific to teacher preparation has also expanded in recent years (Alexander & Golja, 2007; Wilhelm & Confrey, 2005). Expertise and use of technology by instructors at the higher education level has improved with university wide support (Herner, Karayan, Love, & McKean, 2003). This has set the stage for ever more effective uses of technology, specifically web-based tools.

The Use of Web-based Portfolios
The use of electronic portfolios (particularly web-based) for teacher preparation is recent, but they do build on the foundation of traditional paper portfolios. Portfolios generally include a resume, philosophy of education, references, letters of recommendation, reflections on educational theories, personal goals, examples of lesson plans, and unit plans (Aschermann, 1999; Chappell & Schermerhorn, 1999; Ryan, Cole & Mathies, 1997; Wiedmer, 1998). The definitions of portfolios are numerous. They have been defined as a purposeful collection of student work assembled to demonstrate progress and achievement (Barrett, 1999; Bull, Montgomery, Overton & Kimball, 1999; Herman & Morrell, 1999; Tuttle, 1997; Wilcox, 1997). Others believe that portfolios offer the opportunity for multidimensional assessment (Backer, 1997; Cole, Tomlin, Ryan & Sutton, 1999; Jacobsen & Mueller, 1998; Riggby, 1995). They also may be used to enhance teacher preparation course instruction (Corbett-Perez & Dorman, 1999; Moehn, 1997; Purves, 1996; Watkins, 1996). Portfolio creation involves the participant in active learning activities such as problem solving, writing, analyzing and researching (MacDonald, Liu, Lowell, Tsai, & Lohr, 2004). Electronic portfolios also have the possibility to create a program that is more highly visible, with a web presence that aids in student recruitment (Reardon, Lumsdon, & Meyer, 2005).

Web-based Portfolios in Teacher Education
Portfolio content that is stored in a web-based course management database, such as WebCT™, Blackboard™ (In February 2006 WebCT and Blackboard merged) or Desire2Learn™, may be presented to the reader in a variety of formats depending on purpose (Banister, Vannatta, & Ross, 2006). To borrow terminology from the information systems discipline, the user can have a view of the data that fits the user’s particular need (Herner, Karayan, Love, &
Instructors will see artifacts created by a teacher candidate as the program progresses and use them to help form the candidate's development. If the portfolio is maintained in one location throughout the course of studies a final committee reviewing a teacher candidate’s mastery of standards as the candidate nears graduation may see cumulative and summative artifacts and a final reflection created by the candidate. Teacher candidates can create a matrix (see Figure 1) that will assist those responsible for accreditation reporting to see specific course assignments and candidate artifacts produced in response to the assignments that address each standard. As a result, sequencing of assignments and curriculum content to develop the candidate’s mastery of the standard can be evaluated. This is especially important as teacher education becomes more student-centered, and accrediting bodies look for specific evidence of growth (Chambers & Wickersham, 2007). The reasons for using a web-based versus a paper portfolio are numerous. The ease of access for numerous reviewers, the ability to provide graphics and weblinks, as well as learning to model authentic evidence of student progress using this global medium are useful, but one of the most important reasons to use the web-based technology is to learn to problem solve in our technologically driven world (Sanders, 2000).

<table>
<thead>
<tr>
<th>Portfolio Artifacts</th>
<th>Standard I (Domain A)</th>
<th>Standard II (Domain B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I-1</td>
<td>I-2</td>
</tr>
<tr>
<td>July</td>
<td>!</td>
<td>X</td>
</tr>
<tr>
<td>FISH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>August</td>
<td>!</td>
<td>X</td>
</tr>
<tr>
<td>Diversity/Exceptionality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>September</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>September</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Courage to Teach</td>
<td></td>
<td></td>
</tr>
<tr>
<td>October</td>
<td>!</td>
<td>X</td>
</tr>
<tr>
<td>Pedagogy Survey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>November</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Autumn Leaves</td>
<td></td>
<td></td>
</tr>
<tr>
<td>December</td>
<td>!</td>
<td></td>
</tr>
<tr>
<td>Midterm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>January</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Project Wild</td>
<td></td>
<td></td>
</tr>
<tr>
<td>February</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Problem-Posing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lesson</td>
<td></td>
<td></td>
</tr>
<tr>
<td>March</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Opportunity Cost</td>
<td></td>
<td></td>
</tr>
<tr>
<td>April</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ant Story</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Inventions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

X Standard Best met
! Standard met

*Figure 1: Sample Matrix*

**Advancing the Portfolio Process Through the Use of Carmen**

A small regional campus in Ohio, United States, has a five year program for general education teacher preparation that results in a master’s degree and recommendation for an Early (grades Kindergarten- 3rd grade) or Middle Childhood (grades 4th-9th) license certification. During the year of graduate studies teacher candidates participate in a year long series of Capstone seminars culminating in a final portfolio presentation. This study sought to compare the beginning knowledge, skills and dispositions of a group of teacher candidates that used a technology based system for Capstone work, as well as portfolio artifact gathering, with their perceived knowledge, skills and dispositions at the end of the program. Carmen (the name the university gives to their version of Desire2Learn™)
served as the structured technology tool. One of the major goals of technology use was to better mentor master’s students through their graduate work, so that upon completion of the program they show a stronger positive disposition towards using technology and reflecting on their teaching skills, as well as a stronger knowledge base concerning technology skills and state standards.

Comparing Carmen to Other Systems

Carmen is a Desire2Learn™ (http://www.desire2learn.com) product. The instructors have used it for course delivery and storage of information. It is a flexible system that allows for data to be stored throughout the course of the Capstone experience, which lasts one calendar year. It is then available to the teacher candidates for 30 weeks after the end of the course. The collaboration tools it features are discussion boards, group discussion areas, dropboxes and lockers (storage space for group projects). Carmen (http://telr.osu.edu/carmen/) compares well to similar products, such as Blackboard™. Blackboard™ (http://www.blackboard.com/us/index.Bb) allows the user to share files as well as create an e-portfolio with extra software. This system incorporates many of the same functions of Carmen. They both allow for threaded discussions, which can be assigned a grade. In terms of sharing files, Carmen allows easier sharing of files among peers. Both systems include internal email, but Carmen also adds options such as internal course mail only. Both systems allow for real time chat rooms, calendars, working offline and searching within the course. Teacher candidates can also work in groups with their own chatroom and discussion forum. Carmen also offers groups their own dropboxes and lockers. The Desire2Learn™ product was chosen because of its functionality for the university as a whole. The team that compared the course management systems felt that Desire2Learn™ was a more robust system and had an easier to learn format for faculty and staff.

Method

The participants in this study were 41 graduate students, 34 females and 7 males. The group was almost evenly split between those seeking Early Childhood certification (21) and those seeking Middle Childhood certification (20). The instructors of the Capstone seminar created a handbook based on the mission of the College of Education and Human Ecology, the Interstate New Teacher Assessment and Support Consortium (INTASC), the criteria of the National Council for Accreditation of Teacher Education (NCATE), and the standards of the Ohio State License. This handbook included guidelines and scoring rubrics, which were transferred to Carmen.

Instructors have the ability to post documents, video clips, and internet addresses into Carmen for teacher candidates to utilize. There are also discussion rooms and places for work to be uploaded. The instructors of Capstone used Carmen to collect student entries and record comments about teacher candidate work. Carmen also served as a place to post announcements for teacher candidates and for them to look at a personal grade book that kept track of their grades for each assignment throughout the year.

The instructors collected data of teacher candidates’ knowledge, skills and dispositions by reviewing their monthly artifacts and reflections, as well as conducting pre and post surveys. The participants posted an artifact and a reflection paper by the end of each month for one year. By analyzing types of artifacts and ways of posting the artifacts and reflections, the instructors measured the level of knowledge and skills regarding the Carmen system, Microsoft office products, and in most cases, effective use of a scanner and a digital camera. Reading the reflections indicated the level of comfort with state standards. Teacher candidates were required to choose an artifact from their course or field work and reflect on how this artifact advanced their knowledge level regarding teaching standards. The instructors also reviewed the final reflection paper that showed the teacher candidates’ growth in teaching philosophy and understanding of teaching and learning.

This group was also surveyed about the Capstone process in the beginning and at the end of the program. These surveys (see Figure 2) were given electronically by using Flashlight™ (http://ctlsilhouette.ctlt.wsu.edu/ctlsilhouette2_5/), and the results were recorded in a spreadsheet. The survey data were analyzed by the instructors and two graduate assistants.

The purpose of this survey was to indicate teacher candidates’ perceived level of knowledge concerning the use of an online course site (Carmen) and the use of a cumulative portfolio project for students in the Master of Education.
program at The Ohio State University at Lima and it refers to the objectives from the following EDU T & L capstone course. Teacher candidates were asked to mark only one answer and rate their perceived current knowledge in five levels; 0= No knowledge, 1= Little knowledge, 2=Partial knowledge, 3=Sufficient knowledge, and 4=Full/complete knowledge to achieve understanding. In order to save space, we present only survey questions not the original survey, which was in a table format.

1. How much knowledge do you have about portfolio assessment?
2. How familiar are you with Carmen?
3. How familiar are you with online discussions?
4. How familiar are you with posting assignments online?
5. Understanding the content being taught as addressed in content standards for the various disciplines.
6. Becoming familiar with relevant aspects of all students' background knowledge and experiences.
7. Articulating clear learning goals that are appropriate for the students.
8. Demonstrating an understanding of the connections between the content that was learned previously, the current content, and the content that remains to be learned in the future.
9. Understanding and utilizing appropriate educational technology to support student learning in various content areas.
10. Creating or selecting a wide variety of teaching methods, learning activities, and instructional materials or other resources that are appropriate for the students and that are aligned with the goals of the lesson.
11. Creating or selecting evaluation and assessment strategies that are authentic, appropriate for the students and that area aligned with the goals of the lesson.
12. Creating a climate that promotes fairness.
13. Establishing and maintaining rapport with students.
14. Communicating challenging learning expectations to each student.
15. Establishing and maintaining consistent standards of classroom behavior.
16. Making the physical environment as safe and conducive to learning as possible.
17. Including demonstration and use of appropriate safety procedures and adherence to safety rules.
18. Considering ands sensitively addressing the needs of diverse learners, including diversity in academic abilities, physical abilities, culture, race, gender, socioeconomic status and sexual orientation.
19. Making learning goals and instructional procedures clear to students.
20. Making content comprehensible to students.
21. Encouraging students to extend their thinking.
22. Integrating learning across subject areas.
23. Integrating technology and information literacy.
25. Monitoring students' understanding of content through a variety of means, providing feedback to students to assist learning, and adjusting learning activities as the situation demands.
27. Providing culturally responsive teaching for all students.
28. Reflecting on the extent to which the learning goals were met.
29. Demonstrating a sense of personal and professional efficacy.
30. Building professional relationships with colleagues to share teaching insights and to coordinate learning activities for students.
31. Communicating with parents or guardians about student learning.
32. Contributing to the community through community service.
33. Extending your professional knowledge through professional development activities.
34. Demonstrating a commitment to personal life long learning.

We also asked teacher candidates to describe their experiences with technology and web-based portfolio.

35. Did you ever seek assistance with the portfolio process this year? (yes, or no)
36. If you did seek assistance for questions regarding Capstone and the portfolio process how did you seek help?
37. If you sought assistance this year, in what area did you need help?
38. What are the benefits of using Carmen to you as a learner?
39. What are the difficulties for you as a learner when course participation in Carmen is required?
40. What is your idea of participating in online discussions and activities?
41. In what ways is the Portfolio process influencing your growth as a teacher?
42. How prepared do you feel to teach students with diverse learning needs? Why or why not?
Results: Consequences of Using Carmen

The instructors’ goals were to determine the effectiveness of using a web-based course management/portfolio tool for master’s students during the Capstone experience. Reviewing reflection papers and measuring the pre and post surveys of candidates’ perceived knowledge, skills and dispositions gave the instructors a strong indication if the candidates had achieved better knowledge of technology usage and state standards. The final goal was to increase the candidates’ willingness to utilize technology in their teaching as well as for creating portfolios with their own students in future teaching placements. Teacher candidates were specifically asked about the benefits and difficulties of using the Carmen system (see Figure 2) and similar comments are paraphrased and grouped in each section.

Participating in online discussions and activities

In the beginning, preservice teachers’ attitudes toward the web-based portfolio were evenly divided, positive and negative. Some examples of positive comments were

I have never participated in online discussions, but I feel it is necessary to learn.

I think it sounds like a great idea, but I've never done it before.

Some examples of negative comments were

I don't like online discussions. It is easier to accomplish in person. The time we had an online discussion was very distracting with other conversations and late ideas posted that were irrelevant to the subject matter being discussed.

These discussions are typically very stale and do not provide the same depth of benefit/learning that occurs within a class.

Or, some tried to define what online discussions and activities are:
Blogging, submitting questions/answers, getting others ideas, plans, resources, etc...

By the end of the program, their comments concerning participating in online discussions and activities had changed to predominately positive. One student indicated that it is “similar to talking on the phone, but being able to discuss things with more than one other person.” Another student noted that she “gained knowledge through technological advances-videos online.” However, some teacher candidates still preferred in person discussions.

Knowledge and skills of using instructional technology

Teacher candidates who used Carmen report greater skill in using tools such as PowerPoint and Microsoft Word, after being required to use these tools in Capstone. Figure 3 depicts preservice teachers’ perceived knowledge and skills of using educational technology that supports student learning, as well as integrating technology and information literacy; thirty-three percent in the pre and two percent in the post surveys reported having no (NK) or little knowledge and skills (LK); 45% percent in pre and 87% in the post survey rated themselves as having sufficient (SK) or full knowledge and skills (FK). In summary, teacher candidates were more satisfied that they had created a more comprehensive portfolio that showed their growth as a teacher in terms of knowledge and instructional skills.
Knowledge and skills of developing web-based portfolio

Teacher candidates were asked to rate their familiarity with and technological skills for using the on-line tool, Carmen. As can be seen in Figure 4, more than half of preservice teachers already reported having sufficient or full knowledge and skills for using Carmen in the beginning of the program. After completing the web-based portfolio, 85% of preservice teachers reported that they have sufficient or full knowledge and skills for using Carmen, online discussions, and posting assignments online.

Challenges of using a Web-based portfolio

Despite the increasing knowledge and skills with technology, technical difficulties such as problems with their personal computers, Carmen being down, or accessibility from home are reported by preservice teachers throughout the program. These sample comments were typical of the participants:
I liked the assignment submission and online discussion posting because they are not things that I will always use or have used so I am more comfortable with using them. Sometimes I do not like to use things I am not comfortable, but now I am more likely to use them.

It is difficult if your home computer does not allow you to use certain applications such as videos.

Access to another computer when having trouble with ours at home, and wanting to get on Carmen and unable to when Carmen or the university website was down.

The issues preservice teachers reported in the beginning are possibly due to the lack of experience with on-line courses. They may also be habituated in the face-to-face education context. They strongly believe in personal interactions or direct guidance. Some comments made by preservice teachers are as follows:

Simply getting used to using the program.

If you have questions regarding the assignment the response takes awhile, plus sometimes it is easier to understand things in person.

No one to ask questions directly to.

The best instruction still comes through face to face, teacher directed classes.

Not knowing if my assignment was received.

Sometimes postings didn’t work right or the assignments didn’t always show up. Stuff would upload incorrectly.

I would email a professor to ask about an assignment, but couldn’t understand their explanation; Not understanding directions.

By the end of the program, the issue of not having regular face-to-face contacts is still brought up by three teacher candidates. However, most preservice teachers seem to accept the fact that it is a web-based portfolio. The reported difficulties are more technical and practical based on what they experienced.

Access to another computer when having trouble with ours at home, and wanting to get on Carmen and unable to when Carmen or OSU website was down.

It takes a long time for me to upload some artifacts.

Paper. People may have had a hard time not having all the materials in hand.

I would procrastinate because I knew the deadline could be met at night, so I could work on the activity at home. This did not help me organize.

I don’t have a nice portfolio in hand to keep, instead it is online or on CD. I couldn’t really use it to take to interviews.

Benefits of using Web-based portfolio

In the beginning of the program, the majority of preservice teachers saw the course tools as convenient, an easy way to get information, and as a good communication tool. They used these tools to get class information, check and submit assignments, and check grades. By the end of the program, they listed Carmen’s organizational function as the most beneficial. These features are easily viewed from the course homepage (see Figure 4). Students will see messages from the instructors and an indicator on the left side that shows if there are any discussion messages to read. When they click on Content they will find information about the course, including the handbook, assignments, and examples of acceptable entries. When teacher candidates click on Discussions they will have the opportunity to view different forums on topics concerning their portfolio. When Dropbox is selected teacher candidates can load
their monthly reflections or assignments. The final tab is Grades, and this allows teacher candidates to see the grade item, points and grade assigned by the instructor. Other benefits of having the web-based portfolio reported by preservice teachers are summarized below:

- Being able to talk to numerous people
- Gaining knowledge from others
- Improved understanding of the uses for posting and uploading information/questions
- Liked doing discussions.
- Not having to drive to campus to turn in an assignment.
- It is helpful to become more knowledgeable about how technology can be helpful and useful.
- Could always find answers to questions by searching through Carmen
- Saved paper
- Easy to post assignments by the due date and understand what the criteria is
- Using Carmen was an easy way to communicate with the instructors

Discussion and Implications

The goal of this study was to determine the effectiveness of using a web-based portfolio with Master’s students during their Capstone experience. The instructors were hoping that the increased flexibility in course delivery coupled with frequent use of technology would result in stronger knowledge, skills and dispositions for these preservice teachers. Posting an artifact and a reflection paper each month helped the teacher candidates get familiar
with Carmen and practice using technology. The monthly selection and reflection process also helped them view their learning process not as a group of independent subjects but as an aligned whole.

Results indicate that a better grasp of technology is needed, which the instructors will continue to work on. More skill training on using scanners and making Microsoft Office products available to everyone are important issues for future Capstone seminars. Incorporating the use of recent technologies such as Google Docs and discussions about Web 2.0 would be beneficial. The instructors might use Google Docs, a free word processing and spreadsheet site (http://docs.google.com) to link teacher candidates to their documents, and those of their peers, for longer than the calendar year that they are currently available. Google Docs could be very useful for encouraging collaboration throughout the course of study, and it can be easily used in conjunction with the existing Carmen course. Web 2.0 is a bigger concept and involves things such as blogs or wikis. Ways in which using technology to read and write are shared through the World Wide Web. The reason the instructors chose to use Carmen was driven by university policy. The instructors do find Carmen slightly easier to use than WebCT™ or Blackboard™, but similar results may be possible for other instructors using these course delivery systems. A pre and post survey of instructors’ knowledge, skills and dispositions regarding use of the web-based portfolio would also be worthy of future investigation. In addition, content analyses of teacher candidates’ reflection would provide teacher educators with insightful suggestions of using a web-based portfolio.

**Conclusion**

Some Capstone teacher candidates discussed technology use in their teaching placements, but no overt attempt to educate the M.Ed’s about these new technologies took place. This will be important for further technological gains in future classes. The more up to date and pertinent the information, the stronger the possibility to positively affect dispositions will be. Dispositions are hardest to affect, because it takes a long time and frequent exposure to influence attitudes and beliefs. Therefore, the portfolio development process was spread evenly throughout the program. The instructors were pleased with the improved level of reflection about practice, but see a need for further discussion about how the reflection process should work, and a more clear definition of reflection should be made available in the handbook, as well as verbally discussed. The instructors also see a need to make sure each teacher candidate leaves the program with a CD-ROM of the portfolio work they uploaded on Carmen. Teacher candidates were not required to create a paper portfolio, so this CD will give them something to review as they look for employment and reference as they begin their teaching profession. The technology based portfolio project was rated as successful by instructors and participants. As technology use in higher education increases the need for technology training for both teacher educators and preservice teachers will continue. The web-based portfolio is one tool to affect a positive change in preservice teachers’ knowledge, skills and dispositions towards teaching and technology.

**References**


Wilcox, B. (1997). *Intelligent portfolios for professional development, Duquesne University, Pittsburgh, PA, School of Education.*

Evaluating an Intelligent Tutoring System for Design Patterns: the DEPTHS Experience

Zoran Jeremić¹, Jelena Jovanović¹ and Dragan Gašević²

¹FON – School of Business Administration, University of Belgrade // jeremycod@yahoo.com // jeljov@gmail.com
²School of Computing and Information Systems, Athabasca University, Canada // dgasevic@acm.org

ABSTRACT
The evaluation of intelligent tutoring systems (ITSs) is an important though often neglected stage of ITS development. There are many evaluation methods available but literature does not provide clear guidelines for the selection of evaluation method(s) to be used in a particular context. This paper describes the evaluation study of DEPTHS, an intelligent tutoring system for learning software design patterns. The study which took place during the spring semester 2006 was aimed at assessing the system’s effectiveness and the accuracy of the applied student model. It also targeted the evaluation of the students’ subjective experiences with the system.

Keywords
Intelligent tutoring systems, design patterns, learning evaluation, student model assessment, evaluation, adaptive presentation

Introduction
Software design patterns (DPs) have been recognized as very important and useful in real software development since they provide an elegant way of getting around problems that often occur. To be more precise, DPs are common conceptual structures that describe successful solutions to common and recurring software design problems. They can be applied over and over again when analyzing, designing, and developing software applications in diverse contexts (Harrer & Devedzic 2002). DPs are aimed at seamless reuse of software designs and architectures that have already proven their effectiveness in practice. In addition, their application guarantees high quality software solutions that are easy to maintain and extend.

To make use of the benefits that a DP-based software development offers, one has to master DPs both in theory and practice. Accordingly, it is very important to devise an effective approach for teaching DPs. Our experience in teaching object-oriented programming to software engineering students, suggests that there is no single best way to learn DPs. It depends on the students’ knowledge of and experience in the field of software engineering in general and design patterns in particular. For example, beginners in this field are often confused by the number of available patterns, different forms for presenting design patterns and plenty of available sources of information (Raising 1998). It would be best for them to be provided with a step-by-step introduction into the field and strict instructions on how to conduct each step (i.e., what to learn, in which manner, and what resources to use). In this case, sources of information should be carefully selected in order to avoid the threat of using a source providing information that is too difficult for a beginner to understand. On the other hand, for experienced programmers this way of strictly guided learning is not suitable. Rather, exploratory learning would be more appropriate learning strategy to use. It assumes giving a student control over his (since gender-neutral language tends to be imprecise and/or cumbersome, throughout the paper we arbitrarily decided to use masculine pronoun for each generic reference to a student) learning process. In particular, the student chooses on his own what to learn (topics and their order), when to learn and how fast to proceed, at which level of difficulty to learn, and in what way to learn. Such a student has to have access to many different information sources to explore, and thus be enabled to develop his own understanding of the topics explored. Exploratory learners are always intrinsically motivated. It is not possible to force this learning style upon a learner (Holzkamp, 1995). To sum up, students come from different knowledge backgrounds and have different learning styles and preferences. To effectively meet these diversities, an e-learning system should be able to offer different learning experiences to different students.

We found that an intelligent tutoring system (ITS) for learning DPs can successfully address these issues, since it can provide different learning strategies for different types of students, as well as, adapt learning content according to the student’s performance. ITSs use artificial intelligence’s techniques and methods to provide a student with a content that is neither too easy (thus avoiding the drop of concentration), nor above his/her ability to understand (hence not
frustrating the student). ITSs make it possible to deploy a course suitable for students with widely varying needs and levels of motivation.

We have developed DEPTHS (Design Patterns Teaching Help System), an ITS for learning DPs, as a part of our research in the area of teaching DPs in higher education (Jeremic & Devedzic 2004). Our goal was to provide students with the benefits of one-on-one instruction in a cost effective manner. The system was primarily intended for teaching undergraduate students of Computer Science, but it can be equally successful in other education settings, as well. DEPTHS performs adaptation of the teaching material based on the student performance and the usage tracking data collected during the learning session. The quality of the adaptation provided by the system depends on many factors, such as the accuracy of the diagnostic tools that collect and process data, the student model that is used to store the data and the embedded teaching rules. The only way to verify the quality of the DEPTHS’s adaptation functionalities is to evaluate the system in real conditions, with students who are learning the subject material. Evaluation is the central part of quality assurance, as it provides feedback on teaching and learning and thus helps to make the system better.

The evaluation of DEPTHS was conducted in the 2005/06 academic year, in the context of a course we teach at the Department of Computer Science of the Military Academy of Belgrade in Serbia. The aim of the evaluation was to determine how effective DEPTHS is for learning DPs. To this end, the following agenda was set to guide the research:

• Do a literature review in order to determine how the effectiveness of learning can be evaluated.
• Generate an evaluation instrument for determining the effectiveness of learning.
• Conduct an evaluation study to determine how effectively students learn with DEPTHS.

The paper is organized as follows: after giving literature review (Section 2), we introduce the DEPTHS system and give a general description of its functionalities (Section 3). Section 4 explains the DEPTHS architecture in details, whereas Section 5 gives an example usage scenario with DEPTHS. In section 6 we give a detailed description of the evaluation studies that were performed to evaluate the system’s effectiveness and the accuracy of its student model. We give a comparison with some related work in section 7. Finally, we conclude with a discussion of the implications of this work for the next generation of intelligent tutoring systems.

Literature Review

Evaluation of an ITS is a difficult task. In particular, the main difficulty lies in the absence of a widespread agreement on how it should be performed.

The most well-known and used model for measuring the effectiveness of training programs is the model developed by Donald Kirkpatrick in the late 1950s (Kirkpatrick 1979). Since then, it has been adapted and modified by a number of researchers, but the basic structure has remained the same.

The Kirkpatrick’s model defines four levels of evaluation (Kirkpatrick 1979):

• Evaluation of reactions – Reaction is the term that Kirkpatrick uses to refer to how much the students liked a particular training program. An evaluation of students’ reactions consists of measuring their feelings, and does not include a measurement of what was actually learned. A typical instrument for gathering information regarding students’ reactions is an open-ended questionnaire. This information is easy to collect, but does not tell enough about the training success.
• Evaluation of learning – This level of evaluation identifies how well the students understood the facts and techniques presented in the training material. This is much more difficult to measure than reactions. At this level, each student’s learning should be measured by quantitative and objective means. Endres and Kleiner (1990) state that pretests and posttests are necessary when evaluating the amount of learning that has taken place.
• Evaluation of behavior (transfer of learning) – This level of evaluation identifies how well the students apply the acquired knowledge in their everyday practice. This kind of evaluation is more difficult than the abovementioned two and there are very few examples of studies in this area. Feedback from students, their supervisors, and peers as well as some other techniques can be used for collecting information at this level.
• Evaluation of results – The fourth level of evaluation refers to the training results or impact on the organization. Although measuring training programs in terms of results may be the best way to evaluate their effectiveness, this kind of measurement is very difficult to conduct. The major obstacles include the existence of many factors that are impossible to evaluate (e.g., social interaction between employees), and the relative lack of objective, valid tools to use. McEvoy and Buller (1990) question the relevancy of such evaluations. They claim that not all training is result oriented: it can also be used for purposes other than achieving a measurable impact on the performance of an individual employee.

Since Kirpatrick established his original model, other theorists, e.g., Jack Philips (1997), and indeed Kirkpatrick himself, have referred to a possible fifth level, namely ROI (Return of Investment) which is used to evaluate the efficiency of an investment or to compare the efficiency of a number of different investments.

Endres and Kleiner (Endres & Kleiner 1990) use Kirkpatrick’s model as the foundation for their approach to evaluating the effectiveness of management training. They suggest setting initial performance objectives and monitoring accomplishment of those objectives after training. They offer an example in which participants are asked to write their personal and professional objectives at the end of the training experience. These objectives are sent to the participants approximately a week after the training. Two months later they are sent again, and the participants are asked to comment on their performance against these objectives:

• Reeves and Hedberg (2003) criticize Kirkpatrick's approach as overly simplistic for the following reasons:
  • Control groups are rarely feasible in education or training contexts;
  • Paper-and-pencil tests lack reliability and validity in measuring KSAs (Knowledge, Attitude, and Skills).
  • 100% response rate is unrealistic
  • Evaluation results are not the only input to decision-making within an organization

The experience reflected in the literature suggests that trainers often incorporate at least the first three levels of the Kirkpatrick’s model in the design of training programs. In fact, many authors emphasize the importance of considering early in the training design process how each level of evaluation will be addressed. Kirkpatrick’s fourth level of evaluation is still difficult to apply in practice (i.e., do the required measurements). The difficulty lies in the inability to separate training from the multitude of other variables that can impact long-term performance.

**Functionalities of DEPTHS**

DEPTHS (Fig. 1) is an ITS for teaching DPs. The course implemented in DEPTHS gradually introduces the concept of DP and presents students with the most frequently used kinds of patterns. In addition, DEPTHS provides learning materials adjusted to the student’s performance (e.g., his background knowledge and performance in the current domain), and cognitive capacity (e.g., his working memory capacity). Even though it is a tutoring system, DEPTHS allows students to choose between the tutoring and the self-paced learning mode.

The topics of the DEPTHS course are organized in a dependency graph, with links representing the relationship between the topics, such as prerequisite and related topics relationships. A student is ready to learn a topic only if he has completed the prerequisites. Accordingly, the student must achieve sufficient score for the topic in order to qualify to proceed with the course.

DEPTHS performs both content-level (adaptive presentation) and link-level adaptation (adaptive navigation support) (Brusilovsky 1996). Adaptive presentation means that students with different performance levels get different content for the same domain topic. During each learning session, DEPTHS observes the student's progress and adapts the course presentation accordingly. Likewise, the student’s performance is used to adapt the visual representation of hyper-links to related topics. DEPTHS provides two kinds of navigational adaptation through the course material (Brusilovsky 1996):

• Direct guidance – The student is given only one option to continue the browsing activity, that is, just one button enables advancing to the next page. The destination of the “next” button (Fig. 1E) is dynamically determined by the system.
• Link removal – advanced students can choose the topics to learn by selecting appropriate links from the Content menu (Fig. 1A). However, the links that the system considers inappropriate are removed, i.e. they are no longer available.

Figure 1. Web-based GUI of DEPTHS

DEPTHS continually monitors the student’s behavior and performance and stores all the interaction data. This data is used to refine the instructional plan, as well as to generate recommendations for the further work (Prentzas et al. 2002).

DEPTHS architecture

The architecture of DEPTHS (Fig. 2) follows the main guidelines for ITS architecture. In particular, it is widely agreed that the major functional components of an ITS architecture should be: the domain knowledge model, the student model, the pedagogical module, and the user interface (Hartley & Sleeman 1973, Burton & Brown 1976, Wenger 1987). Accordingly the DEPTHS architecture consists of Pedagogical Module, Expert Module, Student Model, Domain Model, and Presentation module (Jeremic & Devedzic 2004b). The user interface component of the traditional architecture is adapted to the web-based environment and is dubbed Presentation Module in our model. In the following subsections we describe each of these components in turn.
DEPTHS is implemented as a classical client-server Web application with a Web browser on the client side and Tomcat 5.0 Web Server as JSP container on the server side (Jeremic et al. 2004).

*Figure 2. DEPTHS system architecture*

**Domain Model**

*Domain Model* is designed as a network of concepts. What we refer to as a ‘concept’ is named differently in different research papers – attribute, topic, knowledge element, object, and learning outcome. We use this term to refer to a topic that describes a single design pattern. Concepts could be related to each other with prerequisite relations. One concept might be a prerequisite for another concept, i.e., the former should be learned before the latter can be presented. For each concept, there is a knowledge threshold (measured through the score obtained on assessment mechanisms) that must be reached by a student before the system can assume that the student has learned that concept. Each concept is decomposed into units – content elements that correspond to a particular web page to be presented to students. The number of units for a particular concept is defined by the teacher. DEPTHS uses the unit variants technique (Beaumont 1994, Kim 1995, Kay & Kummerfeld 1994) to achieve content-level adaptation. This means that the system keeps a set of alternative units for each concept and makes a selection based on the student’s characteristics (e.g., based on the student’s knowledge level: very low, low, good, very good, and excellent). Each unit has an arbitrary number of fragments – chunks of information that can be presented to the student. A unit may also contain fragment variants, i.e. fragments related by the “OR” relationship.

**Pedagogical Module**

*Pedagogical Module* provides the knowledge infrastructure necessary to tailor the teaching material to the student’s learning needs as captured in his student model (see the next section for a detailed explanation of the *Student model*). This module is composed of three functional components (Jeremic et al. 2004):

- Instructional Planner generates an instructional plan as a sequence of instructional plan items (e.g. concepts and lessons) appropriate for the student’s knowledge of the domain.
- Feedback Mechanism is responsible for providing a student with advice and recommendations for further work during a learning session.
- Assessment Component is in charge of collecting and processing data about a student’s learning process and storing it in the student model. The accuracy of the diagnostic tools of this component is one of the main factors that affect the quality of the adaptation process.

All the above mentioned components use Expert Module for making decisions, such as selecting appropriate lesson to teach or questions to ask the student, and selecting advice about the lesson that is appropriate for a student to learn next.
Expert Module

Expert Module’s main role is to make decisions that are used by Pedagogical Module for generating teaching plans and adaptive presentation of the teaching material. In many traditional ITS systems these roles are in the charge of the pedagogical module, but we decided to separate them in two distinct modules in order to increase the flexibility of the system.

One of the main tasks performed by this module is selection of instruction plan items that best suit to a particular student’s knowledge level (Fig. 3). This is a three-step process which starts with the creation of a concepts plan. To create such a plan, Expert Module uses a set of facts about the student (from the Student model), the set of available domain concepts (from the Domain model), and a rule for concepts selection. Expert Module selects concepts having minimal (i.e. entry) knowledge level lower or equal to the student’s current knowledge level and having all of the prerequisites satisfied. In other word, Expert Module will not choose those concepts that it believes the student is not ready to learn yet, neither will it select a concept whose prerequisite (some other concept) the student has not passed yet. This process repeats each time the student knowledge level is changed. The created concepts plan is displayed to the student in the form of adaptive Content menu (Fig. 1A).

When the student or the system selects the concept to be learnt next, Expert Module proceeds to create a detailed lesson plan. In the final step a test plan is created. DEPTHS does not create a detail lesson and test plan for each concept because there is a chance that some of them will not be needed (depending on the student’s performance). These two steps will not be described in detail here since that issue is out of the scope of this paper (Jeremic, 2005).

![Figure 3. The process of creating Instructional plan in Expert module](image)
Student Model in DEPTHS

The student model stores and updates data about the student’s performance in a specific subject domain. It is essential for the system’s operations that adapt instructional material to the student’s characteristics (Devedzic 2001) and comprises both the model of the student and the mechanisms for creating the model (Prentzas et al. 2002).

The student model may keep any number of students’ characteristics, depending on the system requirements. In DEPTHS, three basic categories of the students’ characteristics are used:

1. Personal data – personal characteristics of a student (e.g., name, ID, and e-mail). This information represents the static part of the student model, and is collected during the student’s first learning session with the system, through a questionnaire.
2. Performance data and individual preferences – cognitive and individual characteristics of a student. This part of the student model represents a mixture of static and dynamic data. Static data, such as the desired detail level, the experience level or the preferred programming language, is collected during the registration procedure (through a questionnaire). Dynamic data is derived from the learning sessions and is changed as the student progresses through the course material. Examples of dynamic data include: actual skill level, learning style, reasoning ability, etc. The performance data, both static and dynamic, is quintessential for one of the system’s primary functions – adaptive presentation of the teaching material. This data represents the system’s ‘believes’ of the student’s traits. The quality of these believes directly affects the quality of the content-adaptation process.
3. Teaching history – data related to the student’s interactions with the system during learning sessions. This part of the student model keeps track about everything that the student has done during the learning process. In particular, it keeps data about each student’s session with the system, such as the time spent on solving tests and the student’s success on a particular test. This data is less important for adaptive presentation than performance data, but it is very important for reflective learning which often plays considerable role in a learning process. The system uses this data in order to provide the student with feedback about what he has done well and where he failed, what should be revised and how to make the learning process more successful.

When registering a new student, the system creates his student model and populates it with default values. Miskelly uses a similar approach in Interactive Student Modeling System (ISMS) (Miskelly 1998). In particular, based on the student’s self-assessment, the system classifies the student into one of the following categories: beginner, intermediate or advanced (expert), i.e., it assigns him one of the predefined stereotypes. A learning session then proceeds in compliance with the assigned stereotype until the completion of the first concept, when the student is tested for the first time. Based on the test results, the Pedagogical Module updates the values of the attributes in the Student Model concerning the student knowledge level (actual knowledge level). This data belongs to the performance data, and is updated based on the system’s conclusions about the students’ knowledge level.

DEPTHS uses tests to assess students’ knowledge of each domain topic. A test is created as a set of questions and exercises dynamically selected depending on the student previous knowledge level. The Pedagogical module (Fig. 2), i.e. its Assessment Component to be more precise, is responsible for diagnosing student knowledge level based on the test results. It uses a set of pedagogical rules and domain knowledge in order to assess test results and diagnose the student’s knowledge level based on these results. The result of this diagnosis depends on the three main variables: 1) the test difficulty, which is based on the difficulty of each question as specified by the question creator; 2) the time spent to solve the test, which is compared with the sum of average time for each question as specified by the question creator; and 3) the test result (success). The computation of the student’s knowledge level is based on fuzzy sets (test difficulty, test duration and success) and is calculated as a fuzzy value (ranging from 0 to 6). Thus, computed knowledge level represents a quantitative measure of the system’s belief about the student’s level of understanding of the respective concept. Based on each concept’s assessment, the system creates an overall estimation of the student’s understanding of the domain. This value is calculated as the average value of all concepts degrees of mastery. If assessment on a particular concept is performed more than once, the best score is taken.

DEPTHS in Use – An Example Usage Scenario

When a student runs a session with DEPTHS for the first time, he has to give some basic information about himself, as well as to describe his own experiences in software development as a beginner, intermediate or advanced. After that, he is requested to solve a questionnaire composed of very general questions concerning software development.
Based on the questionnaire results and the student’s self-evaluation, the system categorizes the student into an appropriate stereotype (student) model. Thereafter, the system switches to the teaching mode and develops an initial plan of instructional actions.

In the teaching mode, the student is presented with a concepts plan available as a navigation menu on the left side of the screen (Fig. 1A), as well as the main menu and command buttons available on the top and bottom of the screen (Fig. 1C and 1E). Those links that the system considers inappropriate for the given student are hidden. The student can choose between two options available in teaching mode:
- following the system’s advice (recommended for beginners);
- choosing the concepts to be learned next on his own (suitable for advanced learners).

If the first option is chosen, the command buttons are used for navigation; otherwise the navigation is done through the navigation menu. Each time the student presses the “Next” button or chooses a link from the navigation menu, a new lesson is presented in the main window area. Which lesson is going to be presented depends on the lesson plan created before learning of a specific concept has started. As DEPTHS uses “lesson variants” for adaptation on lesson’s level, it will choose one of the available variants of the lesson based on the student’s previous results. However, the student can choose or change the desired detail level (low, normal, high) any time from the main menu. For example, low detail level will contain only basic description of design patterns without source code examples, diagrams, etc. Student can also choose which programming language will be default for presenting source code examples (Java, C# or C++), or which design pattern form will be used (Gang of Four, Coplien or Alexandrian).

![Figure 4](image)

*Figure 4.* The student can see his performance record, which stimulates reflective learning.
After the presentation of one domain concept is completed, the system switches to the evaluation mode. In this mode, the system provides the student with exercises and tests in order to assess how much the student has learned about the concept. If the student demonstrates insufficient knowledge of the concept, the system suggests better exploration of the current concept and provides the student with an appropriate learning path for that concept. As each question is connected with specific lesson(s), the system can easily provide links to the lessons which has to be relearned. The student can choose either to follow the advice or to move on. After the repeated exploration of the suggested lesson(s), the student is presented with another test addressing the lessons that have been relearned.

After each test, the system displays a page with student-directed feedback, including (but not limited to) the following:

- the student’s performance on each question,
- the correct answer for each question,
- the lessons that should be learned again,
- the student’s overall knowledge of the tested domain concept,
- the correlation between this test result and his previous results,
- many other statistical information, such as the time spent to solve the test, the difficulty level of each question, and points number on each question.

Besides this real-time feedback, the system makes available on-demand statistical feedback. This feedback takes the form of graphical representation of the student’s progress through the course materials, using chart diagrams (e.g., line chart, column chart, and pie chart). In addition, it comprises pages with relevant information concerning the student’s learning sessions with the system (Fig. 4). Besides the abovementioned information about tests (Fig. 4C and 4D), it provides information about each concept the student has passed (Fig. 4A) as well as the lessons passed for each concept (Fig. 4B).

After the completion of each concept, the system has to check the student’s performance, and change the instructional plan accordingly. This change is reflected in the Content menu (Fig. 1A). The system also provides an annotated list of learnt concepts in a form of a navigation menu (Fig.1B). The annotations are used to indicate how well the student has performed on the passed concepts and enabling the student to reflect on his performance. The student can use this navigation menu to quickly access the concept(s) he wants to learn about again.

Evaluation studies

According to Kirkpatrick’s model (described in Section 2), evaluation of Results/Organizational impact needs at least a two year evaluation period. In this study it was not possible to delay evaluation for such a long period as in two years time students would have already left the Academy. Regarding that fact, we decided to use the first two levels (reaction and learning) of the Kirkpatrick’s model to evaluate the preliminary effectiveness of the DEPTHS system and the accuracy of its assumptions about the student.

The method that we applied to evaluate the system’s effectiveness consisted of:

- analyzing the students’ reactions to the training program, and
- conducting an experiment with an experimental and two control groups and comparing pre-test and post-test results of these groups.

The method we used to test the accuracy of the student model consisted of comparing the system’s assumptions about students’ performance level to their results on an external test. In particular, we monitored the progress of each individual student and assessed the quality of tutoring by comparing the student’s records in his student model with his results on the external test. We have used the extent of overlapping of these values as the measure of the accuracy of the Student Model.

DEPTHS was evaluated with a group of 42 students who took part in our course on software development in Spring semester 2006 (from February to July 2006). The study was conducted at the Department of Computer Science of the Military Academy in Belgrade in Serbia, with the computer science students of the third academic year. It was organized during the class hours of the Software development course that offered face-to-face mode of learning. The
students were told that their performance observed during this study will influence their marks for the course. In particular, their performance with DEPTHS participated in the final mark with 40%. The students already had some elementary knowledge in the domain of software design patterns from their previous education.

Reaction evaluation

The participant reaction is a measure of “customer satisfaction” indicating the level of effectiveness and usefulness of a training program at the time the participants are experiencing it (Antheil & Casper 1986). It can be very useful for determining the quality of the learning process. In particular, if the participants’ reaction to the training program is positive, learning is more likely to take place than in case when their reaction is negative. Every teacher is usually anxious to know how students react to his/her teaching efforts and what they have learned. Because the participants are at one place, the required information can be obtained quickly and easily and subsequently used to improve the program. We found that the evaluation of participants’ reaction could be very useful for the overall evaluation of DEPTHS for various reasons:

- It helps us know immediately what worked well and what should be improved (or changed all-together).
- It provides information for improving the current system and designing future versions.
- It shows to the participants and stakeholders that we are interested in improving their satisfaction with the course and that we value their input.

The participants’ reaction can provide us with valuable information, but we have to be very careful with conclusions, and use this information in combination with the information gathered through other evaluation methods. The reason is that in most cases, there is no definite relationship between what people feel about they have learned and what they actually did learn (Dixon 1990, Le Rouzic & Cusic 1998).

We used an interview to collect data about the students’ satisfaction with and attitudes towards learning with the DEPTHS system. The interview was also supposed to reveal the students’ perceptions regarding the effectiveness of learning with DEPTHS. We defined the following rules to ensure accurate measurement:

- the interview was designed in a manner that the collected data can be easily tabulated and manipulated by statistical means. Most of the questions were close-ended questions, based on the Likert scale with five responses ranging from ‘Very much’ (5) to ‘Not at all’ (1);
- the interview was anonymous;
- we used a few open-ended questions in order to encourage the participants to make additional comments not elicited by the given questions.

The questions were divided into three sections based on the type of information we were interested in. The first section gathered questions regarding computer-assisted learning. These questions were designed to gather information about the students’ previous experience with and opinion about computer-assisted education. For example, we asked the students how often they use computers for learning, and for what purposes: to search the Web for relevant learning materials, to access a Web-based learning application, etc. In addition, we were interested in the students’ perception of the relevancy of computer-assisted learning and whether it could (or even should) replace the traditional classroom-based learning.

The questions of the second section were related to the DEPTHS system. We think that a learning system must be easy to use, and must provide students with all tools necessary for successful learning. Students must not be frustrated during the use of the system, because of technical or other problems. We wanted to know to which extent we achieved these goals with DEPTHS and the students’ responses to the second section of the questionnaire were supposed to give us the answer. In particular, we wanted to know what they thought about the system, did they experience some technical problems, did they enjoy (and to what extent) using the system, and if they found it difficult to use. We believed that these answers can help us to improve the design and functionalities of the system in the future.

The third section of the questionnaire was aimed at evaluating the course on design patterns offered by the DEPTHS system. The purpose of this section was to measure the students’ satisfaction with the content of the course. We wanted to know what they thought about the organization of the course content, and if they found the content too
difficult for them. We also asked students about usefulness of the examples that were used throughout the course, and how the course could be improved.

The interview revealed that, in general, the students did not have much experience using similar systems. For example, the majority of them (92.85%) have either no or very little experience with learning management systems (Fig.5 – Q.2c)(all the questions that are mentioned (Q.x, x=1...21) in this section are given in the Appendix). We found that most of them are not familiar with electronic educational tools, and do not use the Internet and web-based learning programs for education (Fig.5 – Q1, Q.2a, b, d). However, the students expressed positive attitude toward computer-assisted learning. More than 70% of them think that computer-assisted learning should be made available to supplement traditional lectures and exercises (Fig.5 – Q.3). However, only few (28.58%) students would not mind using learning programs that are in English (Fig.5 – Q.5). Note that the course offered by DEPTHS is in English while all other lectures are offered in Serbian. The students were supposed to be able to follow lectures in English since they had a mandatory 4 semester long English course as a prerequisite for the course in which DEPTHS was used.

The DEPTHS system received high marks from the students. The majority of the students (78.57%) felt that the DEPTHS system is very easy to use (Fig.6 – Q.8), only 7.14% found navigational tools confusing (Fig.6 – Q.9) and most of them (85.71%) did not experience either any technical problems (64.28%) or at least any significant problems (21.43%) while using the system (Fig.6 – Q.10).

More than 90% of the students thought that the learning goal, outcomes, objectives and expectations for the course taught by the DEPTHS system were clearly stated and that the offered course met them (Fig.7 – Q.12 and Q.13). The majority of the students (85.71%) found the course content well organized (Fig.7 – Q.14), but some of them (28.58 %) thought that the amount of content was not suitable for the length of the course (Fig.7 – Q.15). Nearly two
thirds of the interviewed students (64.28%) reported that they were encouraged to take responsibility for their own learning (Fig. 7 – Q.18). During the course, most of the students (84.72 %) often had a clear idea of where they were going and what was expected from them (Fig. 7 – Q.19). Only 35.72% of the students reported that they were not given enough choices over how they are going to learn in the course (Fig. 7 – Q.20). The majority of the students acknowledged that they had learned a lot about design patterns from the DEPTHS system (92.86 %) (Fig. 7 – Q.21).

Figure 6. The students’ responses to the questions concerning the DEPTHS system

Q.8 Do you find the system easy to use?
Q.9 Do you find navigational tools confusing?
Q.10 Did you experience any technical problems while using the system?

Figure 7. Chart diagrams presenting the students’ answers regarding the course taught by DEPTHS
The students were encouraged to report other advantages and disadvantages of learning with DEPTHS through the given open-ended questions. Having analyzed their answers to these questions, we realized that the majority of the students perceived good organization of the course content and convenient access to the relevant online resources and multimedia materials as the primary advantages of learning with DEPTHS. In addition, two students reported a great satisfaction with the feedback they were receiving during learning sessions. Three students found tests provided at the end of each concept as a great advantage of the system.

Students also noted some negative sides of learning with DEPTHS, which were primarily related to the technical problems that some of them had experienced, and the lack of support for collaborative learning. For example, one student reported that the problems with local area network obstructed his learning with DEPTHS. Four students reported that they would appreciate if they were enabled to communicate with each other or with the teacher during the session. Two students reported the lack of students’ forum to support learning.

To sum up, the main advantages of the DEPTHS system, as perceived by the interviewed students, are a good organization of the course material, simplicity of the use and its ability to let students work at their own pace instead of learning with a tutor led session. Students also found that the feedback messages, which they were receiving during learning sessions, and advice about what should be learned next, were very useful. On the other hand, the major weakness of DEPTHS is its incapacity to support student-student and student-teacher interactions. In addition, the students mentioned that the system could be improved by adding more functionality for knowledge assessment.

Learning evaluation

For this evaluation study we decided to use ‘nonequivalent comparison-group design’ (Marczyk et al. 2005), one of the most commonly used quasi-experimental designs for determining system effectiveness. In order to comply with the requirement of this design, we attempted to select groups of students that are as similar as possible. We were constrained by the number of students and teachers in the Military academy. As this experiment was organized as a part of the regular classes of the Software development course, we were bound to engage all 42 students of the third academic year who were eligible to attend these classes. Since computer laboratories at the Military academy have 15 workstations for students, we had to limit the group size to the maximum number of 15 students. In such a limited environment an inadequate selection of participants (i.e. students) for the experiment is highly probable. In order to prevent this selection threat (an incorrect selection would obscure the experiment results), we decided to use two control groups, rather than the usual one.

Thus, one experimental-group and two control-groups, with 14 students each, were involved. The groups were created based on the students’ final grades in the courses related to software engineering. Our presumption was that these grades indicate the students’ general knowledge of the software engineering domain (column ‘School grade’ in Table 1). In addition, all of the students were tested at the outset of the study (in the so called pre-test phase) to evaluate their knowledge of the subject domain (i.e., software design patterns). The pre-test allowed us to measure differences between groups before their exposure to the intervention. This way we substantially reduced the threat of selection bias, since the pre-test revealed whether and to which extent the groups differed on the dependent variable (i.e., knowledge level of the subject domain) prior to the intervention.

The pre-test consisted of 16 multi-choice and free form questions. The questions were designed to test the students’ elementary knowledge of software engineering in general and design patterns in particular. In addition, this test was aimed at identifying the students’ expectations from the course and its delivery mode as well as their previous experience in the subject matter.

<table>
<thead>
<tr>
<th>Group</th>
<th>School grade (5-10)</th>
<th>Pre-test (1-100)</th>
<th>Post-test (1-100)</th>
<th>Pretest/Posttest Difference</th>
<th>Statistical Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>experimental group A</td>
<td>7.88</td>
<td>57.21</td>
<td>76.57</td>
<td>22.37</td>
<td>P&lt;0.01</td>
</tr>
<tr>
<td>control group B</td>
<td>7.86</td>
<td>55.86</td>
<td>70.21</td>
<td>16.14</td>
<td>P&lt;0.01</td>
</tr>
<tr>
<td>control group C</td>
<td>7.88</td>
<td>58.21</td>
<td>68.14</td>
<td>10.36</td>
<td>P&lt;0.01</td>
</tr>
</tbody>
</table>
Throughout the semester, students in the experimental group (group A) were learning exclusively with the DEPTHS system, whereas students in the control groups (groups B and C) were learning in the traditional way. Three teachers were involved in this experiment, one for each group. Teachers in the control groups employed traditional methods (including multimedia resources), while the teacher of the experimental group was helping students in using the DEPTHS system and by providing assistance when needed.

The course consisted of 10 concepts (i.e., software design patterns) which were gradually introduced, one after the other. During the semester, the system was creating a test for each domain concept (i.e., design pattern) and assessing the knowledge of the students from the experimental group based on the test results (as explained in Section 4.4).

At the end of the semester, the students from each group had to take part in a test created by a human teacher. This type of test (often referred to as post-test) is used to appraise the learning effectiveness after the learning process took place. It consists of a number of questions aimed at a detailed evaluation of a number of issues related to the quality of the learning material and the delivery mode.

In order to analyze the tests’ results, we decided to use a paired t-test to compare differences between pretest and post-test scores for each group (O'Rourke et al. 2005). The obtained results are presented in Table 1. The table reveals that in group A (students who learned with DEPTHS), the pretest score was 57.21, the post-test score was 76.57, and the scores showed significant improvement (t=-9.19, P<0.01). In group B, the pretest score was 55.86, the post-test score was 70.21, and the scores also showed significant improvement (t=-5.88, P<0.01), but in the lesser extent than in group A. Similar results were observed in the other control group (group C): the pretest score was 58.21, the post-test score was 68.14, and even though the scores showed significant improvement (t=-5.92, P<0.01), the improvement was not that great as in group A. In a nutshell, the statistical results showed that the difference between the pretest and post-test scores was greater in group A than in groups B and C. The test score improvements in groups B and C were similar. These findings indicate that DEPTHS has a beneficial effect on the students’ learning and brings in improvements in performance over time.

“One-way analysis of variance (One-way ANOVA)” statistical model (Davies 2000) was used to identify differences among the three groups involved in the experiment. This model was chosen, since it proved as the most appropriate method of statistical analysis of experiments involving more than two groups. It can be used to test the hypothesis that the means among two or more groups are equal, under the assumption that the sampled populations are normally distributed. Our null hypothesis was stating the research question as “no difference” between the experimental group and the control groups.

<p>| Table 2. ANOVA table shows statistical analysis of the post-test scores among three groups |
|---------------------------------|---------------|-------------|----------------|-----------------|----------------|</p>
<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>14</td>
<td>1072</td>
<td>76.57143</td>
<td>192.7253</td>
</tr>
<tr>
<td>Group B</td>
<td>14</td>
<td>983</td>
<td>70.21429</td>
<td>258.1813</td>
</tr>
<tr>
<td>Group C</td>
<td>14</td>
<td>954</td>
<td>68.14286</td>
<td>233.0549</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom (df)</th>
<th>Mean Square</th>
<th>F-value</th>
<th>P-value</th>
<th>F-critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>540.1429</td>
<td>2</td>
<td>270.0714</td>
<td>1.18459</td>
<td>0.316636</td>
<td>3.238096</td>
</tr>
<tr>
<td>Within Groups</td>
<td>8891.5</td>
<td>39</td>
<td>227.9872</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9431.643</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

f (F-value) - measurement of distance between individual distributions

df (degree of freedom) - an equivalent of currency in statistics – you earn a degree of freedom for every data point you collect, and you spend a degree of freedom for each parameter you estimate

P (P-value) - the probability of obtaining a result at least as extreme as a given data point, assuming the data point was the result of chance alone
Having applied the One-way ANOVA model on the pre-test results, we found that these scores did not differ much between the three groups (f=0.108, df=2, P= 0.89). Table 2 presents the results of applying the One-way ANOVA model on the post-test results. Post-test scores among three groups showed no significant differences (f=1.18, df=2, P=0.32), as well (Brown et al 2006). As F-value is lower than the F-critical value, we concluded that statistically there is no significant difference, and the score differences could be explained the best as being present by chance. The outcome of our test was that we accepted the null hypothesis that the students in the experimental group will perform in the similar manner as students who learned in the traditional way (groups B and C).

Even though the results of One-way ANOVA can bring conclusion that DEPTHS does not lead to significant improvements over the traditional learning, we are encouraged with the fact that our system even in its early stages has better results than traditional learning. In particular, the results of t-test performed on the experimental group have shown significant progress of those students who learned with DEPTHS. Moreover, the qualitative results however indicate that the students prefer this kind of learning over the traditional learning of design patterns. Our experiences from this evaluation will lead to many improvements of the system which will make it far more successful in the future.

Student model assessment evaluation

To evaluate the accuracy of the DEPTHS’s student knowledge assessment we compared the students’ knowledge of the domain as captured in their student models with their results on an external test, performed after the course was completed. For each student, the DEPTHS system keeps track of all the tests he has solved throughout the course (i.e., after having learned each domain concept) and evaluates his overall knowledge of the domain from the tests’ results. The external test was designed by the group of teachers, experts in the domain of software engineering. It was given to the students at the end of the course to assess their knowledge of the whole course content.

The results of the 14 participants from the experimental group were analyzed. The questions of the external test were divided into 10 groups, each group covering one domain concept, so that, based on the test results, we could assess the students’ knowledge of each concept individually. Comparing the results of the external test with the student model data (Fig. 8), we found that the external test shows slightly lower knowledge level (15.43%). That difference was expected, because the system (i.e. its student modeling component) assesses student knowledge on each concept immediately after the student has learned it, whereas the external test assesses student knowledge at the end of the whole course when a number of other variables affect the final results. For example, students could not use DEPTHS out of the classroom, but we could not control if they are using some other sources of information (literature or the Internet). Moreover, when estimating the student’s knowledge of a concept, the system takes the best score over all the test attempts related to that concept. However, if instead of taking the best result we took the average test result for a concept (from system’s logs on each concept), we get results that are 11.37% lower than those on the external test. As the result of the student model assessment evaluation performed on DEPTHS, we finally concluded that the comparison between the external test and student model could not be realistic indicator of the student model accuracy, as it is performed during the long period of time (one semester), while system assesses student knowledge immediately after the learning of each concept. Therefore, we decided to analyze the system’s log data in order to get some more information about the student model used.

By analyzing the system’s log data we found that at the beginning of the course students tend to spent more time on each lesson, give more attempts at quizzing and were often revisiting previously learned concepts. The majority of the students (87%) attempted to solve the tests more than once during the first quarter of the course, and 34% gave three or more attempts. This trend was diminishing as the course proceeded. At the last quarter, 49% of the students made two or more attempts and only 11% tried to do the test more than twice. We think that this decrease in the number of test reattempts is due to the fact that as the end of the semester was approaching, the students were having increasingly more obligations on other courses and less time to dedicate to learning with DEPTHS. This is a well known phenomenon, present at the Military Academy for years.

We found an additional interesting issue by analyzing the system’s log data. Some questions, annotated by their creators as very easy, were not performed well by students. Contrary, some other questions described as very difficult were surprisingly well done. Having analyzed this issue, we concluded that it originates from the fact that the Diagnosis module uses data about a question’s difficulty level and time necessary to solve it provided by a
human teacher or question developer. However, a teacher can make a mistake when estimating the difficulty level of
and the required time for a specific question. Such a mistake would lead to the inappropriate adaptation of test, for
example, if a teacher describes a difficult question as very easy, the system will use it to test student with low level
of knowledge. Based on this conclusion we have improved our Diagnosis module. It now performs analysis of the
system’s logs data, compares it with question’s difficulty level and time necessary to solve the question, and changes
it if a variance exists. We strongly believe that this model could significantly improve older solution. However, we
have not evaluated this model yet.

![Graphically presented comparison of the student’s knowledge as estimated by the DEPTHS’s student model and the results of the external test](image)

**Figure 8.** Graphically presented comparison of the student’s knowledge as estimated by the DEPTHS’s student model and the results of the external test

**Related Work**

The main problem in evaluation of intelligent tutoring systems is that the available literature does not suggest any
clear guidelines for an evaluator which methods to use in particular context. In ITS community, the common practice
is to perform experiments with a particular set of users. However, many different approaches to performing these
experiments have been proposed by different authors.

Comparing to our work, we found a similar research presented in (Mitrovic et al. 2002). In this work the authors
presents the results of three evaluation studies performed on SQL-Tutor, an ITS for the domain of SQL database
query language. These studies were primarily focused on the assessment of the Constraint-Based student model that
SQL-Tutor uses (i.e. how well this model supports student learning). In addition, the studies were aimed at collecting
students’ opinion of the system’s usefulness. Overall, the results of these studies were positive. On the regular
classroom examination that followed the experiment, students who had been learning with SQL-Tutor significantly
outperformed students who learned in the traditional manner. However, this experiment was not controlled, and
because some students kept on using the system after the class hours, the results can not be considered as a definitive
proof of the system’s quality. Contrary to this, students who worked with DEPTHS could not use the system out of
the classroom. However, since our experiment lasted for a long period of time, other variables could have affected
our results.

In (Weibelzahl 2003) authors describe an approach for the evaluation of adaptive learning systems in general, which
assumes evaluation on four levels: evaluation of the input data, evaluation of the inference mechanism, evaluation of
the adaptation decisions, and evaluation of the interaction. They have applied this model to an adaptive learning
system, called the HTML-Tutor, to demonstrate the usefulness of the suggested approach. Similarly to the DEPTHS,
they have performed the evaluation of the accuracy of student model. They suggested two methods to perform this:
by comparing its assumptions to an external test and by comparing its assumptions to the actually displayed behavior
of the student. Moreover, they have focused also on the evaluation of the inference mechanism. Their results show
that the evaluation of the accuracy of the system’s assumptions about the student could point to possible
improvements of the inference mechanism. However, they did not present any results about the overall system
effectiveness.
Another similar evaluation was reported in (Miller & Butz 2004). This study has evaluated the usability and effectiveness of Interactive Multimedia Intelligent System (IMITS), a system designed to tutor second year electrical engineering undergraduates. IMITS was evaluated from two perspectives, usability of the software and effectiveness of the software. The Usability Questionnaire was used to gather information concerning students’ reactions to the software. Also, usability data was obtained from the system’s log files. Similar technique as in DEPTHS was used to examine the impact of IMITS on student learning. The authors have used a quasi-experimental design, with a control and an experimental group of students. Analysis of variance (ANOVA) was used to test the hypothesis that the students who were using IMITS learned more than their counterparts under control conditions. Overall, students’ responses were favorable. IMITS improved performance on at least one classroom achievement measure. Regression analyses revealed that the more students used IMITS to learn some engineering concept (with usage defined as percentage of the questions encountered on a particular engineering concept), the better they learned that concept. Comparing to IMITS, we used the similar evaluation design in DEPTHS. We have reported on the systems usability and its effectiveness. Both evaluations have shown that the students’ responses were favorable. However, we go step further in evaluation of the system effectiveness. Contrary to IMITS evaluation of DEPTHS introduce a comparison of pre-test and post-test results in experimental group that lacks in IMITS.

Our approach to evaluation of DEPTHS was influenced by the work reported in (Barker et al. 2002). They have applied a qualitative approach when evaluating their adaptive multimedia tutoring system. The authors have pointed out that an evaluation in real-world settings presents so many uncontrollable variables that traditional quantitative methods can give misleading results. They have suggested combining qualitative and quantitative analysis and have shown the advantages of this combined approach in real-world contexts. Their quantitative ANOVA results have shown that certain differences in mean values are unlikely to stem from random variations among students. Such strong rejection of the random-variations hypothesis would be impossible with a purely qualitative analysis. While agreeing with this approach, we have also adopted Kirkpatrick’s model of evaluation as a foundation of our evaluation design in DEPTHS.

We found an interesting approach for evaluating the accuracy of the student model in (Millan & Perez de la Cruz 2002). The approach is based on the usage of simulated students since the cognitive state of a simulated student, unlike that of a real student, can be precisely determined, so the evaluation results are more accurate. However, as the authors pointed out, simulations could not provide the same degree of validity as experiments with real human students. This work is a good example of separating the evaluation of the accuracy of a student model from the evaluation of the adaptations efficiency based on the student model. By evaluating an intelligent tutoring system modularly with simulated students, the effects of each component can be separated from one another and the weakest component(s) could be easily identified and improved in order to increase the effectiveness of an adaptive system. In the future, we are planning to apply the similar approach with DEPTHS, in order to evaluate its modules, namely student model, pedagogical module, as well as the accuracy of student model assessment and adaptation effectiveness. We believe that this approach could give us valuable results on design and efficiency of each DEPTHS’s module.

Conclusions

This paper presents the evaluation of DEPTHS, an intelligent tutoring system (ITS) for teaching/learning software design patterns. The adopted evaluation approach is quite general, and can be equally well applied for evaluation of other ITSs. In particular, we used first two levels (reaction and learning) from the well-known Kirkpatrick’s model (Kirkpatrick 1979). The conducted evaluation studies targeted primarily the effectiveness of the DEPTHS system as well as the accuracy of its assumptions about the students’ knowledge level. The semester-long evaluation study has provided us with insights into strengths and weaknesses of the DEPTHS system. It has also made clear directions for future actions.

We reported several advantages of the DEPTHS system over the traditional approach to teaching/learning design patterns. Students who learned with DEPTHS found that the system helped them to learn a lot about design patterns. They were especially pleased with the system’s ability to provide them with many useful information, feedback messages and advice for further work. Students’ responses indicated the need for regular communication with teachers and other students as the underpinning priorities for successful completion of online learning.
To test the learning effectiveness of DEPTHS, we used a t-test and compared the pre-test and post-test results of an experimental and two control groups. The statistical analysis showed that the change from the pre-test to the post-test results was greater in the experimental than in the control groups. These findings indicate that students who learned with DEPTHS performed better than students who learned in the traditional way and that learning with DEPTHS brings in improvements in performance over time. We have also applied the One-way ANOVA statistical model to test for differences among the three study groups. The outcome of this test was that we accepted our null hypothesis that the students in the experimental group will perform as well as students who learn in the traditional way. This finding is obviously inconsistent with the result of the t-test, and another confirmation of the difficulty of accurately measuring the effectiveness of a certain educational tool on the students’ performance. However, we are encouraged with the fact that this system even in its early stages has better results than traditional learning.

Finally, we compared the students’ knowledge of the domain as captured in their student models with their results on post-test in order to evaluate DEPTHS’s ability to accurately assess student knowledge of the subject domain. We found that the proposed student model does reflect the students’ knowledge of the subject matter, regardless the observed slight difference between the end-test results and the student model.

Regarding our future work, there are several directions to be explored. Our first goal concerning further improvements of the DEPTHS system includes enabling the Student Model to store not only student’s knowledge, but also students’ cognitive and behavioral characteristics, such as memory capacity and motivation. This data will facilitate even more effective adaptation of the learning material. We are also planning to support interactions among students and the tutor in a group discussion environment, as well as to enable integration of additional supporting software tools, such as ArgoUML (http://argouml.tigris.org/). Regarding that, we are currently developing a Semantic web-based framework that integrates a set of existing tools in order to enhance students learning experience when collaboratively learning about Software patterns on the Web (Jeremic et al. 2008).

References


APPENDIX A

Computer Science and Software engineering,
Military academy in Belgrade, Department of Defense
Serbia 2006

COURSE EXPERIENCE QUESTIONNAIRE

This questionnaire is intended at collecting the information which will help us to evaluate the DEPTHS learning system from the perspective of its technical characteristics and performance as well as from the perspective of the quality and adequacy of the course it provides.

The information you will provide in this questionnaire will be kept strictly confidential. Neither your name nor any details that might be used to identify you will be published. We guarantee that the data will be used exclusively for requirements and statistical analysis.

1=Never/Not at all/Not good at all    5=Always/Very much/Extremely good

<table>
<thead>
<tr>
<th>SECTION: General, e-learning related questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How familiar are you with the electronic educational tools?</td>
</tr>
<tr>
<td>2. How often do you use a computer for learning? To...</td>
</tr>
<tr>
<td>a)...search the Internet for relevant web pages.</td>
</tr>
<tr>
<td>b)...download papers or similar items (with a known Internet address)</td>
</tr>
<tr>
<td>c)...access a learning management system offering a course of interest</td>
</tr>
<tr>
<td>d)... use a web-based learning application</td>
</tr>
<tr>
<td>3. Computer-based learning can replace lectures and exercises</td>
</tr>
<tr>
<td>4. Computer-based learning should be used only to supplement classroom lectures and exercises.</td>
</tr>
<tr>
<td>5. I have difficulties or I just don’t like working with learning programs which are not in my mother-tongue.</td>
</tr>
<tr>
<td>6. Web-based learning should be nothing more than the distribution of notes over the Internet.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SECTION: Evaluation questions regarding the system</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Did you enjoy learning with DEPTHS?</td>
</tr>
<tr>
<td>8. Do you find the system easy to use?</td>
</tr>
<tr>
<td>9. Do you find navigational tools confusing?</td>
</tr>
<tr>
<td>10. Did you experience any technical problems while using the system?</td>
</tr>
<tr>
<td>11. If your answer to the previous question was 5 or 4 please describe the kind of problems you had.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SECTION: Evaluation questions regarding the course</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. The learning goal, outcome, objective, and/or expectations for the course were clearly stated.</td>
</tr>
<tr>
<td>13. The offered course met the stated learning goal, outcome, and/or objective</td>
</tr>
<tr>
<td>14. The course content was well organized.</td>
</tr>
<tr>
<td>15. The amount of content was suitable for the length of the course.</td>
</tr>
<tr>
<td>16. The content of this course was too difficult.</td>
</tr>
<tr>
<td>17. Helpful examples were used.</td>
</tr>
<tr>
<td>18. I was encouraged to take responsibility for my learning.</td>
</tr>
<tr>
<td>19. During the course, you usually had a clear idea of where you're going and what was expected from you.</td>
</tr>
<tr>
<td>20. Students have a great deal of choice over how they are going to learn in this course.</td>
</tr>
<tr>
<td>21. To what extent do you feel you have learned from the program?</td>
</tr>
<tr>
<td>22. This course can be improved by:</td>
</tr>
<tr>
<td>23. Please feel free to make any additional comments about this course.</td>
</tr>
</tbody>
</table>
Digital Students in a Book-Oriented School: Students’ Perceptions of School and the Usability of Digital Technology in Schools

Yifat Ben-David Kolikant
School of Education, Hebrew University of Jerusalem, Mount Scopus, 91905 Jerusalem, Israel // Tel: 972-2-5882056 // yifatbdk@mscc.huji.ac.il

ABSTRACT

Today’s students were born into a world of digital technology. We investigated the impact of computers and the Internet on the learning preferences of students whose schools do not use this technology in class, specifically, (a) the usability they attributed to the technology for tasks requiring the processing of information; and (b) their attitude towards the technology in relation to their attitude towards the type of learning used in school. Our focus was the history class. To this end, surveys were filled out by three different classes that do not use computers in school: one in a high school that advocates beyond-information activities and two in information-focused public (i.e., state) schools—a high school and a middle school. These attitudes were found to be negatively correlated with the legitimacy ascribed to the form of learning used in school. Specifically, the two public-school classes used digital technology in the belief that they knew better than their teacher how to pursue a school information-focused agenda, whereas the third class legitimated the form of learning used in school. In neither case, however, was there any indication of a fertile human-computer partnership, envisioned as the desired form of learning for the digital age.

Keywords
Forms of learning, Learning resources, Usability, Human-computer intellectual partnership

Introduction

Children all over the world use computers and Internet technology, thereby informally learning how to learn with the technology. It is widely argued that technology, computers and the Internet shape how we think (e.g., Bolter, 1984). In school, however, students encounter a form of learning shaped by books. This encounter may feel awkward. Students might doubt the relevance of school as a learning-to-learn institute given their informal learning experience; or they might internalize schooling as a privileged form of learning. In both cases the fertile human-computer interaction needed by for tomorrow’s citizens in a rapidly changing world (e.g. Salomon et al, 1991) is not likely to be established. Understanding how students perceive the relationship among school, learning, and digital technology—the goal of the study described in this paper—is important because schools have a responsibility for preparing the new generation for life.

Digital learning among children

An OECD report (OECD, 2005) shows an increment in children’s use of computers in all 40 countries participating in the survey. In Israel, where the study described was conducted, nine out of ten households with teenagers have at least one computer connected to the Internet, according to a survey on e-readiness carried out for the Ministry of Finance (Mizrahi et al., 2005). This survey also pointed out that teenagers have high e-readiness, i.e., they are able to benefit from digital technology.

Harel (2003) refers to today’s children as Clickerati and as the three X's generation; the latter term emphasizes their computer literacy and their learning through bricolage by eXploring, eXpressing, and eXchanging ideas (hence the three X’s) using technological means. In this way they differ from previous generations that learned from books. These characteristics of children’s learning preferences are also emphasized by other terms used to describe them, such as the digitally fluent (Resnick, 2002), the Net Generation (Tapscott, 1998), and digital learners (Brown, 2002).

Children’s tendency to learn by exploration was observed long before computers became common (Holt, 1964). However, computer and Internet technology made bricolage legitimate or even necessary (Turkle, 1995). Turkle argues that children, like other computer users, “have moved in the direction of accepting the postmodern values of opacity, playful experimentation, and navigation of surface as privileged ways of knowing” (Turkle, 1995, p. 267).
Thus, the use of computers changes not only one’s preferred forms of learning but also epistemological beliefs about learning and knowing. Correspondingly, Shaffer and Clinton (2006) assert that the digital media pose challenges to educators because they encourage the development of mental actions that are not at the heart of schooling.

The school agenda

Resnick (2002) claims that the school agenda is focused on information; thus questions of the following sort will arise with respect to teaching:
- What are the best ways to transmit that information from one person (a teacher) to another (a learner)? What are the best ways to represent and display information so that it is both understandable and learnable? (p. 32)

Resnick and other scholars contend that instead school could and should move beyond information, a phrase that assumes that the citizens of tomorrow who live in a rapidly changing world where knowledge ages quickly do not have to be taught specific information as much as how to effectively consume, create, and communicate knowledge (Papert, 1993, 1996; Scardamalia & Bereiter, 2003, Brown, 2002). Computer and Internet technology should be central to the new pedagogy because of its potential for engaging in an intellectual partnership with humans by sharing the intellectual load and thereby extending human intellectual performance (Salomon et al., 1991; Pea, 1985).

Engaging students in an effective intellectual partnership with computers is not a trivial matter. Merely placing computers near the students is not enough; rather, a suitable pedagogy must be developed. In reality, however, although many schools have set the goal of equipping schools with computers and Internet technology, as can be seen from the OECD report (OEDC, 2005), the school system—unlike other cultural institutions such as banks, hospitals, the entertainment world, and the press—has not re-thought the goals and practices of schooling in the digital age. Instead, computers are perceived by the school system as tools for enhancing the achievement of existing goals (Papert, 1996, 1998). Moreover, Cuban and associates (Cuban, 2001; Cuban et al., 2001; Peck et al., 2002) found that school practices did not change even when students and teachers had ready access to computers and the Internet. Except in a few teachers’ classes, instruction remained teacher-centred and the use of the technology was sparse, mainly involving the social sciences, where students used a word processor or searched the Internet for information. Furthermore, they point out that the school agenda in many developed countries, such as Israel and the US, is focused on achieving high scores on standardized tests; because teachers already have pedagogical tools to accomplish this, they are rightfully dubious about the need for computers.

The encounter between digital and school forms of learning

When students enter school they encounter an institution that was designed before computer and Internet technology was envisioned and thus is based extensively on book technology. Brown (2002) describes the shifts in learning brought about by the digital Internet medium, pointing out the differences between the form of learning used by the generation raised on books and that used by the digital generation. He claims that learning has shifted from being authority-based to being discovery-based. In contrast to the book generation, which tended not to try things unless it was clear how to use them, digital learners are not afraid to try new environments; in fact, they prefer to learn about them through exploration and trial. Digital learners are able to navigate “through confusing, complex information spaces and feel comfortable doing so” (p. 14). They can get used to new, rapidly changing Internet genres, whereas the book generation are used to more stability and hierarchy. Reasoning has shifted from the linear, deductive, and abstract style of the book generation to the bricoleur style, an extreme demonstration of which is that when something “works”, no further justification is required.

The usability of books and the Internet. Given that the teachers themselves were raised on books, the encounter between competing epistemologies and the associated forms of learning might feel awkward to both teachers and students. Teachers and students might attribute different usability to the two types of learning resources, the Internet and books, for students. Usability is defined as an attribute that assesses how easy products are to use (Nielsen, 2000; also see Nielsen’s website). Usability spans five categories, three of which are relevant to this work: (a) Efficiency: assuming that I know how to use an artefact, how quickly will I perform tasks? (b) Errors: How many errors do I
make, how severe are these errors, and how easily can I recover from the errors? (c) Satisfaction: How pleasant is it to use the artefact? The other two categories are learnability (how easy is it to accomplish basic tasks the first time one encounters the artefact?) and memorability (when returning to the artefact after a period of not using it, how easily can one regain proficiency?).

Teachers, although probably aware of students’ affinity for the technology, might be concerned about the possibility of error inherent in navigating confusing, complex realms of information with varying degrees of reliability. Furthermore, being used to teaching in an ordered and hierarchical manner (Brown, 2002), they might find the overlapping presentations confusing and thus inefficient. Moreover, being loyal to the school’s current agenda, teachers may experience computer and Internet technology as unnecessary or even as hindering learning (and thus unsatisfactory), such as when technology distracts students or computers crash, thereby consuming valuable class time. In contrast, books have proven themselves for generations to be reliable and efficient resources.

Students—being used to and capable of navigating the Internet—may perceive the Internet as efficient, especially when working on a paper assignment, because it enables them to choose the most helpful presentation of information from among myriad websites. In contrast, the number of books people have in arm’s reach is usually limited and even then, books are inefficient, since it takes time to locate the required information and more time to manually summarize the contents; thus they are unsatisfactory. The greater assumed reliability of books is probably not valued as much by the book generation because children are used to judging the information they find, though probably not the way the generation that grew up on books legitimize (Brown, 2002).

School as a venue for culture clash, compliance, or acclimation. Given that in the near future almost everyone in the Western world will grow up with computers and thus will experience the “digital” form of learning, it is important to study the nature of the encounter between digital children and the book-oriented, “analogue” school. Papert (1993, 1998) predicts a culture clash. Similarly, Salomon (2000) warns that unless schools change, the encounter between the two forms of learning will lead both teachers and students to consider each other “urban barbarians” as far as knowledge and learning are concerned. Papert (1998) believes students will try to make the school adopt their form of learning. He calls them “an army for change”: “We have an army. It's this army of children, of kids coming into the school with a better image of learning and with the technical knowledge to implement that better image of learning.”

In a previous study in the field of computer science education, it was evident that students’ informal learning experience had brought about a culture clash. Students and teachers differed on what they considered a significant problem, appropriate approaches to problem-solving, and even a satisfactory solution. They delegitimized each other, causing bitterness on both sides (Ben-David Kolikant, 2004). For example, nearly half of the 138 computer science students insisted that a particular program was correct even though they knew it did not fulfil one of the requirements of the assignment. Their claim was that this requirement did not solve any significant or real problem, since a user could not tell whether it had been met. Consequently, the teachers were perceived as being petty. Additionally, consistent with Papert’s prediction, these students negotiated for an alternative curriculum with more emphasis on the exploration of new technologies and questioned the need for learning the theoretical principles (Ben-David Kolikant & Ben-Ari, 2008).

These results inspired the study described in this paper. Further work was required to examine whether the impact of students’ informal experience with digital technology is limited to computer science classes or whether it is relevant to the broader community of educators. One may argue that the situation described in computer science education is unique since much of it involves learning by doing with a computer, an area where students can claim mastery. Other subjects are less closely associated with computers; hence students would not feel that they are experts in the field and would adopt the form of learning presented. Indeed, Shor (1992) claims that students are acclimated to the “culture of silence” of mass education. Teachers are authorities, so they are supposed to do most of the talking. Moreover, students gradually internalize the idea that being a good student means keeping quiet, absorbing information, and agreeing with the teacher. Shor, however, wrote this thesis in 1992. Today, a decade and a half later, things have changed: many students have lived all their lives with computers and the Internet and computers have become common in school.

Consequently, this research project was launched in order to understand the influence of computers and the Internet on the learning preferences of students whose schools’ practices have not been invaded by this technology.
Specifically, two objectives were pursued: (a) exploring students’ perceptions of the usability of computers and the Internet for school, especially in tasks requiring the processing of large quantities of information; and (b) understanding students’ attitude towards the technology in the context of their attitude towards the form of learning used in school, especially the absence of digital technology.

**History as a venue for exploring the encounter**

History, the subject on which this study focuses, is a good venue for exploring the encounter of digital children with an information-focused agenda and its consequences. It is a compulsory subject in Israel, rich with textual information and constantly accused of focusing on information rather than on the practices and thinking processes of “doing history”, i.e., going “beyond information” (e.g., Loewen, 1995; Wineburg, 2001). For example, Loewen (1995) claims that “textbooks are full of information—overly full. These books are huge” (p. 3).

Students can choose to deal with the mass of information by using the Internet instead of or in addition to books. In fact, history can show how an agenda can be transformed even though practices have not changed. In the past, history provided opportunities to learn how to learn as students struggled with the main learning resource—books—thereby practising important skills such as distinguishing between the main idea and ancillary ones. Today, the use of the Internet might absolve students of having to invest the mental effort required when working with books; this may reinforce their impression that school learning means simply acquiring information. Students might also debate the need for and the benefits of school, given that the information acquired in school is accessible online.

This study was preceded by preliminary research, including interviews with 10 history teachers and 20 students from high schools that do not use computers in history classes. The interviews were with five triads of a teacher and two of his or her students (for a total of 15 interviews), plus five other teachers and 10 other students (another 15 interviews). All the teachers taught in urban public schools with good reputations in six different Israeli cities. Interestingly, most of the students thought history classes should remain compulsory. Nonetheless, most also found the Internet usable for history studies even though the teachers did not encourage its use. Students explained that “googling” is easy and comfortable, whereas books were described as exhausting and irritating even for a short summary. The most common use of the Internet was searching for information about an event or personage in order to write a paper or a short summary. Another common use was exchanging information as the matriculation exams approached. All of the students were aware of the reliability issue of using Web sources exclusively. Most of them therefore compared information from about three sites. Finally, most of the students believed that incorporating the use of the Internet in history class was a good idea but were not sure what, besides increasing their motivation, would be the benefit.

The teachers were ambivalent about technology as well as about their students. Some teachers described trying to introduce the use of computers in the hope that their students would learn more and told how they were painsed by their students’ refusal to work hard. The Internet was thus experienced as a hindrance to meaningful learning because it absolved students of having to invest cognitive effort and thereby diminished learning. In fact, some teachers asserted that they now require that students hand in summaries and hand-written assignments because that way something gets into their heads. Students were described, on the one hand, as being smarter than the teachers were at their age but, on the other hand, as being unable or unwilling to do anything requiring cognitive effort beyond memorization. Finally, most of the teachers emphasized the need for books in history as well as their personal love for the medium. Thus, students and teachers ascribed different usability to books and the Internet as learning resources.

**Methods**

Based upon the preliminary research, a three-part survey was composed and filled out by three history classes. The survey and the participants are described below.
Survey

Part A inquired about students’ use of computers and Internet technology after school hours for leisure and for schoolwork. The decision to explore students’ use of computers and the Internet for schoolwork in their free time was based on the following assumption about students’ behaviour during school hours:

Teachers hold the ultimate authority over what occurs in classrooms on a day-to-day basis. Students are thus subject to the pedagogical choices of their teachers. If teachers choose not to use technology, students will receive little exposure to the machines” (Peck et al., 2002, p. 478).

In contrast, after school students have more freedom to make their own decisions regarding how to work on school assignments, and in particular which of the two technologies to use: real-paper technology, such as their books and notebooks, or digital technology, such as the Internet.

Specifically, in order to determine their use of digital technology for schoolwork, we asked students to fill in a table where each row was devoted to one of the following school assignments: studying for a test, doing homework, or writing a paper/doing a project. The students were asked to list all the school subjects for which they had used digital technology at least once in the previous school year and to describe what they had done with the technology.

Part B described the work done by two imaginary students, Yossi and Miriam, on a graded paper assigned in a history class. Yossi works as follows. He (a) types the topic in Google; (b) chooses three sites, usually from the first and second pages of Google’s results; (c) verifies that the content of these sites is similar to make sure he can use the sites ;(d) merges them into one paper; and (e) if necessary, simplifies the language. Miriam works differently. She (a) looks for information in books; (b) looks for additional information on educational websites such as that of Ministry of Education; (c) plans the paper with the information she has in mind; (d) writes the paper; and (e) adds a list of her references. The students were asked to indicate how similar their approach is to Yossi’s and Miriam’s on a scale of 1 to 5 (where 1 means “very different” and 5 means “very similar”) and to list the differences between the way they work and they way each imaginary student does.

We chose writing a paper as the assignment because it leaves more room for the students to make decisions regarding the learning process than homework or tests do and thus could tell us something about students’ perceptions of the usability of books and technology for schoolwork. The choice of history is explained in detail above. The imaginary students were in the typical situation of doing work to be handed in to and graded by the teacher. Yossi’s profile was based on the students’ description of their Internet use during the preliminary research. Thus, Yossi produces his paper by merging three sources (as most students reported doing) that he found on the first or second page of Google results. Then, if necessary, he translates the product. Miriam’s profile was based on teachers’ descriptions of the ideal process for learning history. Miriam uses books as her main source. She uses the Internet as well, but only official sites like that of the Ministry of Education. Also, she plans her paper and then writes it, whereas the emphasis in Yossi’s work is on merging, in which case planning is unnecessary.

Finally, in part C the students were asked to rate, on a scale of 1 to 4, their agreement with statements about their self-perception as independent learners, their learning preferences (in particular the use of computers and the Internet), and their intellectual gain from school and history class. We deliberately asked about the intellectual gain from their history class the previous year in order to minimize the influence of the most recently studied topic on their responses. The statements are listed in table 5. The survey was administered about three months into the school year in order to give students enough time to form an opinion about what school offered.

Participants

The three classes that participated in the survey were carefully chosen to capture different characteristics of schools, including grade levels, the presence of the matriculation exams in the background, and topics taught. The first was Yaakov’s eleventh-grade class comprising 26 students in a well-regarded public school. Like most eleventh-graders in Israeli public schools, they would be taking the summer matriculation exam on Jewish and Zionist history. The second class consisted of Serge’s 29 eighth-graders in a well-regarded public school. This class did not have to take any external exams in history. They, too, were studying Jewish and Zionist history. In Serge’s and Yaakov’s schools, as in most Israeli public schools, project assignments, though they exist, are not at the heart of school life.
Nonetheless, both teachers assign at least one history project during the school year. Typically, both Yaakov and Serge would present a topic in class and explain it. Students were active in these lessons because they were asked to analyze maps and texts, yet the instruction was teacher-centred since the teachers controlled the information presented to the students. Note that both classes were starting their second year in their school. Furthermore, both classes were having the same history teacher for the second year in a row.

The third class, in contrast, was Gidi’s tenth-grade class (19 students) in an elite boarding school where admission is selective and is based on students’ intellectual abilities. This was their first year in the school. This school advocates a learner-centred approach and therefore projects are common. For example, each year students choose a topic and are given free time to investigate it. In addition, teachers assign various projects; for example, in history students did a project on a topic of their choice having to do with ancient Greece. Gidi’s support focused on historical issues, while another teacher helped them with aspects of the work process, such as what constitutes a good research question and how to construct an argument. Additionally, unlike in most Israeli high schools where students take one matriculation exam in history in the tenth grade and another in the eleventh grade, in this school tenth-graders learn about ancient Greece—a topic that is not on the matriculation exam—in the first semester before moving on to topics required by the Ministry of Education, and they take the first exam the next winter (in the eleventh grade) and the second in the summer. The idea is to give students the experience of meaningful learning without the distraction of the matriculation exam.

These classes were chosen after in-depth interviews with the teachers, observations of their classes, and conversations with several students. The first criterion for choosing the classes was good teachers, i.e., teachers whose students do well and consider their teachers good. This criterion was important in order to minimize the impact of students’ rejection of the teacher on their responses. Conversations with arbitrarily chosen students assured us that these teachers qualified. The second criterion was the teachers’ pedagogical decision not to merely force-feed the information required by the curriculum but instead to act as coaches promoting students’ lifelong learning abilities. The in-depth interviews with the teachers as well as the class observations ensured that. Indeed, all the teachers devoted time to class discussions of topics suggested by students and encouraged students to express their opinions and back them with data. Finally, all three teachers had a similar education. Each had a B.A. in history from an Israeli university and a teaching certificate.

**Findings**

**Part A**

The students’ responses to part A reveal that they spend a lot of their after-school time in front of the computer, but mostly for leisure (table 1). These results are similar to those described in the OECD report (OECD, 2005), as well as by the survey on e-readiness in Israel (2005). However, Gidi’s students spend less of their computer time on leisure activities than Yaakov’s and Serge’s students do.

| Computer use | Teacher | | | |
|---|---|---|---|
| | Serge | Yaakov | Gidi |
| In general | 3.46 (2.48) | 2.58 (2.09) | 1.94 (1.57) |
| For school | 0.98 (0.78) | 0.81 (0.62) | 1.18 (0.65) |
| **Table 1**: Students’ use of the Internet for schoolwork | | | |

| Assignment | Teacher | | | |
|---|---|---|---|
| Homework | Serge | Yaakov | Gidi |
| | 97% | 81% | 79% |
| Tests | 7% | 16% | 26% |
| Projects | 38% | 42% | 37% |
The exceptionality of Gidi’s class can be explained by the fact that they live in a boarding school and thus feel less of a need to communicate with their friends electronically as teenagers typically do. Additionally, Gidi’s students have a heavier load of schoolwork than they did in their previous schools and therefore might feel that their computer time should be spent on schoolwork rather than leisure activities. Indeed, three of Gidi’s students commented in the survey that before attending this school they spent much more time on the computer.

Students’ responses regarding their use of technology for schoolwork showed that they only used the computer for word processing (writing papers and lab reports) and occasionally for Internet access. Table 2 shows students’ after-school use of the Internet for three typical school assignments: studying for tests, doing homework, and doing projects; each figure represents the percentage of students who reported using the Internet at least once for that type of assignment. Students’ lists of the subjects for which they had used the Internet were usually limited to one or two, mostly in the social sciences and humanities, such as history and literature.

The students’ descriptions of the work they do on the Internet indicate that they access the Internet for one reason: to obtain information—as they put it, “material”—for projects, papers, and homework assignments in which they have to summarize information about a new concept, person, or event (a type of assignment typical of the social sciences). Few of the students use the Internet to study for tests; those few mainly do so to obtain additional material from the Internet, scan friends’ notes, or look for sample questions to answer. Additionally, 5% use online dictionaries. Only one student used the Internet for science (he asked a question in a forum). The tendency to use the Internet mainly for the social sciences—information-rich subjects—is consistent with the findings of Cuban et al. (2001).

### Part B

The students’ responses regarding the similarity between the procedure they follow when writing a paper and the procedures followed by two imaginary students, Yossi and Miriam, are shown in table 3 and figure 1. Table 3 shows the mean and standard deviation for each class. Figure 1 shows the distribution of each class’s responses; “similar” represents responses of 4 or 5, “somewhat similar” refers to responses of 3, and “different” refers to 1 and 2 on the original scale of 1 to 5.

There was a statistically significant negative correlation between students’ responses to the two procedures (Spearman ($r_s$)=-0.474**). That is, some students identified more with Yossi and others identified with Miriam. As can be seen in table 3, Serge’s students identified with Yossi’s style (M=3.83*, sd=1.136) more than with Miriam’s (M=3.07, sd=1.510). Yaakov’s students demonstrated a similar though less strong tendency (M=3.35*, sd=0.936 for Yossi and M=3.00*, sd=1.058 for Miriam). In contrast to both public-school classes, Gidi’s students identified with Miriam’s style (M=3.37*, sd=0.895) more than with Yossi’s (M=2.84*, sd=0.958).

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Serge</th>
<th>Yaakov</th>
<th>Gidi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yossi</td>
<td>3.83*</td>
<td>3.35*</td>
<td>2.84*</td>
</tr>
<tr>
<td>Miriam</td>
<td>3.07</td>
<td>3.00*</td>
<td>3.37</td>
</tr>
</tbody>
</table>

* A t-test was found significant for p<0.05.

A one-way ANOVA test revealed a significant difference among the three classes’ responses to Yossi’s procedure (F(2,71)=5.376, p<0.01). A Scheffe post-hoc revealed that Gidi’s class differed from Serge’s. Yaakov’s class was somewhere in the middle. This difference is demonstrated in figure 1. A small proportion of the students in the public-school classes found their procedure different from Yossi’s (Serge, 10%; Yaakov, 19%), compared with 37% of Gidi's students. Figure 1 demonstrates a difference between the two public-school classes and Gidi’s class with respect to Miriam's process too. Specifically, only a small proportion of Gidi’s students (21%) thought their procedure differed from Miriam’s, compared with 38% of both public-school classes (almost twice the proportion). However, this difference was not statistically significant.
Identification with Yossi’s procedure

Identification with Miriam’s procedure

Figure 1: Distribution of students’ responses on similarity of procedures

Table 3: Most common objections to Miriam’s and Yossi’s procedures

<table>
<thead>
<tr>
<th>Students’ most common objection</th>
<th>Serge</th>
<th>Yaakov</th>
<th>Gidi</th>
</tr>
</thead>
<tbody>
<tr>
<td>To Miriam’s procedure:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I prefer Internet resources to books/I wouldn’t use only books.</td>
<td>48%</td>
<td>54%</td>
<td>37%</td>
</tr>
<tr>
<td>To Yossi’s procedure:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would use books (or an encyclopaedia) in addition to Internet resources.</td>
<td>24%</td>
<td>27%</td>
<td>47%</td>
</tr>
</tbody>
</table>
The students’ most common response regarding the differences between their procedures and those of Miriam and Yossi concerned the resources (books and Internet) used, as indicated in table 4. The table shows that students approve of both resources; yet the public-school classes tend to prefer the Internet, as reflected in the fact that about half of Serge’s and Yaakov’s students (48% and 54%, respectively) objected to Miriam’s decision to use books primarily and only about a quarter (24% and 27%, respectively) objected to Yossi’s decision to rely exclusively on the Internet. In contrast, Gidi’s class had a more balanced attitude towards the use of books and the Internet, as reflected in the fact that nearly half of the class asserted that unlike Yossi they would use books and 37% objected to Miriam’s choice of resources.

**Part C**

Table 5 presents students’ responses to statements regarding the relationship between computers/Internet and learning as well as the intellectual gains from school and history classes on a scale of 1 to 4, where 1 and 4 represent strong disagreement and strong agreement, respectively. For each statement, the mean (M) and the standard deviation (sd) of each class is shown. An asterisk (*) was added near the mean to indicate significance according to a t-test (p<0.05). A mean greater than 2.75 is defined as agreement by the class with the statement; a mean smaller than 2.25 is defined as class disagreement with the statement. Means between 2.25 and 2.75 are defined as diverse opinions regarding the statement.

Most of the students agreed that they are capable of independent study (statement 1: Serge, M=3.31*; Yaakov, M=3.31*; Gidi, M=3.32*). No class agreed that living with computers and the Internet had made them independent learners (statement 2: Serge, M=2.59; Yaakov, M=2.17*; Gidi, M=2.68) and there was also broad disagreement with the statement that their generation, as a whole, know better how to learn than the generation raised without computers (statement 3: M=2.45, 2.12, 2.37). Thus, it was commonly agreed that the Internet per se does not make one a better learner.

However, differences were found among the classes with respect to their views on the intellectual gains from school in general and history class in particular, as well as the usability of computers and the Internet for schoolwork.

Two polar attitudes about gains from school and history class. Significant differences were found among the classes’ responses to all the statements (7–10) concerning intellectual gains from school and history class. The results of a one-way ANOVA test on these statements were as follows: statement 7, F(2,71)=3.884*; statement 8, F(2,71)=14.137**; statement 9, F(2,71)=4.106*; and statement 10, F(2,71)=10.783**, where * and ** indicate p<0.05 and p<0.01, respectively.

As can be seen from table 5, all three classes agreed that they had gained general knowledge from history class (statement 7). Gidi’s students did not think the class had improved their cognitive and analytical abilities (statement 8: M=1.84*, sd=0.688) or taught them about values (statement 9: M=2.26, sd=0.806). Additionally, Gidi’s students agreed that school as a whole provides them with tools for independent study (statement 10: M=3.43*, sd=0.692).

In contrast, the two public-school classes believed they had gained cognitive and analytical abilities from history class (statement 8: Serge, M=3.03, sd=0.778; Yaakov, M=2.81, sd=0.849) and had learned values (statement 9: Serge, M=2.86, sd=0.953; Yaakov, M=2.96*, sd=0.774). Additionally, unlike Gidi’s students, the public-school classes did not think school was providing them with tools for independent study (statement 10: Serge, M=2.75, sd=1.005; Yaakov, M=2.23, sd=0.765).

The Scheffe post-hoc test found that the public-school students differed significantly from Gidi’s students with respect to the statement that they had gained cognitive and analytical abilities from history class the previous year (statement 8) and the statement that school gives them tools for independent study (statement 10). A similar pattern was found in the responses to statements 7 and 9, that students had gained general knowledge (statement 7) and values (statement 9) from history class the previous year, but there Gidi’s and Yaakov’s students differed significantly, whereas Serge’s students, though their responses were similar to those of Yaakov’s students, did not differ significantly from Gidi’s.
The usefulness of the technology for learning

<table>
<thead>
<tr>
<th>Statement</th>
<th>Serge</th>
<th>Teacher</th>
<th>Gidi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I am capable of independent study.</td>
<td>3.31*</td>
<td>3.31*</td>
<td>3.32*</td>
</tr>
<tr>
<td>2. Living with computers and the Internet has made me an independent learner.</td>
<td>2.59</td>
<td>2.17*</td>
<td>2.68</td>
</tr>
<tr>
<td>3. My generation know better how to learn than the generation that did not have computers.</td>
<td>2.45</td>
<td>2.12</td>
<td>2.37</td>
</tr>
<tr>
<td>4. I am more aware of the potential of the Internet for learning than most of my teachers are.</td>
<td>2.86</td>
<td>2.77</td>
<td>2.26</td>
</tr>
<tr>
<td>5. It is important to use computers in history courses.</td>
<td>3.07*</td>
<td>2.62</td>
<td>2.42</td>
</tr>
<tr>
<td>6. When a topic interests me I would rather learn about it from the Internet than from books.</td>
<td>3.07*</td>
<td>2.73</td>
<td>2.37</td>
</tr>
</tbody>
</table>

Gains from school and history class

<table>
<thead>
<tr>
<th>Statement</th>
<th>Serge</th>
<th>Teacher</th>
<th>Gidi</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. I gained general knowledge from history class last year.</td>
<td>3.34*</td>
<td>3.77*</td>
<td>3.26*</td>
</tr>
<tr>
<td>8. I became better able to think and analyze thanks to my history class last year.</td>
<td>3.03*</td>
<td>2.81</td>
<td>1.84*</td>
</tr>
<tr>
<td>9. I was exposed to my people’s values and universal values in history class last year.</td>
<td>2.86</td>
<td>2.96*</td>
<td>2.26</td>
</tr>
<tr>
<td>10. School gives me tools for independent study.</td>
<td>2.75</td>
<td>2.23</td>
<td>3.42*</td>
</tr>
</tbody>
</table>

a. Cronbach alpha=0.693  b. Cronbach alpha=0.566 (0.691 excluding statement 10)
* p<0.05

Two polar attitudes towards the usefulness of digital technology. The classes’ responses to statements 4–6 concerning the usefulness of the digital technology resembled their responses to Yossi’s working procedure in part B. As can be seen in table 5, Gidi’s students disagreed with statements 4–6, Serge’s students agreed, and Yaakov’s varied with a slight tendency to agree. A one-way ANOVA test followed by a Scheffe post-hoc test revealed significant differences between Gidi’s and Serge’s students with respect to statement 6, that when a topic interests them they would rather learn about it from the Internet than from books. The differences between the responses to statements 4 and 5—that students are more aware of the potential of the Internet for learning than their teachers are (statement 4) and that it is important to use computers in history courses (statement 5)—were not found to be statistically significant, but the low p values (p=0.065 and p=0.05, respectively) can point to a similar trend of two polar attitudes towards the usefulness of the technology.

The relationship between perceptions of usability and perceptions of school and history. Statistically significant correlations were found between the reactions to Yossi’s working procedure, which relies exclusively on the Internet as a learning resource, and statements 3, 4, 5 and 6 (r=0.304**, 0.333**, 0.266*, and 0.398**, respectively; * means p<0.05 and ** means p<0.01), as well as among these statements, all of which concern the advisability of using the Internet and computers for learning purposes. Thus, students’ preferences for using computers and the Internet for learning purposes in or outside of school (statements 5 and 6, and the reactions to Yossi’s procedure) are correlated with the belief that computers make people better learners (statement 3) and that they are more aware of the educational potential of the Internet than their teachers are (statement 4).

In contrast, the reactions to Miriam’s working procedure were found to have a significant positive correlation with statement 10, that school provides tools for independent study (r=0.261*), as well as a significant negative correlation with statement 4, that students are more aware of the educational potential of the Internet than most of their teachers are (r=−0.290*). Thus, students’ decision to rely on books, as Miriam does, is correlated with legitimation of school learning practices. No other significant correlations were found between the reactions to Miriam’s and Yossi’s approaches and the statements in part C.
Discussion

Two polar perceptions of the usefulness of the technology for history classes

The findings depict two polar perceptions of the usefulness of computers and the Internet for history classes. At one extreme are students who think computers and Internet technology are more useful than books. These students reported that they would use the Internet to write a paper and objected to giving priority to books. These students also demonstrated a preference for use of the Internet as a learning resource when a topic interests them (statement 6) and thought it would be beneficial to use computers in history class (statement 5). Moreover, the perception of the usability of the Internet was positively correlated with the belief that the students know more about studying than their teachers do, at least with respect to the Internet (statement 4).

At the opposite extreme were students who found books more useful than the Internet for writing history papers (and thus identified more with Miriam’s approach than with Yossi’s). These students disagreed with statement 4 that the students are more aware of the educational potential of the Internet than their teachers and agreed that school provides tools for independent study (statement 10). In other words, they regarded school as a learning-to-learn venue.

Legitimation of school as an anchor for students’ perceptions of the technology

Interestingly, Gidi’s class, who are in the elite, learner-centred school, identified with Miriam’s approach more than with Yossi’s, whereas Serge’s and Yaakov’s classes, the public middle school and high school classes, respectively, identified more with Yossi’s approach than with Miriam’s. The difference is interesting because Gidi’s students experience learner-centeredness more than the two public-school classes and presumably—given the ready access to digital technology, youngsters’ affinity for learning with the technology, and the freedom inherent in learner-centred activities—should have expressed some of the rebelliousness of Papert’s assumed army of epistemological change (1996, 1998). Should they not, then, advocate procedures different from Miriam’s book-oriented method?

Apparently they do not; although they did not actually reject use of the Internet, they perceived it as ancillary to their method of doing schoolwork. Their responses to part A imply that this preference is not subject-dependent; rather, the computer is not central to any of the subjects in their curriculum. One might argue that Gidi’s school advocates the use of books and that Gidi’s students go along with the school message, perhaps because they want to belong to the prestigious school. If so, however, why do the students readily admit that they do not think they know better about the educational potential of the Internet than their teachers? Furthermore, no matter what message the school conveys, the fact is that Gidi’s students agreed that school provides them with tools for learning and did not consider themselves more knowledgeable with respect to the Internet. Thus, these students’ are not compliant; they are simply legitimizing school as a learning-to-learn venue. Living with computers thus has not changed these students’ epistemology in any crucial way. Presumably, they appreciate their school’s beyond-information agenda and are unaware of the learning potential inherent in partnering with the technology to pursue this agenda.

In contrast, a majority of students in the two public-school classes favoured Yossi’s procedure and believed they knew more about the educational potential of the Internet than their teachers. Given that the focus of school, and specifically history classes, is on information (e.g., Resnick, 2002; Salomon, 2000; Scardamalia & Berriter, 2003), the students will presumably find the Internet a useful learning resource. Furthermore, students would quite likely attribute the school’s decision to pursue its information-oriented agenda through books rather than through the Internet (or any digital medium) to ignorance of the potential of the Internet. Thus, keeping the school agenda focused on information leads students to propose the Internet as a better way to pursue this agenda, thus acting somewhat like Papert’s army of change. This claim is consistent with the fact that a majority of Serge’s and Yaakov’s students did not think school provided them with tools for independent study and it implies that the Internet influences students’ perceptions of school.

However, these students agreed that they had gained cognitive and analytical abilities from their history class in the same school. If history is focused on information, what are these abilities? Probably students believe that dealing with masses of information requires the ability to think and analyze; alternatively, perhaps they regard their teachers’ explicit instruction as teaching them how to deal with masses of information. Gidi’s students gave the opposite response, probably because of their impression of their school’s emphasis on beyond-information learner-centred...
activities. Specifically, the history classes they had experienced before coming to this school had resulted from an information-focused agenda and were thus less highly regarded.

The intellectual partnership between humans and computers as developed in history class

I claim that all the classes have failed to develop a fertile intellectual partnership with the computer. Gidi’s students do not have an intellectual partner; they carry their intellectual load by themselves. They could benefit from the technology. For example, authoring tools might help them organize and re-organize their thoughts while they are engaged in their projects. They could also use the Internet to give feedback to their peers, to get feedback from them, and to communicate with experts outside school. The fact that these tools were not mentioned strengthen Cuban’s argument that students’ exposure to technology is limited if the teacher does not advocate its use.

Serge’s and Yaakov’s students do share their intellectual tasks with digital technology; however, they merely use the technology for word processing and information searches and do not spontaneously form an intellectual partnership that would broaden their intellectual abilities. Furthermore, in an information-centred school, the technology might diminish students’ intellectual efforts. For example, an assignment to write summaries used to be an opportunity to develop the intellectual abilities required for dealing with the main knowledge resource back then—books. Due to the linear presentation of information in the books, the students invested mental effort in distinguishing the essence from the ancillary or their hand would have hurt from writing. These days this assignment can be accomplished without the need to invest such cognitive efforts because with the Internet one can rapidly retrieve the exact information needed from a vast number of sources. Furthermore, living in a world with easy access to information, students might get the impression that these homework assignments are all about information rather than learning to learn. (Of course, teachers can insist that students manually copy the material they find on the Internet into notebooks, as some of the interviewees in the preliminary research reported, thereby forcing them to summarize. This, however, is artificial; obviously students could simply find a shorter explanation of the material on the Internet. It would not prevent the students from perceiving schooling as merely creating a copy of someone else’s knowledge in their heads.)

Given that students used computers and the Internet very little and that history was one of the most commonly mentioned subjects for which the Internet is used, it may very well be that their responses would have been dramatically different for other subjects in the sense that they would believe that the Internet is more useful for the other subject. Further work is required, however, to explore students’ perception of the school’s general attitude towards computers and the Internet.

Conclusions

To some extent, the results of this study are reminiscent of the culture clash found in computer science education (Ben-David Kolikant, 2004; Ben-David Kolikant & Ben-Ari, 2008). In both cases, students’ informal experience with the technology led the students to believe that they knew more than their teachers. In computer science education, this belief led students to delegitimize what their teachers had to offer. For this reason, the results of this study, especially the belief among a significant number of public-school students that they know more about the educational potential of the Internet than their teachers, should alert the system. It may be losing its legitimacy as a learning-to-learn institution. However, more work is required to explore students’ attitudes towards school learning. Additionally, further work is needed to shed light on teachers’ attitudes towards students’ use of the Internet at home and its impact on students’ beliefs and schooling practices.

The results show that the Internet enters school through the back door. Specifically, it was evident that a significant number of students prefer using the Internet when they have a choice. Ignoring this fact could bring about a culture clash. Therefore, schools should rethink how to encourage students to engage in a fertile intellectual partnership with the technology to enhance their learning.

Acknowledgement

This study was funded by the Israel Science Foundation, grant 1108.
References


A Markov-based Recommendation Model for Exploring the Transfer of Learning on the Web

Yueh-Min Huang1, Tien-Chi Huang1, Kun-Te Wang1 and Wu-Yuin Hwang2

1Department of Engineering Science, National Cheng Kung University, No.1, University Road, Tainan City, Taiwan, R.O.C. // Tel: +886-6-2757575 ext. 63336 // Fax: +886-6-2766549 // huang@mail.ncku.edu.tw // kylin@easylearn.org // taito@easylearn.org
2Graduate school of Network Learning Technology, National Central University, No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan, R.O.C. // wyhwang@cc.ncu.edu.tw

ABSTRACT

The ability to apply existing knowledge in new situations and settings is clearly a vital skill that all students need to develop. Nowhere is this truer than in the rapidly developing world of Web-based learning, which is characterized by non-sequential courses and the absence of an effective cross-subject guidance system. As a result, questions have arisen about how to best explore and stimulate the transfer of learning from one subject to another in electronically mediated courses of study. In this study, we argue that online learners would benefit from guidance along applicable group-learning paths. This paper proposes use of the learning sequence recommendation system (LSRS) to help learners achieve effective Web-based learning transfer using recommendations based on group-learning paths. We begin with a Markov chain model, which is a probability transition model, to accumulate transition probabilities among learning objects in a course of study. We further employ an entropy-based approach to assist this model in discovering one or more recommended learning paths through the course material. Statistical results showed that the proposed approach can provide students with dependable paths leading to higher achievement levels, both in terms of knowledge acquisition and integration, than those typically attained in more traditional learning environments. Our study also identified benefits for teachers, providing them with ideas and tools needed to design better online courses. Thus, our study points the way to a Web-based learning transfer model that enables teachers to refine the quality of their instruction and equips students with the tools to enhance the breadth and depth of their education.

Keyword
Groups-learning paths, Transfer of learning, Markov chain model, Individual learning, Entropy.

Introduction

Recent advances in network bandwidth availability have increased use of Web-based education. In Web-based educational settings, information and communication technology (ICT) provides access to various resources, such as audio recordings, videos, images, and text. ICT has therefore had a dramatic impact on both learning processes and outcomes. For learners, the need for large amounts of information is increasing. However, not every learner can effectively organize and choose between these learning resources. A good learning management system (LMS) should provide learners with guidance that will assist them to make the best choices in their management of the available resources. Therefore, the selection of learning contents and recommended learning sequences have been recognized as two of the most interesting research topics in Web-based learning systems (Chen et al., 2005; Tseng et al., 2007).

Sharable Content Object Reference Model (SCORM) a model of web-based learning content and sequence standards, was proposed in 1997 at the 4th International forum on research and technology advances in digital libraries. This model was not innovative, but combined many existing international standards, such as the Content Packaging, Simple Sequencing (IMS, 2003), and LOM (IEEE LTSC, 2002). According to SCORM, learning contents, such as multimedia, images, slides or others learning objects, can be packaged and organized based on the SCORM standards to form a learning sequence. Notably, the SCORM-based learning sequence (ADL, 2004) offered a way to string several learning objects together and design instructions easily. In spite of the digital content learning sequence specifications instituted by SCORM, the pedagogical meaning of a learning sequence also needs to be recognized in e-Learning environments.

Generally speaking, the goals of education are not only the transfer of skills and knowledge in one context, but also the transfer and application of those skills and knowledge to other contexts. “Transfer of learning”, which refers to the expansion and generalization of learning outcomes, is one major criterion of learning efficacy. The theory of
transfer of learning was introduced by Thorndike and Woodworth (Thorndike & Woodworth, 1901). They explored similarities in the features of transfer of learning in different contexts. The theory asserted that effective transfer of learning depends on identical elements shared between the learning task and the transfer task. Several studies (De Corte, 2003; Haskell, 2001; Richman-Hirsch, 2001) have noted that transfer of learning is fundamental to learning because it generally aspires to impact on contexts quite different from the context of learning (Perkins & Salomon, 1992). In Web-based learning settings, however, there are two limitations that obstruct the transfer of learning. First, there is no guide to assist learners to conduct cross-subject transfer of learning in a Web-based learning environment. Second, while considerable attention has been paid in the past to research issues related to Web-based course design (Chan et al., 2006; Chen et al., 2005; Kabassi & Virvou, 2003; Zhu, 2007), the designer seldom takes the transfer of learning into consideration during the process of course design. Therefore, learning transfer is a significant issue for educators to take into consideration during the course-designing process for Web-based learning environments.

As noted before, transfer of learning is the process whereby skills learned in one situation or under one set of conditions are demonstrated in a different situation or under a different set of condition (Borich & Tombari, 1995). However, since transfer effects are not concrete, how to achieve and evaluate transfer of online learning is a challenging issue. In this study, we argue that transfer of online learning can be effectively achieved from group-learning sequences. Intrinsically, a learning sequence which describes a learner moves from one context to another is a concrete presentation of transfer of learning. Herein, constructing a learning sequence can be regarded as the process of completing a set of several learning transfers. Thus, developing an ideal learning sequence may help to trigger the transfer of learning. On these grounds, the purpose of this paper is to explore the idea that certain learning sequences could trigger transfer of Web-based learning. By developing a learning sequence recommendation system, the present study proposes the idea of applying an online learning sequence to trigger the transfer of Web-based learning.

This study proposes that the best way to deal with learning transfer exploration is by analyzing group-learning experiences in the field of intelligent tutoring systems (ITS). In this study, the group-learning experience is the sequence pattern from the different subjects’ (contexts) learning. The proposed system, the Learning Sequence Recommendation System (LSRS), analyzes group-learning experiences to predict and provide a personal learning list for each learner by tracking others’ learning patterns regarding certain topics. This will provide learners opportunities to improve their transfer of learning. For example, some learners have studied the course “Management Information System: MIS”, and then moved on to enroll the course “Data Structure”. It is clear that both courses are in different domains. Since both courses are not closely correlated in terms of course continuity, it’s difficult to achieve the integration in learning and the transfer of learning. So far as this problem is concerned, LSRS will provide a relationship, which is represented as the same concept across the two different domain subjects.

In this paper, we propose a novel learning mechanism by using the Markov chain model to analyze the prediction of time-series recommendation problems. The Markov chain model is the simplest mathematical model for random phenomena evolving in time and it has been widely used in sequence prediction and navigation analysis (Heller et al., 2004; Rajapakse & Ho, 2005; Sarukkai, 2000; Tian et al., 2000). However, few studies have used the Markov chain model to predict the recommended learning sequence. Using this model, this study focuses on exploring dependable group patterns. The outcome of the Markov chain model can also illustrate the structure of the learning sequences. These explored learning sequences have strengths that may be reinforced by the responses of the learners, which provide the LMS with suitable recommendations as to which sequences facilitate effective and efficient learning.

Additionally, using ranking techniques, we can also record learners’ feedback for sorting group learning paths. We propose an entropy-based approach to automatically compute the average entropy value of each predicted learning sequence and estimate its quality as well. Consequently, the system provides a re-ranked recommendation list for the learner to create an adaptive learning sequence recommendation, which then provides him/her with the guidance of learning transfer.

The main goal of this paper is to investigate the impacts of group-learning patterns on the transfer of learning via a proposed learning system that has ability to provide each learner with an adaptive learning sequence list. As mentioned earlier, the following two research questions in web-based learning settings need to be answered:

- Could online learners achieve appropriate transfer of learning by recommended learning sequences?
Could recommended learning sequences provide an e-Learning course designer with guidance or suggestions to design online courses, which should possess effects of transfer of learning?

The next section describes the mathematical methodology used in this study to examine group-learning patterns. The recommendation prediction mechanism is also proposed, which uses an entropy-based approach to re-rank the recommended results to accommodate each learner with an appropriate sequence list. An illustrated example is given to explain the recommendation process in the following section. After that, the developed learning sequence recommendation system (LSRS) description and the diagrams of manipulation are displayed. The evaluations are then presented, with a thorough description of the participants in the research and of the instruments and procedures used. Finally, the results are discussed and the conclusions are drawn.

The proposed model for predicting and recommending dynamic learning sequences

In this section, we first describe the multi-subject learning sequencing model that has been adopted to define the sequencing model in our study. Then, we detail the learning sequences prediction and recommendation mechanisms in subsections.

Multi-Subject Learning Sequencing Model

The concept-mapping technique is usually applied to show the structure of a specific knowledge field. It uses a graphic representation of the knowledge base so that learners can easily see the entire learning structure at a glance. So far, many studies that discussed concept-mapping only focused on a single subject or domain (Chen et al., 2008; Hsu & Hsieh, 2005; Tseng et al., 2007). However, a general learning process requires learners to go through several subjects to achieve a learning goal. For example, if a learner needs to learn the “Markov Chain Model”, he or she should possess prerequisite knowledge about subjects “Direct graph”, “Probability and Statistics”, “Engineering Mathematics”, and “Discrete Mathematics” (Figure 1). Hence, a learning map associated with several subjects has the capability of clearly presenting the entire learning process. From Figure 1, it is evident that some courses or learning objects are connected to each other, so a learner can recognize and follow the appropriate learning sequence.

![Figure 1. A number of prerequisites and chapters for a course on the “Markov Chain Model”](image-url)
We observe that nowadays many students cannot apply what they have learned, or are learning to existing problems or contexts, and many students cannot appreciate the applicability of their reading materials. That is to say that the transfer of learning seems not to be effectively present among students. Through the multi-subject learning sequencing model, learners can clearly recognize the relationships between several courses and see how to apply what they are learning to other fields or courses. In addition, by seeing the connections between the subject being studied and the subjects in which they experienced the most success, they are better informed in choosing other related subjects to study in the future.

Definition

The application of the Markov chain model in a learning system has the advantage that the likelihood of the group-learning patterns is simply a product derived from the collective likelihoods of individual instances. It can provide learners in need of assistance with several learning sequences that are based on actual learners’ experiences, and in accordance with group-learning patterns. In a Web-Based learning environment, an LO is first used by individual learners as a learning course related to a topic or issue. LOs are placed in a learning repository and the learners’ profiles are recorded in a transaction database. Thus, we can retrieve the learning data of each learner, which can be viewed as a learning sequence. Moreover, we give a definition that an LO, denoted as $O_j$, belongs to the record of a learner, and each learner produces their own learning sequence $<O_1,...,O_n>$ during the learning process for learning a specific concept. Furthermore, the dynamic learning sequence can be formally defined as:

**Definition 1: Dynamic Learning Sequence (DLS)**

Let $O$ be a set of LOs from a learning repository. Given a set of LOs $X = \{x_1, x_2, ..., x_m\}$, $x_i \in O$, $1 \leq i \leq m$, then the learning sequence at time $t$ is denoted as $LS_t = x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow ... \rightarrow x_i$, and the end is $x_i \rightarrow x_j$ at the time $t+1$, where $1 \leq i \leq l$.
Analysis of sequences prediction mechanism

First, we illustrate the learning mechanism, which is useful for the learning platform. All of the learning activities of learners on the platform can be monitored and dealt with as a transaction database, and a transition matrix expresses all of the sequences of learners from there. Thus, the matrix can be seen as a base set of states, which accumulates the probability of each node (as an LO) using the Markov chain model. We utilize the model to provide the predicted learning sequences. Additionally, by using an entropy-based approach, we turn ranked learning sequences into recommendation lists, which will be described in a later section. For individual learning, a weighting scheme is utilized to determine the final recommended sequences. We ask learners to write down opinions for feedback in order to adjust the weight values of the recommended sequences and reorder them by using a combination sort function. The methodology of the proposed learning mechanism is shown in Figure 2.

Transition Matrix

The learning sequence of each learner is obtained from the learning objects that reside in the learning corpus. Next, the sequences of all learners are used to form a transaction dataset. To signify the LS of each learner, we give an n-dimensional matrix $A$ including transition probabilities from the transaction dataset, which is positive and non-irreducible. The transition matrix is shown in Figure 3.

$$
\begin{bmatrix}
0 & p_{ab} & p_{ac} & \cdots & p_{an} \\
p_{ba} & 0 & p_{bc} & \cdots & p_{bn} \\
p_{ca} & p_{cb} & 0 & \cdots & p_{cn} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
p_{na} & & & \cdots & 0
\end{bmatrix}
$$

Figure 3. States transition matrix

In Figure 3, the $p_{ij}$ is the transition probability from one state to the next state. Notably, in the matrix $A$, the diagonal elements all are zero, that is, there is no self-loop in the graph.

Markov Chain Model

The state spaces of LS are a case of the Markov chain model, which is to say, a chain of random LOs in which each subsequent LO depends only on the current LO. Thus, by estimating the learning process of the $k$-th learner from their learning profiles $X^k = \{x_0, x_1, \ldots, x_n\}$, we can obtain sequences of different lengths in state $i$. All learners have their own learning sequence, which is shown as follows:

$$
\begin{align*}
x_0^{(i)} \rightarrow x_1^{(i)} \rightarrow x_2^{(i)} \rightarrow \cdots \rightarrow x_m^{(i)} \\
\vdots \\
x_0^{(k)} \rightarrow x_2^{(k)} \rightarrow x_3^{(k)} \rightarrow \cdots \rightarrow x_n^{(k)}
\end{align*}
$$

If a learning sequence includes duplicate learning objects, it belongs to a repetitive learning, otherwise, it is held by a progressive learning.
If the current state is \( i \) and the next state is \( j \), the transition probability is represented as

\[
 p_{ij} = \Pr\{x_{t+1} = j \mid x_t = i\}, \quad t \in Z^+ ,
\]

where \( p_{ij} \geq 0 \) and \( \sum_{j=1}^{n} p_{ij} = 1, i = 1, 2, ..., n \). According to the Chapman-Kolmogorov equations (Gross & Harris, 1998), which are shown as formula (1), we can easily find the matrix \( P^{(n)} \) formed by the elements \( p_{ij}^{(n)} \).

\[
 p_{ij}^{(n)} = \Pr\{X_{n+m} = j \mid X_m = i\}, \quad \text{and} \quad p_{ij}^{(m)} = \sum_{k=0}^{m} p_{ik}^{(n)} p_{kj}^{(m)}
\]  

According to equation (1), the matrix \( P^{(n)} \) can be conducted as multiplying the matrix \( P \) by itself \( n \) times. That is, \( P^{(n)} = P \cdot P \cdot \cdots \cdot P = P^n \).

**Stability of the learning sequence**

In order to discuss the stability of the learning sequence, we need to consider the long-term behavior of Markov chains. This means that we need to find the steady-state probabilities of the Markov Chain after a long period of time. Therefore, we first consider a discrete Markov chain, which is ergodic and represented as equation (2).

\[
 \lim_{n \to \infty} p_{ij}^n = \pi_j, \quad \forall i, j \geq 0
\]  

where \( \pi_j \) is the steady state distribution. Also, it is independent of initial probability distribution and exists uniquely with the state. In order to verify whether the transition achieves equilibrium, \( \pi_j \) can be checked using equations (3) and (4).

\[
 \pi_j = \sum_{i \in S} \pi_i p_{ij}, \quad j \geq 0
\]  

\[
 \pi e = 1
\]

where \( \pi_i \) stands for the initial probability that \( X_0 = i \). Equation (3) indicates that if the transition approaches the steady state, the distributions will not change repeatedly. Equation (4) is a vector notation where \( \pi = (\pi_0, \pi_1, ...) \) represents the limiting probability vector and \( e \) is a matrix whose elements all equal one. It implies a boundary condition (i.e. \( \sum_j \pi_j = 1 \)).

**Adaptive model of learners’ feedback**

After each learning process, learners will give feedback information to the LMS. According to the given feedback, the adaptive model will enhance the choosing probability of the learning sequence. This model provides a weight for each learning sequence. By changing these weights, the rank of learning sequences will also change. From Figure 4, \( Y = f(S_1, S_2, ..., S_n \mid \Phi) \), where \( f \) function is a combining function that sorts the weighted learning sequences, \( S_i \) represents the \( i \)th learning sequence and \( \Phi \) is the set of weights, \( w_1, w_2, ..., w_n \). If the feedback from the learner is positive, the weight of the learning sequence is increased, and vice versa.
The advantage of the adaptive model is that the rank of learning sequences not only follows the training data, but it is also evaluated by feedback from other learners.

**The entropy-based approach for learning sequence recommendation**

Having obtained the steady probability transition matrix, one has also acquired several steady learning sequences. In this subsection, we propose an entropy-based approach for sorting these group learning sequences so that they can be substantiated as dependable learning recommendations for learners.

Since a learning sequence consists of a few learning objects, the links among learning objects are important clues for determining the significance of the sequence. Therefore, it is necessary to design an approach to discover the most informative learning sequences. First of all, Shannon’s information entropy (Cover & Thomas, 1991) is applied to calculate the link entropy based on the above-mentioned steady state matrix. According to the definition of entropy, it can be expressed as $-K \sum_{i=1}^{n} p_i \log p_i$, where K is a positive constant which is set to one in this study. $p_i$ is probability of $i$th event and n is the number of events. To calculate the average information within a learning sequence, the equation is expressed as follows:

$$E(LS_i) = \left( -\sum_{j=1}^{n} p_j \log p_j \right) / n$$  \hspace{1cm} (5)

where $p_j$ is the probability from one LO to another, which can be obtained from the above-mentioned steady transition matrix, and n is the number of links within a learning sequence. According to the definition of entropy, the most informative learning sequence should have the least entropy value. Therefore, the equation used to evaluate the amount of information of learning sequence $I_{LS_i}$ is used to define and represented as follows:

$$I_{LS_i} = 1 - E(LS_i)$$  \hspace{1cm} (6)

After calculating the average information amounts of the learning sequences, the adaptive model is applied to extract more informative learning sequences based on learners’ feedback. As mentioned in section 2.3, every learning sequence has a given weight that is regulated by learners’ feedback. The feedback is either positive or negative and determines whether the weight value is increased or decreased. The combination function is utilized to sort the every weighted entropy set $\{w_1I_{LS_i}, w_2I_{LS_i}, ..., w_nI_{LS_i}\}$. Afterwards, LMS will recommend the re-ranked list to the learners. Through the repeated feedback process, a more precise learning sequence list will be provided to the learners.
An Illustrative Example

We now present an example that demonstrates how we can find the steady state probability distribution. For simplicities sake, we present four learning objects, a, b, c and d, with 400 learners in this learning process. Table 1 shows the initial state and the first state shifting.

<table>
<thead>
<tr>
<th>Learning Objects(Learners before shifts)</th>
<th>Learning objects shifts</th>
<th>Learners after shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>a (100)</td>
<td>0  40  40  20  0  40  60  25  125</td>
<td></td>
</tr>
<tr>
<td>b (100)</td>
<td>40  0  30  30  40  30  0  20  85</td>
<td></td>
</tr>
<tr>
<td>c (100)</td>
<td>60  20  0  20  40  30  0  50  120</td>
<td></td>
</tr>
<tr>
<td>d (100)</td>
<td>25  25  50  0  20  30  20  0  70</td>
<td></td>
</tr>
<tr>
<td>total (400)</td>
<td></td>
<td>400</td>
</tr>
</tbody>
</table>

In the initial state, each learning object has 100 learners. After one shift we assume that the learners will not stay at the same learning object and, as a result, the number of learners staying at the same object is indicated by zero. To find out the number of learners at each learning object, the matrix is transposed and the sum of each row is calculated. After the first shift, the number of learners staying at the learning objects b and d decreases, while that of learners staying at a and c increases. As time passes, the transition will approach equilibrium. In Table 2, we can observe the transition process gradually going to steady state.

<table>
<thead>
<tr>
<th>Learning Objects</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0  1  2  3  4  5  6  7</td>
</tr>
<tr>
<td>a</td>
<td>100 125 110 98 105 115 122 122</td>
</tr>
<tr>
<td>b</td>
<td>100 85 95 100 94 85 90 90</td>
</tr>
<tr>
<td>c</td>
<td>100 120 110 103 110 111 113 113</td>
</tr>
<tr>
<td>d</td>
<td>100 70 85 99 87 80 74 74</td>
</tr>
</tbody>
</table>

Table 2. Approaching a steady state

Figure 5. The instance of the probability transition matrix of learning objects trends towards to a steady state.
The figures 5a and 5b illustrate the transient and steady states respectively. In Figure 5(a), the number of learners is mapped to state 5 (described in Table 2) while Figure 5(b) is mapped to the above-mentioned states 6 and 7. Before state 6, the number of learners is in a transient state, since the number of learners in any one learning object will vary. The number of learners, however, reaches equilibrium after state 6, implying that the amounts of inflow and outflow between two learning objects are the same. It also demonstrates that the learning process has achieved a steady state.

From the steady state table (see Table 2), the steady state distributions can be computed as \( \pi_a = 0.305 \) (=122/400), \( \pi_b = 0.225 \), \( \pi_c = 0.283 \) and \( \pi_d = 0.185 \). As the transition approaches a steady state, we can also compute the steady probability transition matrix. Table 3 shows the first shift in transition probability, and the summary of each row satisfies the constraint \( \sum_j p_{ij} = 1 \).

**Table 3. Probability Transition Matrix at first shift**

<table>
<thead>
<tr>
<th>Learning Object</th>
<th>To a ( 0 )</th>
<th>To b ( 40/100=0.4 )</th>
<th>To c ( 40/100=0.4 )</th>
<th>To d ( 20/100=0.2 )</th>
<th>( \sum_j p_{ij} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td></td>
<td>40/100=0.4</td>
<td>40/100=0.4</td>
<td>20/100=0.2</td>
<td>1.0</td>
</tr>
<tr>
<td>( b )</td>
<td>40/100=0.4</td>
<td>0</td>
<td>30/100=0.3</td>
<td>30/100=0.3</td>
<td>1.0</td>
</tr>
<tr>
<td>( c )</td>
<td>60/100=0.6</td>
<td>20/100=0.2</td>
<td>0</td>
<td>20/100=0.2</td>
<td>1.0</td>
</tr>
<tr>
<td>( d )</td>
<td>25/100=0.25</td>
<td>25/100=0.25</td>
<td>50/100=0.5</td>
<td>0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

By the same procedure, the remaining transition probabilities can be calculated. Hence, we can obtain the steady probability transition matrix, which is able to assist in the prediction of the learning sequence associated with a new learner. The steady probability transition matrix can be obtained from the following table.

**Table 4. Steady Probability Transition Matrix**

<table>
<thead>
<tr>
<th>Learning Object</th>
<th>To a ( 0 )</th>
<th>To b ( 0.45 )</th>
<th>To c ( 0.28 )</th>
<th>To d ( 0.26 )</th>
<th>( \sum_j p_{ij} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td></td>
<td>0.45</td>
<td>0.28</td>
<td>0.26</td>
<td>1.0</td>
</tr>
<tr>
<td>( b )</td>
<td>0.304</td>
<td>0</td>
<td>0.376</td>
<td>0.31</td>
<td>1.0</td>
</tr>
<tr>
<td>( c )</td>
<td>0.731</td>
<td>0.124</td>
<td>0</td>
<td>0.14</td>
<td>1.0</td>
</tr>
<tr>
<td>( d )</td>
<td>0.165</td>
<td>0.226</td>
<td>0.609</td>
<td>0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

The user can then give source and destination LOs so that the system can provide the recommendation list. In this example, if the user sets LOs \( a \) and \( d \) as source and destination respectively, the system will calculate the average entropies associated with the learning sequences from \( a \) to \( d \). Table 5 shows learning sequences and their average entropy values, which are calculated according to the steady transition probabilities in Table 4. For instance, the average entropy of sequence \( a \rightarrow b \rightarrow c \rightarrow d \) can be computed by the entropy equation (Eq. 5).

\[
E(\text{LS}) = \frac{(-0.45 \times \log 0.45) + (-0.376 \times \log 0.376) + (-0.14 \times \log 0.14)}{3} = 0.81
\]

**Table 5. The average entropy of each learning sequence**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a \rightarrow b \rightarrow c \rightarrow d )</td>
<td>0.81</td>
</tr>
<tr>
<td>( a \rightarrow c \rightarrow b \rightarrow d )</td>
<td>0.61</td>
</tr>
<tr>
<td>( a \rightarrow b \rightarrow c \rightarrow d )</td>
<td>0.639</td>
</tr>
<tr>
<td>( a \rightarrow c \rightarrow d )</td>
<td>0.654</td>
</tr>
<tr>
<td>( a \rightarrow d )</td>
<td>0.505</td>
</tr>
</tbody>
</table>

Finally, equation (6) is used to calculate the information amount in each learning sequence. Furthermore, the adaptive model is employed to evaluate the effect of the learners’ feedback on the recommended sequences. In this example, the weights are all initially set to 0.5. Therefore, the ranked recommendation list can be obtained shown as Table 6 (e.g. The weighted information amount of the sequence \( a \rightarrow d \) is computed as 0.5*(1-0.505) = 0.248, where 0.505 is the average entropy of the sequence).
### Table 6. The ranked learning sequences with weighted information amount

<table>
<thead>
<tr>
<th>Re-ranked learning sequences</th>
<th>Weighted Information amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a \rightarrow d$</td>
<td>0.248</td>
</tr>
<tr>
<td>$a \rightarrow c \rightarrow b \rightarrow d$</td>
<td>0.195</td>
</tr>
<tr>
<td>$a \rightarrow b \rightarrow d$</td>
<td>0.181</td>
</tr>
<tr>
<td>$a \rightarrow c \rightarrow d$</td>
<td>0.173</td>
</tr>
<tr>
<td>$a \rightarrow b \rightarrow c \rightarrow d$</td>
<td>0.095</td>
</tr>
</tbody>
</table>

It can be shown that the more information a learning sequence has, the more weighted information amount value becomes. It must be noted that the provided learning sequences present the progressive learning type; that is, repetitive learning information is eliminated from the recommendation list.

### System Development

Our proposed system was incorporated in a learning portal run by lecturers to create learning contents and manage course processes, as well as to offer learners a useful learning sequence suited to their current individual learning status. These courses are based on the specifications of SCORM; that is, the contents are packaged into learning objects by specific authoring tools and reorganized into the course structure as related to a topic. The learning process is comprised of system portal logging, learning sequencing exploration, and the delivery of suitable learning objects; after which, the metadata is generated. This is be demonstrated in the following subsections.

#### Learning interface

First, we launched a learning portal for a dynamic learning sequences model, which provides a personalized learning interface for learners. After registration, learners enter this system using their user identity and password. The portal offers learners a set of comprehensive learning services such as personal information, a personal directory, course lists, and reference lists. When learners’ lists are available, they can study each course by clicking the courses’ hyperlink. Figure 6 presents an example of the layout of a learning course. It is noteworthy that each course was produced using the standardized course generation process (Huang et al., 2008). Hence, each course in our learning platform corresponds with the SCORM standard. Several multimedia file types, such as streaming video and audio files, Macromedia Flash files, and Microsoft PowerPoint files are included within a designed e-learning course.

![Figure 6. The layout of a course presentation](image)
Learning sequence recommendation system (LSRS)

From the portal login page, learners can proceed to the learning sequencing exploration. As shown in Figure 7, the interface of LSRS consists of three operation areas: the source subject inputting area is located in the upper-left side, the presentation area for recommended subjects and the concurring concept list is located on the left side, and the active area, which represents recommended learning sequences, is located on the right side of the window.

![Figure 7. Annotated screenshot of the learning sequence recommendation system](image)

The source subject specifies the source course of the learning sequences. After inputting the source subject and pressing the Start Button, the system will provide recommended learning sequences. A recommended list displays all subjects contained in the recommended sequences. For example, if the recommended sequences contain four topics, SQL, Programming, Java, and JDBC (like that shown in Figure 7), the learner selects a prerequisite topic, SQL, as the starting point. Using information gained by tracking the groups’ learning (i.e. learning sequences), the learner acquires a recommended sequence list that consists of these four topics. At this time, the learner clearly sees the relationship between the recommended subjects and the source subject and chooses these subjects for study.

The concurring concepts are shown in the concept list while the learner selects one of recommended subjects. These concepts are very useful for learners in finding the same concepts shared between two subjects, which is the main concern in transfer of learning. This is to say that the system applies the concept list to help learners find the relations between two subjects for achieving transfer of online learning. However, the system does not point out the locations of shared concepts in the learning object. The goal is to train learners to carefully find out where the shared concepts located in and narrowly learn them. This is also a part of whole learning process. Recall our earlier example, in which a learner gets a recommended list, he can see the concurring concepts after selecting a subject ‘SQL’. When he selects another subject ‘Java’, he finds a concurring, concept ‘Database Access’, shared between the two subjects. At this time, he needs to ascertain and learn the relationship between the learning objects.

Moreover, the most important information is shown in the largest field, which is called the active area for the presentation of learning sequences. It represents some of the successive paths, which were calculated by the above-mentioned process and are listed from top to bottom. The learners can easily see the inferred learning sequences after the start button has been pressed. The source subject drawn in the active area is denoted as the root. A number is also given to each sequence, which stands for the priority of the recommended sequences. Of course, the learner could re-
select different source subjects according to their own needs or interests. In the active area, the learners can also click expansion and contraction symbols to expand or condense the sequences.

In addition, after the learning process, learners can express their own opinions regarding the recommended results and give the feedback to the system, using one of three rating options: positive, negative, or undecided. Figure 8 demonstrates a feedback-giving example. In this manner, the score of the learning sequence can be recalculated, which is beneficial for recommendation to other learners. For further enhancing the transfer of learning between specific two subjects, any new relationship found from their own studying could also be recorded in the feedback area. The recommendation result metadata is exportable in XML format (see Figure 8) so that it may be imported into other SCORM-based learning platforms. The main benefits of the metadata focus on two points: operational efficiencies of the learning portal, including reuse and curriculum analysis, and interoperability of course materials among different learning portals.

In essence, this work establishes explicit guidelines for group paths to ensure that learning objectives are met and to ensure that the learners’ honesty and integrity are protected by the transparency of the process. Besides outlining such learning sequences, the learners can use the recommendations to manage their learning plan, to guide their study of how to fit course material into the “bigger picture” of real world applications and implications, and to generally widen and deepen the quality of their learning.

Research Methodology

Research Design

The purpose of the learning sequence recommendation system is to assist learners to achieve effective transfer of learning in web-based learning settings. As noted above, since learning transfer is a crucial educational concern, and Web-based learning is often non-sequential, guiding online learners to achieve effective transfer of learning is an important research problem. With the aid of LSRS, this study needed to investigate whether valuable learning sequences are beneficial for transfer of learning, and to further assist transfer of learning even if there is no guide in Web-based learning settings.
In this study, forty subjects were monitored while using the Web-based learning system (personal learning portal site) during the 2005 fall semester. Each experiment participant was asked to complete three courses during the semester, and each course lasted for six weeks. That is to say total learning process on the Web lasted for 18 weeks. To estimate the effectiveness of LSRS in assisting learners to achieve transfer of learning with regard to Web-based learning, a randomized subject control group design was used. The participants were randomly assigned into the experimental group and control group respectively, after they submitted registrations on the learning system. The only dissimilarity between the two groups is that the experimental group learners were given recommended learning sequences along with their corresponding primary concepts via LSRS. In other words, the second and the third subjects that the experimental group learners studied were decided by LSRS, while control group learners followed their inclinations to choose three subjects without concept suggestions. To evaluate the learning outcome and the usability of the LSRS, a post-test was conducted in both groups after the 18th week. The post-test consisted of the evaluation examination and the questionnaire introduced in the instrument section.

Participants

The experiment was performed in one large university in Taiwan, and the target population was comprised of 113 college sophomore students and 23 teachers. 76 of the student participants were male and 37 were female. The reason for choosing sophomore students is that they are in the process of learning a number of fundamental courses, which are the foundations of higher-tier, more specialized courses and fields of knowledge. 57 students served as the experimental group, which studied online learning courses and were asked to follow the learning sequences recommended by LSRS throughout the experiment period. The other students (56 persons) served as the control group, which was exposed to the same online learning environment as the experimental group, without the aid of LSRS.

Instruments

In this section, we evaluate the effectiveness of the proposed models and the usability of the learning sequence recommendation system. In order to evaluate the effectiveness and impact of using LSRS, and its assistance to course designers for designing online courses, two different research instruments were devised, one for learners and one for course designers. The following two subsections introduce these research instruments in more detail.

Evaluation system

This section describes how we measured the transfer of learning after the whole learning process was complete. The devised evaluation system (see Figure 9) was used to evaluate whether the effective transfer of learning of learners who used the LSRS was achieved in the web-based learning setting. The instructions section explains how the system is used. There are three more sections in the evaluation system including a subject selection section, a concept selection section, and an explanation section. In part 1, learners chose the three subjects that they had studied during the experiment. Next, in part 2, the most important or impressive concepts associated with each course were chosen. Finally, learners had to explain the reason why they chose these concepts and whether relationships exist among these concepts. Each learner’s results were evaluated by four teachers, who were all experts with regards to their respective subjects. To avoid the experiment effect, the scorers were not told which group participants belonged to. Teachers graded with regard to learners’ responses in the explanation section. Students’ responses included three sections, solidity, relation, and rationality. The solidity section described the completeness of the description given regarding the chosen concepts; the relation section measured whether a relationship existed between the chosen concepts; and the rationality section indicated the accuracy of the explanation given regarding the chosen concepts. As mentioned before, the evaluation system was conducted after an 18 week learning period and the evaluation process lasted for two weeks.

Questionnaire

In order to evaluate the usability of LSRS for learners, and its effect on course design for teachers, two semi-open ended questionnaires were devised. In addition to the questionnaires a Likert-type scale was adopted to measure
learners’ and teachers’ degree of satisfaction, ranging from strongly approve to strongly oppose. Each response was assigned a corresponding score as follows: 5 for strongly approve, 4 for approve, 3 for neutral, 2 for oppose, and 1 for strongly oppose. The participants were required to write down the reason for their choice, with a two hundred word limit, so that we could gain a better understanding of what students received from, and how they felt about their interactions with, the system. Tables 8 and 9 respectively show the questions and results of the semi-open ended questionnaires.

![Figure 9. Screenshot of devised evaluation system](image)

Regarding teachers, question 1 in Table 8 evaluates whether or not LSRS can provide effective learning transfer for teaching in the Web-based learning environment. In question 2, we assess whether the recommended learning sequences constructed knowledge of the subjects from the teachers’ perspective. Then, in the third question, we assessed whether or not teachers referred to the learning transfer information for designing future courses. Finally, according to the results assessed in the evaluation system, the teachers told us whether or not LSRS enhanced students’ transfer of learning.

For the learners, the first two questions in Table 9 assess the usability of the proposed system, which includes the clearness and understandability of the user interface interaction and the learning sequences. The third question evaluates whether or not the learning sequences provided by LSRS rendered the connections between the existing knowledge and new subjects, and the fourth question evaluates whether or not the recommended learning sequences offered a macro view, similar to a learning map, during the curriculum learning process.

**Results and discussion**

We first compared the evaluation scores of the two groups, which each comprise three independent scores, solidity, relation, and rationality. As mentioned above, according to the results from the evaluation system, a group of four teachers who are experts in the subjects graded the scores. An independent-sample t-test was conducted to detect any significant difference between the two groups. The next part of the analysis used the questionnaires to examine the potential of LSRS to promote learning transfer and curriculum design. We then analyzed the perceptions of the experimental group learners toward the usefulness of LSRS. Finally, we separated our findings regarding learning transfer by gender.
The potential for LSRS to promote learning transfer and curriculum design

As shown in Table 7, data analysis of the evaluation scores with independent-samples t-tests were almost all statistically significant (Evaluation score: \(t(111)=-9.588^*, p<0.05\); relativeness: \(t(111)=-13.308^*, p<0.05\); rationality: \(t(111)=-11.360^*, p<0.05\)), except for the solidity score (\(t(111)=.274, p=.785\)).

<table>
<thead>
<tr>
<th>Score</th>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation score</td>
<td>Control</td>
<td>56</td>
<td>65.82</td>
<td>-9.588*</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>57</td>
<td>75.09</td>
<td></td>
</tr>
<tr>
<td>Solidity</td>
<td>Control</td>
<td>56</td>
<td>77.46</td>
<td>.274</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>57</td>
<td>77.00</td>
<td></td>
</tr>
<tr>
<td>Relativeness</td>
<td>Control</td>
<td>56</td>
<td>69.59</td>
<td>-13.308*</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>57</td>
<td>84.37</td>
<td></td>
</tr>
<tr>
<td>Rationality</td>
<td>Control</td>
<td>56</td>
<td>59.98</td>
<td>-11.360*</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>57</td>
<td>74.25</td>
<td></td>
</tr>
</tbody>
</table>

\* \(p<0.05\)

On further investigation we find that the experimental group gets a higher score in the relation aspect (\(M_{\text{exp}}=84.37; M_{\text{control}}=69.59\)). The results may be explained by the assistance of LSRS. Since the recommended learning sequences possess high public acceptance and the structural characteristic, they can assist learners to recognize relationships between subjects. Hence, experimental group learners had the ability to choose courses with more correct relationships than control group learners. In addition, the experimental group also attained higher scores in the rationality aspect (\(M_{\text{exp}}=74.25; M_{\text{control}}=59.98\)), which implies that the metacognition of experimental group learners could be reinforced. Since the clear connections between subjects facilitated learners in giving more attention to these connected concepts, learners were able to accurately and reasonably describe the connections between the chosen concepts.

According to the hierarchy of cognition ability, recognition is easier than description or lucubration. Additionally, describing contents in one’s own words presents a process of elaboration in information processing. Therefore, learners could not clearly describe the meaning of subject connections unless they were certain that there was substantial meaning in the subject connections. The statistic significance on rationality aspect supports this claim that LSRS assists learners in exploring the meaning of the subject connections. However, it is worth noting that the experimental group attained higher scores in the relation aspect than in the rationality aspect (84.37 vs. 74.25). This may imply that the system needs to provide more explicit information instead of only providing implicit connection information. More detailed discussions will be presented after the questionnaire results are shown.

As for the solidity aspect, the results revealed that there was no significant difference between two groups. A partial explanation for this may lie in the fact that the experiment adopted the random assignment method and that this aspect mainly focuses on the completeness of description of concepts for one single subject. Therefore, there is almost no effect on the variance of the solidity score between both groups, which also matches our expectations.

Despite the fact that the statistical data supplies positive evidence to support the LSRS’s potential in promoting learners’ learning transfer in the Web-based learning environment, the results evaluated from questionnaires provide more meaningful information and explanations of participants’ perceptions. The results of two semi-open questionnaires render partial explanations of the statistical results, as shown in Table 8 and Table 9.

70% of the teachers agreed that the recommended learning sequences actually facilitate transfer of learning in a Web-based environment (Question 1, Table 8). They admitted to the limitations of Web-based learning mentioned in the introduction, including the absence of cross-subject guiders and neglect of learning transfer during the course-design process. They expressed the view that that LSRS provides connections between each pair of subjects, which usually have identical elements or similar concepts. Hence, 65% of them thought that students could track the related elements among different courses (Question 4, Table 8). However, few teachers agreed that LSRS provided the
knowledge structure that connects the subjects (Question 2, Table 8). The reason they pointed out is the deficiency of helpful information, which is able to construct the knowledge structure within a curriculum. Also, they signified that LSRS should provide the substantial meaning behind recommended learning sequences rather than only provide the concepts that are similar between each pair of subjects. Thus, the transfer of learning would perform well after taking the meaning of learning sequences into account. Other than that, over half of teachers decided to take the learning transfer into consideration during future course-design, but some of them thought that the situation should depend on the features of the subjects (Question 3, Table 8). The opposite teachers believed that some subjects possess very high levels of independence from, rather than connections with, other subjects. Therefore, from the teachers’ perspective, LSRS would be better if it provided more meaningful learning sequences, which should depend on the different features of each subject.

Table 8. The research questions to teachers and their responses

<table>
<thead>
<tr>
<th>Teachers’ Choices</th>
<th>Question</th>
<th>Strongly approve &amp; approve (%)</th>
<th>Neutral (%)</th>
<th>Oppose &amp; Strongly oppose (%)</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) The LSRS made teaching for learning transfer possible in the Web-based learning environment.</td>
<td>70</td>
<td>26</td>
<td>4</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>(2) The LSRS appropriately illustrated the knowledge structure of the subjects.</td>
<td>26</td>
<td>31</td>
<td>43</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>(3) I will take the suggestions provided from LSRS into consideration while I am designing the courses for the next semester.</td>
<td>57</td>
<td>4</td>
<td>39</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>(4) With the system’s assistance, students can track the related learning elements (concepts) among different subjects.</td>
<td>65</td>
<td>31</td>
<td>4</td>
<td>3.9</td>
<td></td>
</tr>
</tbody>
</table>

Table 9. The research questions to learners and their responses

<table>
<thead>
<tr>
<th>Learners’ Choices</th>
<th>Question</th>
<th>Strongly approve &amp; approve (%)</th>
<th>Neutral (%)</th>
<th>Oppose &amp; Strongly oppose (%)</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) My interaction with LSRS was clear and understandable.</td>
<td>88</td>
<td>10</td>
<td>2</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>(2) The learning sequences provided by LSRS rendered an understandable relation between each of the two subjects.</td>
<td>51</td>
<td>5</td>
<td>44</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>(3) The learning sequences introduced by LSRS effectively helped me to explore a new subject or knowledge domain.</td>
<td>61</td>
<td>30</td>
<td>9</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>(4) During this semester, LSRS provides a macro view of the curriculum which consists of three subjects chosen.</td>
<td>44</td>
<td>30</td>
<td>26</td>
<td>3.3</td>
<td></td>
</tr>
</tbody>
</table>

The results indicate that 61% of learners gave positive views of the recommended learning sequences, which can help them explore a new subject (Question 3, Table 9). Some of them expressed the view that they had no idea how to choose the most useful subjects, while the fact that only three subjects can be chosen was stipulated in advance. Herein, LSRS provided them a few directions to help them choose the subjects and connect their existing knowledge from previous subjects with a new subject or knowledge domain. However, 9% of students expressed the opinion that the learning sequences only displayed the connections between each pair of subjects rather than explaining how to actually apply the learned knowledge or skills to the next subject. The results also corresponded with the statistical findings mentioned earlier, which were regarded as a limitation of LSRS. Other than that, 44% expressed positive views about the assistance of LSRS for enhancing their cognition in a semester curriculum (Question 4, Table 9).
They pointed out that LSRS helped them to choose a coherent curriculum. Despite the fact that nearly 50% of students gladly accepted the curriculums recommended by LSRS, 26% of students were still not predisposed to choosing the curriculums arranged in advance. They expressed the idea that the subjects within the recommended learning sequences were usually found in the similar domain, while regarding subjects from different domains, LSRS seemed not to offer a macro view of the curriculum. The most important finding from this result suggests that LSRS should improve its recommendation ability for cross-domain subjects. According to above mentioned limitations, additional research focusing on these aspects would be of great interest and value in understanding the role of cross-domain subjects and their influence on transfer of learning.

The usability of LSRS for experimental group learners

The responses to questions 1 and 2 (see Table 9) summarize the results obtained for evaluation of the perceived usability of LSRS, and the clearness and understandability of the learning sequences. Opinions extracted from questions 1 and 2 in the semi-open ended questionnaire denoted how the learner perceived the system user interface and the objects display presented in LSRS. 88% of learners responded that the LSRS user interface is very friendly and that the system is easy to use (Question 1, Table 9). Additionally, 51% of learners expressed the positive views about the clear display of the learning sequences while 44% of learners disagreed on this point (Question 2, Table 9). The mixed results imply that the system can still be greatly improved. Some of the opposite learners expressed that, although LSRS provides a clear display of learning sequences and the connections among subjects, the presented learning sequences should include more information about the linking relations, such as which textbooks have caused the relations or related assessments between two subjects. In this regard, more research is needed on the effects of more displayed information toward effective transfer of learning.

The LSRS's effectiveness for learners of different genders

Gender based differences in learning and behavior have long been recognized as an important issue on educational research (Herring, 1994; Scanlon, 2000; Weinman & Cain 1999). Furthermore, the importance of investigating whether or not gender differences exist in web-based learning settings is also reported in the literature (Garland & Martin, 2005; Lee & Tsai, 2004; McSporran & Young, 2001). However, the findings of these studies have been mixed. In this study, an independent-sample t-test was conducted to examine whether or not differences in learning transfer related to gender exist in the web-based learning environment. The results revealed no significant difference between males and females. Perhaps future research could examine different samples from different learning domains, and provide more detailed results, which may differentiate these views from one another.

Overall, the above evaluations reflect positively on our efforts since most respondents gave positive feedback regardless of whether they were teachers or learners. Both positive and negative feedback gives us more information to analyze so that we can offer more precise levels of individualization for each user in the future.

Conclusion

When effective group patterns are successfully applied, the likelihood of positive learning achievement and students’ satisfaction with group activities is significantly increased. Alternatively, when students encounter obstacles and feel frustrated because they do not use the experienced group-learning patterns for assistance, or do not believe that continued searching the available material is useful, web based learning can become a cause of stress, and its benefits to users are reduced. Therefore, it is all the more important for teachers to have a well planned learning sequence in hand when designing a course and ensure it leads learners to achieve transfer of learning.

In this paper, we have proposed a dependable learning sequence recommendation system. By employing the Markov chain model and an entropy-based approach, we discovered learning sequences and formed a recommended learning map that allowed learners to set their own learning pace in Web-based educational settings. The proposed approach successfully discovered applicable learning sequences so that learners could avoid redundant or otherwise ineffective learning behaviors.
From the evaluation results, we found that the use of the learning sequence by learners was able to significantly promote their learning transfer, a fact that has several important pedagogical implications.

For teachers, the system offers learning sequencing support by providing good recommendations and constructing a learning picture that is able to clarify the learning objectives for students. The evaluation results showed that most lecturers responded positively to the use of the system for their teaching, as it allowed them to effectively utilize the limited time they had available for student assistance. In the future, we will make an effort to provide sufficient learning objects to allow instructors to create more meaningful learning maps.

For learners, applicable learning sequences can be dynamically generated to target individual learners. Learning sequences inferred by the Markov chain model can narrow down and focus the whole learning process. The results revealed that there was not only a statistic significant difference in evaluating learning effectiveness; but that the learners also felt that they received benefits from the experienced learning sequences. Therefore, such learning recommendations in particular, may reduce the workload involved in searching for and accessing information, and lead to the development of better time management skills in a professional context. In addition, it is important to emphasize that the method of investigation is not without problems. Although we have measured whether learners acquired the essential elements of learning transfer, the detailed cognition and knowledge transfer processes seem not indicated from the results of the present study. We are hopeful that future research will provide more detailed results which may clearly reveal the transfer process of knowledge during students’ learning.

By observing patterns of group learning, the possible pitfalls in terms of student perceptions of the worth of the learning sequences become clear, and it is obviously important that efforts should be directed towards the provision of high quality materials, rather than the provision of large quantities of material.

We also may expand to a number of unanticipated nodes (or paths) for a subject, which means that the provision of learners’ selection would need to be carefully arranged so that all individual learners would be assured of some viable opportunities for extra learning over their course of study. Therefore, we will weight up the options for these recommended learning sequences and deploy the proposed system to the real-world learning environment found in our university e-learning system. Then, we believe much more research work is still needed to investigate for the scalability, performance, and fault-tolerance of the system.

Given a system capable of exploiting learning patterns in e-Learning settings, statistical heuristics might provide a default track to groups’ paths. To promote learning, we will hopefully find more detailed and precise criteria to accommodate their needs, and then provide a foundation upon which an even more intelligent and responsive system can be built.

Acknowledgements

This work was supported in part by the National Science Council (NSC), Taiwan, ROC, under Grant NSC 95-2221-E-006-307-MY3.

References


ABSTRACT

With the gradual adoption of “Web 2.0” technologies, one of key Web 2.0 technologies, blog, has become a popular and wide-accepted Web application. Although mobile device users can access the Web whenever or wherever the need arises, there is not an easy way to publish their thoughts and experiences via blog articles. In this study, we offer a solution by designing a mobile blogging system which enables mobile bloggers to publish their comments in authentic context anytime and anywhere. We show that with the help of the mobile blogging system, we can establish a collaborative learning model for students in virtual classrooms. The results revealed from the learning outcome are positive and encouraging regarding the effectiveness of the supported collaborative learning model. In the conclusions, we discuss the findings and applications of the proposed system in collaborative learning.

Keywords

Collaborative learning, W/H device, Blog, Information retrieval

Introduction

By the introduction of CSCL (Computer Supported Collaborative Learning) to e-learning, many studies have revealed CSCL to be a successful learning scheme in the field of pedagogy. Using this kind of pedagogy, learners not only can concentrate on their own assignments and projects but also can work collaboratively with other students to achieve goals, and in so doing, enhancing their own learning (Koschmann, 1996; Dillenbourg, 1999). For mutual benefit, the learners are inspired to achieve efficient learning. A positive interdependence is produced among the learners, since they are aware that other members are working together with them towards a common outcome (Kravcik et al., 2004). In recent years, due to the rapid development of information technology (IT), collaborative learning now takes place on the Internet as well as in traditional classrooms (Carchiolo et al., 2002; Serce & Yildirim, 2006; Gogoulou et al., 2007).

Due to the rapid emergence of wireless communication technology and mobile devices, the use of handheld technology in education has increasingly been the object of study in recent years. From laptops to wireless phones and handheld devices (or W/H devices for short) (Motiwalla, 2005), the nature of higher education has been altered through the massive infusion of computing devices and the rapid improvement of Internet capabilities (Green, 2000). Via wireless technology, W/H devices can be synchronized with or connected to computers or wireless networks (WiFi, GPRS, 3G, etc.). Research has shown that W/H devices provide new opportunities for communication and innovative learner interaction both in and out of the learning setting (Chen et al., 2003; Tatar et al., 2003; Clough et al., 2007). However, there is no clear consensus on how to evaluate the benefits of applying wireless technologies to collaborative learning environments.

This paper presents an implementation of W/H devices in support of a blogging learning system on a real curriculum. The goals of the study are not only to investigate the effects of learning process through the use of W/H devices, but to explore students’ learning outcomes from three aspects: (1) the learning effects of the mobile blogging system in a collaborative learning model, (2) whether being a mobile blogger can generate positive emotion, (3) whether the interactivity in blogging learning environments can also be seen in mobile blogging learning settings.

We have organized the rest of this paper in the following way: the related works are introduced in the next section and the research methodology is described in the following section, addressing both research design and learning activities. Following, we describe the system description which consists of architecture and implementation details. Afterward, the evaluation of the system with the designed learning activity is outlined; the results and discussions for the various analyses are shown following each of these descriptive sections, and finally, conclusions are presented.
Related Efforts

An Internet forum is a discussion space utilized by individuals interested in a specific topic. However, it is difficult to personalize content within a forum and a Web-based client is usually a user rather than an administrator. This deficiency can be made through the use of an emerging application, the weblog, more commonly known as a blog. A blog is an interactive website that both publishes and collects personal information, according to the purposes and plans of the blog editor (Katerina and Chronis, 2007). Blogging systems generally include the following five features: content which is separated from presentation; presentation templates which are usually provided for bloggers to easily produce blog articles; blogging application programming interfaces (APIs) also help blogging clients use other software, such as Microsoft Word, to publish content to blogs; many information management tools which are offered in blogging systems for content management; and Really Simple Syndication (RSS) a mechanism which offers a subscription to the blogging clients who are interested in specific topics (Lindahl and Blount, 2003). Blogging systems have the ability to support informal communication, shared learning, group reflection, and community building (Nardi et al., 2004; Divitini et al., 2005). Because of these characteristics, blogging systems have begun to be used in educational settings in recent years. An implemented blogging system has been proven to be feasible in an international distance course (Lin et al., 2006). Students were able to use the system to document their learning, and to share experiences and knowledge. Meanwhile, a blogging system is able to provide useful informative auxiliary materials for assisting students’ learning (Huang et al., 2008). In addition, blog articles were applied to construct a learning map called blog-based dynamic learning map (Wang et al., 2008). It is designed to provide informative blog articles to assist students’ learning. In China, researchers applied blogs to involve new ideas of student assessment. The blog applications help them be toward new functions and goals of authentic and formative assessment previously disregarded in China (Chen & Bonk, 2008). Moreover, difficulties around the issues of direct student-student and student-tutor interactions were addressed. However, some delivery features of a blog increased anxiety among some students, resulting in some cases, in the inefficient use of learning time (Dron, 2003). Through the application of different structural and methodological techniques, these problems can be overcome. Using the learning research methods and learners’ experience presented in a blog (Giarrè and Jaccheri, 2006), researchers or learners have been able to study the relationship between the research process and results (Giarrè and Jaccheri, 2005). With blogs, users can exchange information, discuss research problems, and by drawing upon the expertise of each other, collaboratively find solutions to problems and new applications for concepts without any constraints in terms of time (Mai, 2005; Nicola and Giuseppe, 2006). Nevertheless, most of blogging actions take place on PCs or laptops, which causes the temporal and spatial limitation of publishing blog entries. Under these constrains, blogging learning could not make students notice the features of problems encountered in authentic context. Hence, students could not make particular relevant actions to solve the problems encountered.

Since W/H devices can support cooperative and collaborative learning environment, it greatly enhances mobility, coordination, communication, and organization of materials, negotiation, and interactivity in ways not possible in conventional collaborative learning environments (Zurita and Nussbaum, 2004; Lai and Wu, 2006). Zurita and Nussbaum (2004) indicated the weaknesses of non-tech collaborative learning settings and described how W/H devices improved them in a mobile computer supported collaborative learning environment. Additionally, Roschelle and Pea (2002) have suggested five reasons for using W/H devices in collaborative learning environments: (1) augmenting physical space, (2) leveraging topological space, (3) aggregating coherently across all students’ individual contributions, (4) facilitating evaluation of student performance, and (5) providing an easy, instant way to archive student work for future reference. The mobility and connectivity of W/H devices allows the group members to work on a task and to discuss topics with each other even while they are away from the group. Overall, W/H devices provide a more natural, mobile collaboration environment (Imielinsky and Badrinath, 1994; Danesh et al., 2001; Kurti et al., 2007).

Methodology

Research Design

In traditional upper level classes, the teacher would give a lecture on a platform in front of a group of students for around an hour. During the lecture, students could take notes and record key points of the presentation. Once the session ended students could discuss the points of the class amongst themselves and share their notes with each other, but they can no longer access the content of the lecture itself. Because of this, and because of a great deal of
variation in the ways in which students attend to a lecture and record their impressions, after-class discussions could just easily lead to confusion. Nowadays, with the advance of information technology (IT), the lectures could be recorded in digital form and uploaded to a specific e-learning web site. Thus, students can review the key points of a lecture long after the class has ended up, and by making reference to the original material, correct some of the more obvious deficiencies in their notes. What is more, blogging is now being widely used to facilitate discussions amongst students and to publish students’ comments on an e-learning web site. Lecturers can initially upload learning content or a lecture video in a blog system; students can then access the information and post their responses to it. However, the extent to which digital technology can contribute to learning has not yet been fully realized, and the next step in this exciting field is to find ways to free students from the limitations imposed by the need for access to desktop computer facilities. As it stands today, it is difficult if not impossible for most students to engage in meaningful digitized learning without cumbersome computer hardware, in a specific and fixed location. To address this issue, we proposed a mobile blogging learning environment to support learning activity via W/H devices, which aid in the construction of a truly modern collaborative learning environment. Then we proposed an activity design which is described as follows.

This section introduces the research design concept of the paper which applies the mobile blogging system to facilitate the learning activity in a collaborative learning environment. When we look at technology-rich learning environment, we need to look at activity from both a ‘Design’ and ‘Use’ perspectives (Wasson & Ludvigsen, 2003; Wasson, 2007). Figure 1 illustrates the ‘Design’ and ‘Use’ perspectives which direct a technology enhanced learning environment to deep thinking and analysis. Besides, we add two perspectives ‘Mobility’ and ‘Blogging Application’ to emphasize a mobile enhanced learning environment to make a more concrete description. The pedagogical view of collaborative learning can be regarded as the theoretical fundamental of design perspective and technology also supports the design view of the system. After students manipulate the mobile blogging system in a learning activity, the use from collaborative and technological perspective should be observed in the experimental process which can further influence the design aspect by evaluating the learning effect of students. When constructing the learning activity from design and use perspective, the concept of mobility and blogging application are embedded into the learning activity. Implementation of the design perspective from mobility and blogging application aspects brings collaborative learning environment a technological-enhanced issue and influence result of the activity brings a discussion room by this study. According to the use perspective, we designed an experiment in which students carry out the designed learning activity in a collaborative learning environment. After the experiment, the evaluation which is inferred from collaborative and technological perspective can reveal the learning outcome of students.

**Design of the Learning Activity**

This activity design was intended to investigate the assistance of using handhelds to support a normal learning activity in a collaborative learning environment. The experiment stresses a need to place students in a mobile collaborative learning environment in which they can engage in effortful interactions with blogging system and learning companies in order to achieve the learning goal. The designed learning activity took place over two months in a semester, and the participants included 40 college students who major in engineering science of National Cheng Kung University in Taiwan. The participants are all sophomore students in college and they all have mobile devices which support network communications. The students are all familiar with the operation of personal computer and mobile devices so that they can soon be familiar with the operation of the mobile blogging system. We developed the learning activity based on collaborative learning conception and constructed the technological environment, designed the help and assistance for the deployment. The detail collaborative learning activity is described as follows.

The testing course is called data-structure which depicts the way of storing data in a computer so that it can be used efficiently. The basic concept of data-structure contains *Arrays, Stacks, Queues, Linked Lists* and *Tree*, and they should be learned step by step so that students can learn the entire concept more concrete. In the beginning of the learning activity, the students were divided into five groups and each group was assigned a topic from the five types of the course. During the experimental period, each group was asked to contribute the learning resources and helpful discussion content for the assigned topic. Thus, there were five blogging topics on the mobile blogging system; each group maintained the blog content of a topic and all groups needed to read the blog content of five topics on the mobile blogging system. They can also discuss about the related topic on mobile blog as well as solve the question posted on it. The lecturer conducted the learning activity in class from the sequence of *Arrays to Tree*. Students can both learn the course topic from the lecture as well as the mobile blogging system. The students were all equipped.
with handheld devices and they were asked to view and post the articles/questions on the blogging system. In this learning activity, the group was to:

1. Organize a collaborative group effort to contribute the related topic of the course as well as promote the understanding of the entire concept.
2. Produce a report of the assigned topic which contains the learning material, lecture note and discussion result.

In addition to participating in the learning activity, each student needed to accomplish the questionnaire which reflects the perception of being a mobile blogger.

The technological aspects of the mobile blogging system comprise three modules, mobile blog, mobile RSS aggregator and mobile IR. Mobile blog, a blogging system deployed in the mobile setting environment, was used as the main information delivery and discussion technology. Mobile blog provides a well-integrated platform which enables group members to contribute personal knowledge, discuss posted questions, and upload learning materials (web links, files, pictures or video clips) any time and any where that construct a group collaboration work. With a various type of information presentation, mobile blog help students get more authentic context example of knowledge. Furthermore, mobile blog facilitates the personal communication with email-like function, that is, students can mail the question or comment to anyone in the system which is a personal interaction support function. From the collaborative learning perspective, mobile blog enables collaboration and supports the independent interaction between group members. Mobile RSS aggregator records the RSS feed and receives the updated information for mobile blogger. Mobile RSS aggregator assists the share of information between group members and promotes the collaborative learning activity. Mobile IR is a search engine which integrates the information retrieval (IR) technology and responds the searching request with more precise answer. Mobile IR improves the efficiency for students’ searching request and help students acquire more related lecture information. The detail mobile blogging system description is illustrated in the following section.

**Implementation of the Learning Activity**

The designed learning activity was deployed for 2 months (8 weeks) including testing phase, training phase, main experiment and evaluation phase. The testing phase took 1 week, and students test their account and group mail address in this period. The training phase lasted for another 1 week. In this phase, the teacher assistance instructed the use of the mobile blogging system and introduced the learning activity which includes the idea of how to achieve the collaborative learning in the mobile setting. The main experimental learning activity took place in the next 5 weeks. There were 5 groups and each group was assigned a topic. The topics included *Arrays, Stacks, Queues, Linked Lists* and *Tree*. The students collaborated to complete the learning activity and contributed the information,
thoughts, and comments for the data-structure course. The final week was reserved for students to complete the report and the questionnaire.

System Description

Architecture

Since mobile devices are widely used in everyday life, we hoped to show that incorporating their use in a formal learning model can motivate students to take full advantage of the resources available on the mobile blogging system. Through increased frequency of discussion with others, we expected learners to not only learn assigned material more thoroughly, but also learn valuable social and communication skills, such as contributing to a group effort, exercising creativity and showing initiative. The mobile blogging learning system described in this study supports a mobile blog discussion module. We implemented a mobile blogging learning architecture utilizing the support of the RSS mechanism which is able to automatically update learning content so the waiting time for receiving learning content could be reduced and the learning efficiency could be expected to greatly raise.

Really Simple Syndication (RSS 2.0) is specified in XML format. RSS gives a blog article its information in an XML file and the file is named a RSS feed. Bloggers can subscribe to a blog RSS feed or RSS aggregator and automatically receive the latest postings instead of searching for them on different blog servers. An RSS feed contains the individual posted blog entries on a blog server. The content of an RSS feed contains the title and the hyperlink of a blog entry. There are also many other items that describe each blog article such as the publishing date, description, and the related file location.

The system architecture and the operation workflow of the system are shown in Figure 2. The mobile blogging system provides mobile bloggers with a flexible web environment for discussion and communication. Mobile bloggers can post blog articles using W/H devices and these articles are subsequently stored in the blog server. Additionally, through the devised transformation agent, posted blog articles can be read on different platforms such as personal computers, laptops, and other mobile devices. The agent renders different presentation of learning content by calculating different device screen sizes. If the platforms possess powerful presentation capabilities, the
agent would direct learners to a learning interface which could present considerable multimedia information. Conversely, the agent directs learners to a simple interface which renders mobile bloggers more readable learning interface. The streaming server provides video streaming which can provide rich multimedia support for posted articles. Mobile bloggers can receive other blog articles from other individuals by using the devised mobile RSS aggregator which collects all RSS feeds. The functionality of mobile RSS aggregator will be introduced in the system implementation section in more detail.

System Implementation

The mobile blogging learning system runs on a computer AMD Athlon 64 3000+ CPU at 2.0 GHz and 1G memory, using Microsoft C#.net 2005. Dopod CHT 9000 was chosen to be the testbed mobile device. This section contains the system presentation of a mobile blog, mobile RSS aggregator, and mobile information retrieval (MIR), which are presented graphically in Figure 3. RSS aggregator and MIR can be regarded as two components in the mobile blog. Moreover, in order to prove the system is able to run on real mobile devices, pictures were taken from real mobile devices instead of using screenshots of a mobile simulator.

![Figure 3. Three main implemented components](image)

**Mobile Blog**

As Figure 4 shows, there are several links on the main page of a mobile blog, which include information retrieval (IR), RSS feeds, and the entry link of a personal blog, classification by groups and classes, and new account registration. Each posted article is listed below those links.

![Figure 4. The system presentation of a mobile blog](image)
A mobile blogger can directly edit blog articles on kinds of mobile devices at any place where has wireless signals. Furthermore, if the W/H devices support the Multimedia Messaging Service (MMS), mobile bloggers can also post multimedia files such as images, video, or audio files to their blogs. It should be noted that if mobile bloggers want to post blog articles with multimedia files, they must use MMS to send the edited messages to a special number that is assigned from an Internet Service Provider (ISP) and connected to an Integrated Messaging Service Platform (IMSP) server which deals with message services such as Short Message Service (SMS) and MMS. This particular server is often managed by an ISP. In order to retrieve the messages from an IMSP server, we developed an application called IMSP Decoder which is able to decode messages composed of text, images, and 3GP files, as shown in Figure 5. After pressing the button “Import IMSP”, the decoder executes and then gets messages from the IMSP server continuously. The content of the messages are shown in the decoder, including the date, source, ID, original text, image, and 3gp files. Meanwhile, when our system receives a message, it is recorded in a log file so that the further analysis would be dealt with. After the IMSP Decoder decodes the messages, related information will be stockpiled in the blog server and then the blog articles will be published automatically. Figure 6 shows the process of publishing blog articles using personal mobile devices.
Mobile RSS Aggregator

In order to subscribe an RSS feed into a personal blog, we developed an RSS aggregator. It extracts several key tags included in an RSS feed file, such as channel, title, link, description, etc. Figure 7 shows two RSS feed examples from five feeds: Linked Lists and Tree. The devised RSS aggregator can link to our blog server where stockpiles all managed RSS feeds and blog articles posted. Because of this, the content of blog server is constantly being updated and is always current. An example of aggregated blog RSS titles and item content are given in Figure 8 and Figure 9 respectively. Mobile bloggers could view blog articles under a specific topic. In this case, the “Linked Lists” topic is selected and five blog articles also called items are included under this topic.

![Figure 7. Two examples of an RSS feed](image1)

![Figure 8. Choosing an article from the RSS Aggregator](image2)

![Figure 9. The content of a chosen article in the RSS Aggregator](image3)
**Mobile IR**

When blog articles are stored in the blog server, IR techniques are adopted to proceed with tokenization. All blog articles are tokenized as terms that are viewed as words. Subsequently, the correlations between terms and articles are calculated, using vector models of term frequency-inverse document frequency (TFIDF) measure. Finally, we adopt the Wikipedia way to display the tokenization results, as shown in Figure 10. Each underlined word links to related articles, which makes mobile bloggers easily find interesting discussion topics and other related articles.

![Figure 10. A search result from mobile IR](image)

**Evaluation of the use of the Mobile Blogging system**

The proposed mobile blogging system was regarded as an assistant tool in a collaborating learning perspective that emphasizes the mobility and blogging application in the process of learning activity. We examined the effectiveness of the mobile blogging system on students' learning attitude using an online questionnaire. The response portion of each question in the questionnaire was designed using a 5-point Likert scale. Typically, an item in a Likert scale is given as a statement to which the user must respond using a scale from 1 to 5, in which 5 stands for “strongly agree” and 1 stands for “strongly disagree”. The 5 level responses also stand for the score of each question thus we can calculate the mean value of each item. We also examined the students’ attitude from open ended question which can present more realistic response. The questionnaire item includes:

1. Help me satisfy the urgency of learning need of the lecture on the mobile blogging system. (5-point likert-scale)
2. Help me acquire the related lecture information. (5-point likert-scale)
3. It is easy and convenient for me to access blog articles. (5-point likert-scale)
4. The interactivity of the learning process is efficient for me. (5-point likert-scale)
5. The mobile blogging system provides the authentic context of the learning material for me. (5-point likert-scale)
6. The mobile blogging system provides the organized learning content for me. (5-point likert-scale)
7. Please comment the learning experience as a mobile blogger. (open ended question)
8. Please comment the interactivity you can do during the learning activity. (open ended question)

As we stated in the introduction part, we wanted to explore students’ learning outcomes from three aspects. The above items can separately reply the three aspects. Item (1) to (6) can depict the first aspects from a statistic view, and item (7) and (8) can illustrate the second and third aspect according to the responses of students. The responses data from students were collected from online questionnaire system and it can be further analyzed and discussed. The statistical results were presented in Table 1. The 5th column describes the percentage of each item score that are greater or equal to 4. The responses to first item indicate that most students were satisfied with the mobile blogging system (Mean=4.350, 82.5%). This reflects that the mobile blogging system can be used for an urgent matter of
learning, such as posting question and searching information. The responses to second item indicate that most students can acquire the related lecture information articles on the mobile blogging system (Mean=4.425, 90%). This reflects that mobile IR in the mobile blogging system can respond the related information based on student’s requests in time. The responses to 3rd item indicates that to access blog article is easy and convenient (Mean=4.375, 85%). According to the 3rd item, the educational practice can be performed conveniently and easily any time and any place. The responses to 4th item indicates that the mobile blogging system provides smooth using experience for students (Mean=4.350, 87.5%). Therefore, students can notice the features of problem situations that make particular replies relevant. The responses to 5th item indicates that the mobile blogging system can provide more authentic context learning material/example for students (Mean=4.350, 85%). The responses to 6th item indicates that the mobile blogging system provides more authentic context learning material/example for students (Mean=4.350, 87.5%). Therefore, students can notice the features of problem situations that make particular replies relevant. The responses to 6th item indicates that the mobile blogging system can also provide integrated learning material with well organized (Mean=4.275, 80%). This feature facilitates complex and ill-structured learning content to be understood easily. Besides, the responses from item (7) and (8) were summarized in Table 2. According to the responses of 7th item, the mobile blogging system was regarded as a helpful, convenient tool in a collaborative learning activity. The mobile blogging system can provide the desired information for students and it can stimulate students to look for more information on the system. The responses of 8th item states that the interaction to use blogging system on a mobile device is similar to the use on a personal computer. Students wrote that the successful use of the mobile blogging system was not just tied to the convenient of use, rather, that it provides the necessary function to satisfy the collaborative learning activity.

**Table 1. Questionnaire Result**

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
<th>Stand Deviation</th>
<th>Variance</th>
<th>Score &gt;= 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.350</td>
<td>0.769</td>
<td>0.592</td>
<td>82.5%</td>
</tr>
<tr>
<td>2</td>
<td>4.425</td>
<td>0.675</td>
<td>0.455</td>
<td>90%</td>
</tr>
<tr>
<td>3</td>
<td>4.375</td>
<td>0.740</td>
<td>0.548</td>
<td>85%</td>
</tr>
<tr>
<td>4</td>
<td>4.350</td>
<td>0.735</td>
<td>0.541</td>
<td>85%</td>
</tr>
<tr>
<td>5</td>
<td>4.350</td>
<td>0.699</td>
<td>0.489</td>
<td>87.5%</td>
</tr>
<tr>
<td>6</td>
<td>4.275</td>
<td>0.784</td>
<td>0.614</td>
<td>80%</td>
</tr>
</tbody>
</table>

**Table 2. Open Ended Question Summary**

<table>
<thead>
<tr>
<th>Item</th>
<th>Summary</th>
</tr>
</thead>
</table>
| 7    | (a) ...I want to share more helpful online resources with my classmates....
      | (b) ...Mobile blogging system is convenient for me to post article, upload pictures any time and anywhere, even when I was in bathroom...
      | (c) ...I can find the article I want through the searching engine...
| 8    | (a) ...I can discuss the question with my group members privately before I post the question on the mobile blog...
      | (b) ...I can write the article, post my comment and upload the figure which is similar to the interactivity in my additional blog on the web...
      | (c) ...I can subscribe the RSS feed on mobile blog which is convenient for me to gain latest article...
      | (d) ...when I see an event which is related to the learning concept, I can take a picture by the mobile device and upload it using mobile blog immediately...

The mobile blogging system is useful and it furnishes students with well learning effect in collaborative learning environment which is working with peers in the classroom and in the field as well as with students from other schools or countries towards a common outcome. In a mobile blogging environment, students need not to meet each other but can make the collaboration effort to accomplish the learning goal of the lecture. Additionally, according to the result of questionnaire, student as a mobile blogger can generate positive emotion (see Table 2, response of Item 7(a) and 8(b)) during the collaborative learning activity. The mobile blogging system not only provides the blogging application for students but also the convenience with no limitation in time and position. In our observation from the experiment, students gradually learned the coordination of activity between group members to solve an identified problem. Particularly, the coordination process took place in the distributed situation where students need not to get together. In this interactivity, the mobile blogging system encourages students’ cognitive activity and self-regulation in the collaborative learning activity, engaged students to learn and discuss during the activity. Furthermore, a mobile blogging system in collaborative learning activity can train students to provide constructive comment because that everyone can post, ask and comment a question before the posted information is useful. Neutral comments are
perceived to be less useful than constructive or supportive comments (e.g., suggesting what the next step or encouraging comments). The mobile blogging system provides more convenient to free the tie from sitting together and meeting face to face so that students have more courage and time to consider posting an article/comment. Consequently, students learn to contribute constructive comment in the mobile blogging system.

Our general findings in the experiment include:

- In collaborative learning setting, the mobility and blogging applications play an important part.
- The mobile blogging system offers more authentic context learning example for students. For example, students took the picture of people stand in a line as an instance of Queues and shared it to group members.
- Coordination issue is a challenge for collaboration with distributed students, and the mobile blogging system provides a possible solution for it.
- Blogging applications provide potential possibility to facilitate different kind of learning activity. However, the mobile blogging system owns the advantage in mobility of learning setting.

From the above findings, we can summarize to reflect the response of the three perspectives mentioned in the Introduction section as follows. (1) With the benefit of mobility and blogging applications implemented in the collaborative learning setting, students’ learning attitude is enthusiastic and their responses are appreciative of the blogging system. Students can find useful articles in the blogging system and enjoy as a mobile blogger. Additionally, this interactivity in collaborative learning setting brings learning effect a positive impact. (2) In the responses of questionnaire item 7, students thought the mobile blogging system is a helpful and convenient tool to acquire more authentic context learning examples. They are willing to discover knowledge in the system and share their own learning materials with other peers. In the process of discovering and sharing, their learning emotion is positive. (3) According to responses of questionnaire item 8, students can interact with the mobile blogging system similarly as they do in a general blogging system. Moreover, the mobile blogging system brings coordination issue a solution by its advantage in mobility.

**Conclusions**

This study aimed to investigate the learning effects of the mobile blogging system in a collaborative learning model as well as to explore the learning behavior of mobile blogger. We highlight the importance of mobility and blogging applications in a collaborative learning environment. Thus, an educational mobile blogging system is implemented for this purpose. The designed learning activity focuses on exploring students’ learning outcomes from the proposed three aspects in the Introduction section. The results of the conducted learning activity were evaluated by the questionnaire, which reveal our observations and findings. We found that the mobile blogging system can provide more authentic context learning example and help to solve the coordination issue in a collaborative learning environment. In addition, the developed mobile blogging system established a mobile blog-based learning environment which brings students a similar manipulation of web-based blogging system in daily life and ties no position and time limitations.

Technology that supports the development of pedagogic activity should be deployed according to the concerned issue of pedagogic activity. In this study, the experimental activity joints intellectual effort by the participants from different position and time which is the characteristics of collaborative learning activity. By the reflection of students’ learning outcomes, mobility and blogging applications bring them an unlimited discussion space so that they can learn in a more free and easy manner.

Learning with related and appropriate example helps students in reasoning and deliberation of learning concept. A good learning example should be easily realized and make students feel familiar with daily life. The result of the study implies that the mobile blogging system can be used for providing authentic context of learning example. For instance, student can share a picture of a train which can be an example of Linked Lists concept and upload it to mobile blog. After students read the blog article, they can discuss and comment it. One student can illustrate the concept of the figure, and the others can look for different example more appropriate. By the manner, students learn to coordinate their efforts toward consist and correct outcomes and elaborate the learning material in a collaborative learning setting.
With novel technological support, more tools that emphasize aspects of mobile supported learning environment could be made available for lecturers or students to facilitate their learning process. However, it is important to put right technology into right pedagogic application. This study implemented a mobile blogging system, and took the characteristics of it to improve the learning effect in a collaborative learning environment. Additionally, we believe design a technological supported learning environment should consider along with an appropriate pedagogic theory. In this manner, technology supported learning system not only helps to amplify the feature of the pedagogic theory, but also stimulates students’ learning emotion to learn enthusiastically.

Acknowledgements

This work was supported in part by the National Science Council (NSC), Taiwan, ROC, under Grant NSC 95-2221-E-006-307-MY3.

References


Affective e-Learning: Using “Emotional” Data to Improve Learning in Pervasive Learning Environment

Liping Shen¹, Minjuan Wang² and Ruimin Shen¹

¹Computer Science and Engineering Department, Shanghai Jiaotong University, China // ²Department of Educational Technology, San Diego State University, USA // lpshen@sjtu.edu.cn // mwang@mail.sdsu.edu // rmshen@sjtu.edu.cn

ABSTRACT
Using emotion detection technologies from biophysical signals, this study explored how emotion evolves during learning process and how emotion feedback could be used to improve learning experiences. This article also described a cutting-edge pervasive e-Learning platform used in a Shanghai online college and proposed an affective e-Learning model, which combined learners’ emotions with the Shanghai e-Learning platform. The study was guided by Russell’s circumplex model of affect and Kort’s learning spiral model. The results about emotion recognition from physiological signals achieved a best-case accuracy (86.3%) for four types of learning emotions. And results from emotion revolution study showed that engagement and confusion were the most important and frequently occurred emotions in learning, which is consistent with the findings from AutoTutor project. No evidence from this study validated Kort’s learning spiral model. An experimental prototype of the affective e-Learning model was built to help improve students’ learning experience by customizing learning material delivery based on students’ emotional state. Experiments indicated the superiority of emotion aware over non-emotion-aware with a performance increase of 91%.

Keywords
E-Learning, Affective, Emotion detection, Pervasive computing

Introduction
In the past decade, e-Learning has evolved from Computer Aided Instruction, through Intelligent Tutoring System, to Smart Classrooms, and to Mobile Learning (e-Learning with mobile devices). Today, e-Learning becomes heavily learner-centered, emphasizing pervasive and personalized learning technologies. Also known as ubiquitous or ambient learning, pervasive learning refers to learning that is available anywhere anytime (www.pervasive-Learning.org). Pervasive learning is supported by wireless communication and wearable computing. “In combination with session and service mobility as well as device independency, the arising learning environments should have the potential to be accessed by anyone at any place and any time. Learning wherever and whenever needed is to become possible” (http://wwwra.informatik.uni-rostock.de/perel08/). E-Learning should not only generate good learning outcomes, but also better engage learners in the learning process. From engaged learning perspective, truly engaged learners are behaviorally, intellectually, and emotionally involved in their learning tasks (Bangert-Drowns & Pyke, 2001; Wang & Kang, 2006).

The influence of emotions on learning is still under-emphasized. Recently, a growing body of literature (e.g., Currin, 2003; Dirkx; Hara & Kling, 2000; Kort, Reilly & Picard, 2001; Wang & Kang, 2006) has begun to espouse the central role of emotion to any learning endeavor and outcomes, especially in online learning. Continuous and increasing exploration of the complex set of parameters surrounding online learning reveals the importance of the emotional states of learners and especially the relationship between emotions and effective learning (e.g., Kort, Reilly & Picard, 2001). Research (Isen, 2000) also demonstrates that a slight positive mood does not just make you feel a little better but also induces a different kind of thinking, characterized by a tendency towards greater creativity and flexibility in problem solving, as well as more efficiency and thoroughness in decision making. These findings underscore the important effects of emotions on learning. Human brain is not just as a purely cognitive information processing system, but as a system in which both affective functions and cognitive functions are inextricably integrated.

In this paper we describe a state of the art e-Learning platform based at Shanghai and we have augmented this with the research of affective computing coined by Picard (1997). Affective computing describes computer methods that are related to, derived from or deliberately designed to influence, emotions. It involves two areas: emotion synthesis used to artificially imitate some of the physical or behavioral expressions associated with affective states, and
emotion analysis which is often employed in decision making for interactive systems. Emotion synthesis is useful to develop ways to communicate with humans at a subjective level involving social participation, for example using robots. Emotion analysis could be used to monitor the emotional state of a subject, taking actions based on the type of individual feeling being experienced. Some computing systems are capable of displaying immediate reactions to people’s feelings by incorporating a combination of both emotion detection and emotion synthesis (Garzon, Ankaraju, Drumwright, & Kozma, 2002; Morishima, 2000). “We believe that existing and future affective and cognitive research needs to be adapted and applied to actual learning situations. Thus far, most research on emotions does not bridge the gap to learning” (Picard et al., 2004). This study intends to fill in this gap, by discussing how physiology-based emotion sensing is integrated into the e-Learning platform. The goal of this study is to help improve students’ learning experience by customizing learning material delivery based on the learner’s emotional state (e.g. curiosity, confusion, frustration, etc).

This article describes the development of an affective e-Learning model, and demonstrates the machine’s ability to recognize learner emotions from physiological signals. In the following, we first introduce related work and our pervasive e-Learning platform; we then describe our “emotion-integrated” e-Learning architectural model and its theoretical foundations. Afterwards, we present the preliminary experiments, emotion classification, and data analysis. Finally, we summarize the findings and describe further research planned for the near future.

Related Work and the Pervasive e-Learning Platform

The extension of cognitive theory to explain and exploit the role of affect in learning is still in its infancy (Picard et al., 2004). Studies carried out by the AutoTutor Group presented evidence that confusion and flow were positively correlated with learning gains, while boredom is negatively correlated (Craig, Graesser, Sullins, & Gholson, 2004). Stein and Levine (1991) identified a link between a person’s goals and emotions. Their model adopts a goal-directed, problem-solving approach and predicts that learning almost always occurs during an emotional episode. Research revealed that emotion can also affect learner motivations (Keller & Suzuki, 1988). Kort, Reilly and Picard (2001) proposed a four quadrant learning spiral model in which emotions change when the learner moves through the quadrants and up the spiral. They also proposed five sets of emotions that may be relevant to learning. However, empirical evidence is needed to validate the learning spiral model and to confirm the effects that these emotions might have on learning.

The Affective Computing Group at MIT’s Media Lab is investigating the interplay of emotion, cognition, and learning as part of its “Learning Companion” project. This project aims to develop an ‘affective companion’ prototype that will provide emotional support to students in the learning process, by helping to alleviate frustration and self-doubt (Burleson, Picard, Perlin, & Lippincott, 2004). Studies carried out by the AutoTutor Group discovered that learning gains correlate positively with the affective states of flow and confusion (Craig, Graesser, Sullins, & Gholson, 2004). According to Fowler’s (1997) study, the relationship between learning performance and the arousal is a type of inverted-U curve, and people learn best when their emotions are at a moderate optimal arousal level. Shen et al. (2007) validated this relationship between performance and arousal and revealed that arousal remained relatively stable during learning.

For user emotion modeling, researchers and developers widely refer to Russell’s (1980) two-dimension ‘circumplex model of affect’, where emotions are seen as combinations of arousal and valence (Craig et al., 2004; Fagerberg, Stähle, & Höök, 2004; Kort et al., 2001; Leon, Clarke, & Callaghan, 2007; Picard, Vyzas, & Healey, 2001). Another model, known as OCC (Ortony, Clore, & Collins, 1990), is also established as the standard cognitive appraisal model for emotions. This model specifies 22 emotion categories based on emotional reactions to situations constructed either as a) goals of relevant events, b) actions of an accountable agent, or c) attitudes of attractive or unattractive objects. Conati and Zhou (2002) used the OCC model for recognizing user emotions in their educational game Prime Climb. Katsionis and Virvou (2005) adapted the OCC model to map students’ emotions when they played an educational game. Emotions are also used to design and model learning content. Papert (1996) conducted a project that he described as ‘Instead of trying to make children love the math they hate, make a math they’ll love’. Participants were involved in designing things-to-learn so as to elicit emotions in ways that will facilitate learning. Other non-educational settings also began to tap into the influence of emotions. For instance, video content with emotion tags were modeled to support the automatic generation of ‘video highlights’ or personalized recommendations for video films (Hanjalic & Xu, 2005). Sundström (2005) proposed an interesting concept known
as an ‘affective loop’, which refers to an affective interaction process or cycle where emotion plays an important role in interaction involvement and evolution.

Emotion recognition is one of the key steps towards affective computing. Many efforts have been taken to recognize emotions using facial expressions, speech and physiological signals (Cowie et al., 2001; Healey, 2000; Picard et al., 2001). The identification and classification of emotional changes has achieved results ranging from 70–98% on six categories of facial expressions exhibited by actors (Bassili, 1979) to 50-87.5% for speech recognition (Nicholson, Takahashi, & Nakatsu, 2000). The successes in physiological emotion detection include Healey (2000)’s 80–90% correct classification for eight emotions, Leon et al.’s (2007) 85.2% for three valence states real-time recognition, Haag et al.(2004)’s 90% for three valence states and Picard et al.’s (2001) 81% for eight emotions. It is suggested however that, because physiological measures are more difficult to conceal or manipulate than facial expressions and vocal utterances, and potentially less intrusive to detect and measure, they are a more reliable representation of inner feelings and remain the most promising way for detecting emotions in computer science.

Pervasive e-Learning Platform is one type of e-Learning platforms that provide “always on” education. It aims to support pervasive learning environments where learning resources and tools could be accessed by students anytime anywhere. It differs from the previous platforms by using wireless computing and pervasive computing technologies. The pervasive e-Learning platform (Figure 1) developed at an online college of a major University in Shanghai delivers fully interactive lectures to PCs, laptops, PDA, IPTV and mobile phones. The core of the platform includes a number of “smart classrooms” distributed around Shanghai, the Yangtze River delta, and even in remote western regions of China such as Tibet, Yan’an, Xing Jiang, and Nin Xia. These classrooms are equipped with smart devices/sensors and specially developed software for distance learning. For example, the touch screen of the room displays presentations (e.g. PowerPoint), while also acts as a whiteboard for handwriting. The instructor can write on materials projected on the screen using a laser E-Pen. To optimize the video quality, a pan-camera can follow the instructor when he/she moves around in the classroom. RFID (Radio Frequency IDentifier) tags are used to identify and track students. Another tracking camera is mounted in the front of the classroom and it captures students’ attention status by recognizing the ‘blink frequency’ of their eyes. During the class session, instructors can load their pre-prepared PowerPoint and Word documents and write on the whiteboard (even when they are away from the whiteboard). The students can also write individual notes on the instructors’ handwriting window. All these classroom activities are being transmitted through various technologies to students at a distance. They are also recorded and archived for later review. Using this hi-tech environment, the teacher can move freely in the room, use his body language to communicate, and also interact with learners naturally and easily as in a traditional face-to-face classroom.

Figure 1. pervasive e-Learning platform in Shanghai
This online college has about 17,000 students, and 99% of them are working professionals who attend school part time. Therefore, their academic backgrounds, knowledge, and skills vary a great deal. Given such diversity, it is important to provide personalized learning services. The Shanghai system has harnessed data-mining technologies to organize learning communities and provide learning content recommendation based on student profiles (L. P. Shen & Shen, 2005; R. M. Shen, Yang, & Han, 2003). The large number of students in this college and its expansive course delivery systems make it an ideal place to test new and emerging technologies.

**Affective e-Learning Models**

The goal of this study is to understand how learners’ emotions evolve during learning process, so as to develop learning systems that recognize and respond appropriately to their emotional change. This paper also proposes a prototype of an affective e-Learning model that combines learner’s emotions with the Shanghai e-Learning platform. As Picard (2004) stated, theories of emotional impact on learning need to be tested and further developed. Until today, there is still a lack of comprehensive and empirically validated theories about emotion and learning. In this experiment, we examined several of the existing emotion theories in learning to help construct our affective e-Learning model. We used Russell’s ‘circumplex model’ to describe user’s emotion space. We then used the Kort’s ‘learning spiral model’ as the starting point to explore the affective evolution during learning. Following is the description of these models and our proposal for a prototype of an affective e-Learning model.

![Figure 2. an example of the basic learning emotion space](image)

**Russell’s Circumplex Model of Affect**

In our search for an emotion theory we focused on dimensional models instead of cognitive appraisal model for user emotion modeling because they cover the feeling of emotional experience both on a low level and a higher, cognitive level. One well-established dimensional model is psychologist Russell’s circumplex model of affect (Russell, 1980) where emotions are seen as combinations of arousal and valence. In Russell’s model, emotions are distributed in a system of coordinates where the y-axis indicates the degree of arousal and the x-axis measures the valence, from negative to positive emotions. The Russell’s model is widely used in recent researches (Craig et al., 2004; Fagerberg et al., 2004; Kort et al., 2001; Leon et al., 2007; Picard et al., 2001). And most of these just explored from three to eight basic emotions. Though Kort et al. (2001) proposed five sets of about thirty emotions that may be relevant to learning, however, we believe that skilled human tutors and teachers react to assist students based on a few ‘least common set’ of affect as opposed to a large number of complex factors; thus, we carefully select a basic learning emotion set which we deem most important for shaping our affective learning model. The basic set includes the most important and frequently occurred emotions during learning, namely, interest, engagement, confusion, frustration, boredom, hopefulness, satisfaction and disappointment. They might not be placed exactly the same for all people when put in the Russell’s two-dimension emotion space, because this model focuses on subjective experiences.
Figure 2 is an example of two-dimensional basic learning emotion space. We anticipate revising this emotion set when the study progresses.

**Kort’s Learning Spiral Model**

Kort, Reilly and Picard (2001) proposed a four quadrant learning spiral model (Figure 3) in which emotions change while the learner moves through quadrants and up the spiral. In quadrant I the learner is experiencing positive affect and constructing knowledge. At this point, the learner is working through the material with ease and has not experienced anything overly puzzling. Once discrepancies start to arise between the information and the learner’s knowledge structure, he/she moves to quadrant II, which consists of constructive learning and negative affect. Here the learner experiences affective states such as confusion. As the learner tries to sort out the puzzle but fails, he might move into quadrant III. This is the quadrant of unlearning and negative affect, when the learner experiences emotions such as frustration. After the misconceptions are discarded, the learner moves into quadrant IV, marked by ‘unlearning’ and positive affect”. While in this quadrant the learner is still not sure exactly how to move forward. However, he/she does acquire new insights and search for new ideas. Once the learner develops new ideas, he/she is propelled back into quadrant I; thus, concluding one cycle around the learning spiral of Kort’s model. As the learner move up the spiral, cycle after cycle, he/she become more competent and acquire more domain knowledge.

Kort, Reilly and Picard (2001) also described the empirical research methods to validate this spiral model, and promised a future paper to report the results of the empirical research. There were some efforts of empirical research on this model, e.g. Kort & Reilly (2002) stated that “we are in the process of performing empirical research on this model”, and “ideally, the Learning Companion should observe and try to understand the processes a learner experiences during all of these quadrants; however, this is currently beyond the capabilities of the technology”(Kapoor, Mota, & Picard, 2001), and “The process of cycling through these emotions during learning is currently being investigated in the present project”(D'Mello et al., 2005). But to our best knowledge, there have been no empirical validation report about this model. We used this ideal learning spiral model as the starting point to explore learner’s emotional evolution during the learning process.

![Figure 3. Kort’s Learning Spiral Model](image)

**The Proposed Affective e-Learning Model**

For our study, we used Russell’s ‘circumplex model’ to describe learner’s emotions detected from biophysical signals, and used the Kort’s ‘learning spiral model’ as the starting point to explore learners’ emotional evolution during the learning process. Finally, based on our previous work (L. P. Shen, Leon, Callaghan, & Shen, 2007; L. P. Shen & Shen, 2005), we proposed a model of affective learning (Figure 4) focusing on how we could make use of the information when we have got the learner’s emotion states and their evolution. This affective learning model considers the contextual information of the learner and the learning setting, and generates appropriate responses to
The affective learning model used a combination of (a) a cognitive appraisal approach to affective user modeling (which infers emotions according to situations experienced by the user as well as the observable behavior of the user) and (b) a physiological approach. Figure 4 shows a high-level description of this affective learning model, which only displays the general factors involved. The upper part of the model was modified OCC (Ortony et al., 1990) cognitive appraisal model for emotions, and the lower part of the model was the physiology recognition of emotions, where these two models converge. One of the pattern recognition methods—Bayesian Networks (Duda, Hart, & Stork, 2000)—were employed to model the relations between the emotional states and their causal variables.

This affective model indicated that the user’s emotional states were related to learner profiles (e.g., learning preferences, cognitive skills, knowledge structure), goals of learning, and learner interaction with the learning system. The user’s emotional states would in turn influence the measurements of the available sensors. The advantage of having a model based on a Bayesian Network was that it could leverage any evidence available on the variables related to emotional states to make predictions for any other variable in the model. For instance, the user’s emotional state could be assessed using existing information on learner profile and the learning context, even in absence of reliable sensors. Or, we could assess both emotional state and learner profile from reliable sensors, learner behavior and system reactions. The emotion detection system helped the system to provide timely help or adequate content based on the emotions the learner is experiencing at a particular moment.

To keep it simple, our model included only a subset of the factors that could be taken into account to assess learner’s emotional reactions in e-Learning. This model had the following major components:

1. Learner Profile: only two learner characteristics were considered in this model: learner preference and competency. Learner preference included learning content type (e.g., introduction, examples, cases and exercises), interactive type (e.g., active, expository and mixed), learning habit (individual, collaborative) and cognitive traits, etc. These factors were generated from multiple intelligences theory (Gardner, 1993), which argues that people have differing analytic capabilities, some are good at mathematical equations, others prefer verbal descriptions, and others may prefer to manipulate graphical representations. Information about learner
preference was gathered through data mining technologies and learner surveys, and could also be modified by this learning model dynamically. Learner competency described learner’s current knowledge structure, which would be used to compute the learning gap between current and anticipated competencies. This gap analysis was used to identify appropriate learning content that could help the user progress towards his objective competencies.

2. Learning Goals: We identified four types of learning goals for students in Shanghai’s e-Learning test bed, that is, an online college that has about 17,000 students. These goals included: acquire-objective-competency, problem-solving, learn-randomly, and have-fun. Our student characteristics survey revealed that these were the common goals of the majority of the students. Among these four goals, the first two were about anticipated events, which the OCC model referred to as ‘Prospect Relevant’ and that affected the “Prospect Based” emotional category (hope, fear, satisfaction, relief, disappointment), with the other two concerning unanticipated events, which the OCC model refers to as Prospect Irrelevant and that affect the “Well Being” emotional category (joy, distress).

3. Learning Events: We developed three kinds of actions/services to help achieve the learner’s goal effectively: learning content recommendation, quiz, and entertainment content delivery (e.g. playing music). The learner’s mouse movements were also collected as an indication of engagement or confusion. We had completed a study about learning content recommendation (L. P. Shen & Shen, 2005), an intelligent recommendation mechanism using competency gap analysis and a learning object knowledge base. The entertainment content delivery was included because, when the learners become deeply frustrated or bored, entertainment such as jokes, music and video could help to lift their spirits in the same way as an expert teacher might do in the classroom.

4. Learner Emotions: This affective learning model included eight basic learning emotions: interest, engagement, confusion, frustration, boredom, hopefulness, satisfaction and disappointment, among which the last two emotions are prospect based. The two lower level emotion variables were: valence and arousal.

5. Biosensors: four kinds of physiological data were used in our system: heart rate (HR), skin conductance (SC), blood volume pressure (BVP), and EEG brainwaves (EEG).

6. Learning Object Knowledge Base: The knowledge base included the subject knowledge structure, ontology, learning objects, and an answering machine (R. M. Shen, Han, & Yang, 2002). It also included the affective video collection tool, which stored video clips with emotion tags.

7. The Emotion Detection system gave feedback to the system and helped the system to provide timely help or adequate content for the learner experiencing certain emotions.

The model we described above was used to incorporate the emotion states into the interaction between the learner and the learning system. Emotions could also be feed back to teachers in the classroom and back to the fellows in a team work. Expert teachers are able to recognize the emotional state of their students and respond in ways that positively impact on learning in face-to-face traditional classroom, but in the e-Learning case, there are large numbers of remote students in distributed classrooms and mobile users; thus the challenge is, how could a teacher circumvent this? We would provide a solution for such problems via the real-time feedback of students’ emotional states to the teacher that the teacher could then adapt the lecture style, speed and content accordingly. What’s more, as emotion plays an important role in interaction involvement and evolution, the teacher should be aware of the students’ emotional states when organizing discussion groups so as to enhance the information flow within the group by smoothing the emotion flow.

Experiments and Results

We used the same method as used in the Picard study (Picard et al., 2001), where data were gathered from a single subject over many weeks of time, standing in contrast to efforts that examine many subjects over a short recording interval (usually single session on only one day). Single subject experimental design is a recognized method in educational and psychological research. It is typically used to study the behavioral change an individual exhibits as a result of some treatment (Gay & Airasian, 2003, page 383). Reliability of single subject experiment can be ensured by using reliable instrumentation, repeated measures, and also by describing the experimental conditions in details (i.e., the conditions of measurement and the nature of the treatment) (Gay & Airasian, 2003). Our study carefully followed these guidelines.

Considering the versatility of the variables being studied (i.e., emotion), it seems appropriate to use one-subject design in this preliminary experiment. Ekman and his colleagues (1983) acknowledge that even simply labeled
emotions like “joy” and “anger” can have different interpretations across individuals within the same culture. For instance, the same emotion could elicit different physiological patterns from different subjects. For a pilot study, more than one subject will pose great challenge in finding significant physiological patterns. Through repeated measures, this study generated a larger data set than traditional affect recognition studies involving multiple subjects. A repeated measure of one subject in a longer period of time ensured consistent interpretation of the variables (emotions), and also showed the evolution of emotions during the learning process. In addition, in pervasive learning environment, learning is individualized and emotions might be too. The average responses from a group might not apply to individuals.

The subject in our experiments was a healthy male graduate student in the e-Learning lab of Shanghai Jiao Tong University. This experiment focused on gathering physiological data for real-time emotion detection, to explore the affective evolution during learning, and to study the effectiveness of emotion when adding to existing e-Learning platform.

**Experimental Method**

Collecting reliable affective bio-data can be challenging when comparing to computer vision or speech recognition. Cameras and microphones are reliable and easy to use, but a few factors could affect the reliability of the biosensors. For example, whether the subject just washed his hands, how much gel he applied under an electrode, how tight the electrodes were placed, and even the air humidity could affect the readings. What’s more, emotion is often subjective and even the subject might not be completely aware of his own feelings.

We carefully designed the experiment to obtain high quality physiological data for emotion analysis. The experiment was conducted in a learning lab, and continued for two weeks and totally twenty sessions, with each session lasting about 40 minutes. The subject was a senior undergraduate student from computer science department. He sat in the quiet comfortable lab, where he regularly worked and studied. One of his classmates worked in the same lab and helped him when needed, for example, fetching a book because the subject could not move too much with the sensor on, or discussing with the subject when he had some questions. A few other people were nearby but they were quiet and did not interrupt him during the experiment. The subject was preparing for graduate entrance examination, namely, the courses of Data Structure and Object Oriented Programming. He was free to select the learning content, textbooks, exercises, search online, or coding. He was asked to completely focus on studying, and was not allowed to take a break during a single session. During the experiment, his emotion was naturally elicited, responding to the learning situation. Therefore, this research setting was more natural and closer to the real-world settings, comparing to the experiments using guided imaginary technique (Clynes, 1977) and classified emotion picture set (Center for the Study of Emotion and Attention, 2001). In many of their experiments, emotions were initiated by the subjects (that is, subject purposefully elicit emotions such as actors), and might be just external expressions instead of internal feelings. The subject reported his own emotion from the emotion list whenever he felt any change, which was used to label the data. Even though the experiment took place in the subject’s regular place of study, he had some “unnatural” feelings. For example, he was not allowed to move greatly because movement could bring noise into EEG waves through the wired bio-sensors. In addition, he had to remember to report his emotion changes, which he considered as interruption to his learning.

Although our e-Learning model had eight frequently occurred emotions, in this preliminary experiment, we explored four distinctly different emotions: engagement in the quadrant I of Russell two-dimensional affective model (positive valence, high arousal), confusion in quadrant II (negative valence, high arousal), boredom in quadrant III (negative valence, low arousal), and hopefulness in quadrant IV (positive valence, low arousal). The smaller the set, the more likely we are to have greater classification success by the computer.

The data were collected from three sensors: a skin conductance (SC) sensor measuring electrodermal activity from the middle of the three segments of the index and ring fingers on the palm side of the left hand, a photoplethysmograph measuring blood volume pressure (BVP) placed on the tip of the middle finger of the left hand, and a pre-amplified electroencephalograph sensor measuring EEG activity from the brain whose electrodes were placed on PCz, A1 and A2 according to the 10-20 International System of Electrode Placement. In our case, three EEG electrodes were sufficient (Lèvesque, 2006). Sensors and sampling were provided by the Thought Technologies ProComp5 suite (Figure 5), chosen because the suite was small enough to attach to a wearable computer. Signals were sampled at 256 Hz (namely 256 samples every second). The ProComp5 could automatically
compute the heart rate (HR) as a function of the inter-beat intervals of the blood volume pressure, BVP and could also separate different kinds of brainwaves into $\delta$, $\theta$, $\alpha$, low and high $\beta$ with filters. The frequencies and relationships with emotion of each brainwave were listed in Table 1. Each sample of data set comprised 3 raw data (SC, BVP, EEG), the HR from BVP, five brainwaves from EEG, five power percentages of the brainwaves and the label with time tag.

![Figure 5. the Subject Wearing ProComp5 in the Experiment](image)

<table>
<thead>
<tr>
<th>Wave Type</th>
<th>Frequency</th>
<th>When wave is dominant</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$ Delta</td>
<td>0-4 Hz</td>
<td>Deep sleep</td>
</tr>
<tr>
<td>$\theta$ Theta</td>
<td>4-8 Hz</td>
<td>Creativity, dream sleep drifting thoughts</td>
</tr>
<tr>
<td>$\alpha$ Alpha</td>
<td>8-13 Hz</td>
<td>Relaxation, calmness, abstract thinking</td>
</tr>
<tr>
<td>Low $\beta$ Beta</td>
<td>15-20 Hz</td>
<td>Relaxed focus. High alertness, mental activity. Agitation</td>
</tr>
<tr>
<td>High $\beta$ Beta</td>
<td>20-40 Hz</td>
<td>anxiety</td>
</tr>
</tbody>
</table>

We also conducted sixteen 30-minute sessions for emotion-aware (EA) and non-emotion-aware (NEA) content recommendation comparison. During these sessions, we used the emotion detection system introduced following to feedback the subject’s emotion states to the content recommendation system and recorded the user interaction (interventions) with the system. In order for the comparison to be as accurate as possible, the two systems were conducted at a pseudo randomly selected sessions and with the same probability. As the result, eight sessions were selected for EA system and another eight sessions were selected for NEA system.

**Data Preprocessing**

For emotion recognition, we have got 20 sessions, with each session lasting 40 minutes and the sampling rate 256Hz. Finally we select 18 sessions, removing the other two because of the bad placement of the bio-sensors. Then we got a total of 11059200 samples (18 session * 40minutes/session * 60seconds/minutes * 256samples/seconds = 11059200). Such big data set would make data training and data classification very time-consuming. To make it more efficient, and according to the fact that emotion won’t change so frequently as much as 256Hz, we merged n samples into 1 sample. When n=256, then we would have 1 sample every 1 second; when n=2048, then we would have 1 sample every 8 seconds. We employed very simple fusion method that we computed the means of the non-oscillating signals (SC, BVP, HR, the power percentages of the 5 brainwaves) and the FFTs of the oscillating signals (EEG, 5 brainwaves from EEG) as the corresponding values of the resulting sample.
Classification and Results

We have got two sample sets for emotion detection: one sample every 1 second (data set I) and one sample every 8 seconds (data set II). And then two pattern classification methods were tested: support vector machine (SVM) and K-nearest neighbor (KNN) (Duda et al., 2000). SVM maps training vectors into a higher dimensional space and then finds a linear separating hyperplane with the maximal margin in this higher dimensional space. The mapping function is called the kernel function. We selected to use the radial basis function (RBF) kernel:

$$K(X_i - X_j) = e^{-\gamma ||X_i - X_j||^2}, \quad \gamma > 0 \quad (1)$$

The problem of SVM then requires the solution of the following optimization problem:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i$$

subject to $y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0 \quad (2)$

Where $K(X_i, X_j) = \phi(x_i)^T \phi(x_j), l$ is the number of samples in the training set, $x_i$ is the attributes, $y_i$ is the label. By using the software LIBSVM (Chang & Lin, 2007), we first found the best parameter $C$ and $\gamma$ with cross-validation, then used the best parameter $C$ and $\gamma$ to train the whole training set, and finally tested over the testing set. KNN was tested with $k=1$ to 10 with the function provided by MATLAB (http://www.mathworks.com/).

Results of the classification

The two pattern classification methods: LIBSVM and KNN were tested over the two data sets. Table 2 is the classification results we obtained where we only listed the better results testing on data set I and II. The data shows that when we just used SC, BVP and HR attributes, the rates were 68.1% and 60.3% respectively by SVM and KNN, and when we just used the power percentages of 5 EEG brainwaves, the rates were 67.1% and 62.5%. However, when we used these two groups of attributes together, the rates were as high as 86.3% and 75.2% respectively. We also found that the brainwave power percentages contributed more than the sheer EEG powers from FFT. The best recognition rate 86.3% was achieved by LIBSVM tested on data set I, that is, the data set of one sample every one second. Though other researches have got as high as 90% correct rate over 3 valence levels, we were the very few experiments detecting the real learner’s emotions in a natural and close to real world settings. And because emotions won’t change too much during learning (L. P. Shen, Leon, Callaghan, & Shen, 2007), it’s harder to detect emotions in the learning process.

Table 2. Recognition rates of four learning emotions

<table>
<thead>
<tr>
<th>Attribute Space</th>
<th>LibSVM</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC, BVP, HR</td>
<td>68.1% (data set I)</td>
<td>60.3% (data set II)</td>
</tr>
<tr>
<td>EEG power% for Brainwaves</td>
<td>67.1% (data set I)</td>
<td>62.5% (data set II)</td>
</tr>
<tr>
<td>SC, BVP, HR EEG power% for Brainwaves</td>
<td>86.3% (data set I)</td>
<td>75.2% (data set II)</td>
</tr>
<tr>
<td>SC, BVP, HR EEG FFTs</td>
<td>60.8% (data set II)</td>
<td>59.2% (data set II)</td>
</tr>
</tbody>
</table>

Affective Evolution Results

Kort suggested that learning behavior would manifest itself in a spiral-like form i.e. a series of linked cycles separated in time. In order to learn how emotion evolves during learning, we computed the emotion distribution for all the 43,200 samples of data set I (Table 3), where engagement and confusion were the most important and frequently occurred emotions in learning, and frustration was the least. This is consistent with the findings from AutoTutor project (Craig, Graesser, Sullins, & Gholson, 2004). The transition distribution (Table 4) showed that the subject’s emotion changed frequently between engagement and confusion in both directions, and then frequently from confusion to boredom. There occurred no loop during single session and one loop within two successive
sessions. Of all the 18 learning sessions there were only one loop which could not be used to confirm the Kort’s learning spiral model. However, we hoped these initial results would prove encouraging to others who had speculated on this relationship and hopefully would motivate more detailed work on this aspect.

### Table 3: The emotion distribution for all the 43200 samples

<table>
<thead>
<tr>
<th>Emotions</th>
<th>Sample Numbers</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>17625</td>
<td>40.8%</td>
</tr>
<tr>
<td>Confusion</td>
<td>13738</td>
<td>31.8%</td>
</tr>
<tr>
<td>Boredom</td>
<td>7992</td>
<td>18.5%</td>
</tr>
<tr>
<td>Hopefulness</td>
<td>3845</td>
<td>8.9%</td>
</tr>
<tr>
<td>Total</td>
<td>43200</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Table 4: The transition distribution for all the 86 emotion transitions

<table>
<thead>
<tr>
<th>Transitions to</th>
<th>Engagement</th>
<th>Confusion</th>
<th>Boredom</th>
<th>Hopefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>×</td>
<td>32</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Confusion</td>
<td>18</td>
<td>×</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Boredom</td>
<td>1</td>
<td>5</td>
<td>×</td>
<td>6</td>
</tr>
<tr>
<td>Hopefulness</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>×</td>
</tr>
</tbody>
</table>

### Affective e-Learning Model Results

The affective e-Learning model we implemented was based on our previous intelligent content recommendation work (L. P. Shen & Shen, 2005), and added four emotion-aware recommendation rules as displayed in Table 5. We conducted eight 30-minute sessions for emotion-aware (EA) and non-emotion-aware (NEA) content recommendation respectively.

Enhanced user satisfaction is one of the most important objectives of the learning content recommendation system. User satisfaction could be evaluated by analyzing the number of times the user had to interact with the system in order to adjust the content recommended by the system. Manual adjustments meant that the system failed to deliver the content what the user expected and needed at different situations. Experiments indicated the superiority of EA with only 11 user interventions for the entire 8 sessions, whereas NEA was adjusted 21 times, an increase of 91% (Table 5).

### Table 5: The Emotion-aware Recommendation Rules and User Interventions

<table>
<thead>
<tr>
<th>Emotion states</th>
<th>Rule for that emotion</th>
<th>Number of user interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>Normal as NEA according to subject’s learning progress and objective</td>
<td>3</td>
</tr>
<tr>
<td>Confusion</td>
<td>Deliver examples or case studies for current knowledge</td>
<td>4</td>
</tr>
<tr>
<td>Boredom</td>
<td>Deliver video/music according to subject’s preference to ease the tension</td>
<td>1</td>
</tr>
<tr>
<td>Hopefulness</td>
<td>Deliver music according to subject’s preference to enhance meditation</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>11</td>
</tr>
</tbody>
</table>

### Conclusion

Using physiological signals to predict emotions, this study explored the emotion evolution during learning, and proposed an affective e-Learning model. Physiological data were gathered from one subject over many weeks in a natural and close-to-real world setting. The best-case classification rate (86.3%) was yielded by SVM (Support Vector Machine) with raw data, opening up the possibilities for instructors to understand the emotional states of remote learners. When brainwave signals were added to the other peripheral physiological data, the classification rate...
rose significantly from 68.1% to 86.3%. This increase suggested the close relationship between brainwaves and emotions during learning. When learner emotions detected were fed back to the affective e-Learning model in real-time, the content recommendation system could deliver proper learning content or entertainment content according to learner’s current emotional states and other learning contexts. Our experiments indicated the superiority of emotion-aware system over non-emotion-aware system with an increase of 91%, which suggested that using emotional data could greatly improve the performance of e-Learning system, particularly in the categories involving user-centered learning. Other noteworthy findings included that engagement and confusion were the most important and frequently occurred emotions in learning, and that using the power percentages of brainwaves yielded better results than using the sheer FFT powers of brainwaves. Since most current affective learning researches focused on identification of the users’ emotions as they interact with computer systems, such as intelligent tutoring systems (Craig et al., 2004), learning companion (Burleson et al., 2004) or educational games (Conati, 2002). There is a lack of research efforts in detecting emotions during learning in real-time, the emotion evolution during learning and using the emotional feedback to better design e-Learning system. Findings of this study should make a contribution in the research and development of affective e-Learning systems.

While this study has generated encouraging results, it has some limitations. Following are the issues that should be addressed in future studies.

1. Learner’s emotions are detected through biophysical signals only. Future studies should use multi-modal pattern analysis of signals, from face, voice, body and learners’ surroundings, to achieve more accurate results in emotion recognition.
2. In this study, the subject freely chose his learning content. To better explore emotion evolution during learning, researchers should design well-structured consistent lessons and materials focusing on learning a specific knowledge or course that can be used with more than one subject.
3. The affective e-Learning model resulting from this study was only for self-pace learning, it needs to be broadened to include other ‘types’ of learning such as classroom lecturing, group discussion etc.
4. The affective e-Learning model we proposed combined emotional feedback with learner profiles and the existing pervasive e-Learning platform. Future studies should explore how to leverage the affective information into the new pedagogical strategies. For example, at which emotion states will the learners need help from tutors and systems (when they are confused or frustrated)? What kind of help do they need at a specific emotion state?

In order to increase the reliability of results, future studies should also include a larger sample. At this exploratory stage, single-subject experiment is more feasible. The value of our work is to confirm a few general principles related to affective learning. And we anticipate a more systematic and in-depth study in the near future.

Acknowledgement

Affective e-Learning is a collaborative research project conducted between the Digital Lifestyles Center, University of Essex, UK (http://digital-lifestyles.essex.ac.uk) and the e-Learning Lab, Shanghai Jiaotong University, China (http://www.dlc.sjtu.edu.cn). During the formation of the research ideas and models, we have enjoyed many discussions with our colleagues whose intellectual contributions we are most grateful for. We would especially like to thank Vic Callaghan for his inspiring ideas and Enrique Leon for his valence recognition work. This work was carried out as part of the ‘Research on Context-aware Multi-modal Interaction Model and Key Technologies’ Project Supported by the High-tech Research and Development Program of China (National 863 project) under Grant No. 2007AA01Z157.

References


Wireless Handhelds to Support Clinical Nursing Practicum

Cheng-Chih Wu¹ and Chin-Yuan Lai²

¹Graduate Institute of Information and Computer Education, National Taiwan Normal University, 162, Heping East Road, Section 1, Taipei, Taiwan // Tel: +886 2 23622841 ext. 13 // chihwu@ntnu.edu.tw
²National Taichung Nursing College, 193, Sanmin Road, Section 1, Taichung, Taiwan // Tel: +886 4 22196964 // yuan@mail.ntcnc.edu.tw

ABSTRACT
This paper reports our implementation and evaluation of a wireless handheld learning environment used to support a clinical nursing practicum course. The learning environment was designed so that nursing students could use handhelds for recording information, organizing ideas, assessing patients, and also for interaction and collaboration with peers during an on-site clinical practicum. Our wireless handheld learning environment was field tested during a three-week practicum session. Analysis of data showed that both the instructor and the student benefited from using the Personal Digital Assistant (PDA) environment. The handhelds not only provided students with scaffolds to enhance learning but also facilitated peer cooperation and interaction with the instructor. Issues resulted from our implementation included things like the capacity of handhelds, network access, and participants’ preconceptions on using PC-based systems.

Keywords
Wireless, Handhelds, Clinical nursing practicum, Nursing education, PDAs

Introduction

“I was lucky to find a seat on the bus on my way to the hospital. I took out my PDA and checked the concept map I had drawn last night. I found a link was missing! I added it. …”

“At the daily morning report, a doctor and his nurse reported how they handled a patient who cried for leaving the hospital and disrupted the ward last night. I recorded their report and wrote down key terms with my PDA. … I later looked up the terms in my PDA’s Nursing Dictionary.”

“The patient was curious about our PDA. We explained that it was for recording conversations and writing assignments. … After the clinical interview, we returned to the nursing station and used the PDA to log into the course server to fill out the Patient’s ABC record. … I saw my patient getting up from the bed and chatting with his neighbors, and then walking to the yard. He looked to be weighted down with a heavy heart. I switched my PDA to the Symptom Assessment page and record what I had observed. Unfortunately, the speed of the network was somewhat slow, and I couldn’t finish the task. …”

“The instructor first asked everybody to transmit the concept maps to her notebook and then discussed with everyone individually. … At the end of the meeting, the instructor asked us to download reading materials from her notebook to our PDA.”

“On the bus going home, I didn’t use the PDA for fear of running out the battery. The first thing I did after going home was to charge my PDA and to synchronize it with my PC ….”

(Adopted from students’ journals in the study)

The use of handheld computers in clinical settings has been on the rise in recent years. Handheld computers have served not only as personal organizers but also as facilitators of efficient clinical diagnostic reasoning (Rempher, Lasome & Lasome, 2003). Surgeons, for example, have used PDAs to simplify data entry and retrieval at the patient’s bedside (Fowler, Hogle, Martini & Roh, 2002). Family practice residencies have also used handhelds as medication reference tools, electronic textbooks, clinical computational programs, and tracking of patient information (Criswell & Parchman, 2002). Healthcare professionals use PDAs to find appropriate medications, check for drug interactions, calculate dosages, and consult other clinical references (Tooley & Mayo, 2003). Today’s nurses have even used handheld-computer-based support systems to obtain patients’ preferences for functional bedside performance (Ruland, 2002). The use of handheld technology in clinical practice has been well documented. Huffstutler, Wyatt, and Wright (2002) suggested that these advanced technologies must be integrated into the nursing curricula to foster students’ proficiency in adapting to varied and expanding complex information systems.
The theory-practice gap in clinical nursing education has long been documented in the literature as a critical area that needs improvement (e.g., McCaugherty, 1991; Hewison & Wildman, 1996; Landers, 2000). Landers’ (2000) overview of the literature suggests that nursing instructors need to identify methods where students can more readily recognize the theoretical constructs underpinning successful nursing practice. Nursing instructors also need to play a key role in facilitating students’ application of clinical learning practices. Traditional clinical teaching has previously required one faculty member to supervise many students in a variety of settings, which has often included remote/distant sites. Geographic dispersal has posed challenges to faculty and students in the realm of timely responses to student needs, clinical monitoring, and student learning (Ndiwane, 2005). Even though the clinical area is considered to be at the center of nursing education, students still often express dissatisfaction with their clinical training: citing inadequate support from teachers in the clinical setting (Clifford, 1992; Lee, 1996).

Many initiatives have been undertaken in an effort to bridge the theory-practice gap. An important innovation is to support students’ learning in the clinical setting with modern technology. The incorporation of emerging technology into clinical areas has demonstrated its potential to provide more student learning opportunities, create innovative teaching practices, and promote current, accurate information retrieval systems for nursing care (Jeffries, 2005). In recent years, handheld technology has been considered as a possible application for supporting students’ learning in clinical practice. For example, Lehman (2003) used handheld computers to observe and to maintain records of students’ performance of clinical tasks in a clinical setting. Ndiwane (2005) used the Nightingale Tracker system, which connected a server and handheld devices with telephone lines, to allow timely communication between clinical faculty and students in community-based clinical settings. Their system also allowed for easy transmission, retrieval, and management of student-generated clinical data. Faruque et al. (2005) applied geospatial information technology to support academic practice, faulty outreach, and education initiatives at a nursing education institution. Studies in nursing education have found that personal digital assistants (PDAs) were an effective student learning resource, especially for reference materials, during students’ clinical practicals (Miller et al., 2005); and PDAs may facilitate the application of evidence-based knowledge to clinical practice (Stroud, Erkel, & Smith, 2005). Greenfield (2007) provided nursing students with PDAs equipped with a drug information program and found that some medication errors could be reduced at the point of care. Farrell (2008) found that most nursing students used PDAs in real time at the bedside primarily to access drug information and they agreed that PDAs enhanced their pharmacological knowledge (but not their contextual knowledge). The small and light-weight handhelds are easy to carry throughout learning activities, which makes it particularly useful for using in clinical sites. As White et al. (2005) pointed out, PDAs were probably the most useful tools for supporting students in clinical settings, where the pace is fast and resources may not be readily available.

A recent guest editorial by Billings (2005) in the Journal of Nursing Education, in response to the increased use of handheld devices by health care agencies and classrooms, urged nurse educators to take advantage of the power of mobile wireless technologies to create more learner-centered teaching. Wireless technologies such as WiFi, Bluetooth, and Infrared (IR) extend the power of handhelds so that students can use the devices to access Internet, to share and exchange files, and to submit and download data. Handhelds may also promote interactions among students and teachers (Lai & Wu, 2006). In our survey of articles in nursing education, we found that very few studies have incorporated the wireless capability of handheld devices in clinical training practices. Our literature search found only two studies that used the wireless features of handhelds in clinical training. White and et al. (2005) used only the wirelessly beam (IR) feature in their study, and students beamed their document to the instructor’s PDA in the clinical site once a week. The study by Garrett and Jackson (2006) equipped students’ PDA with WiFi and cell phone/GPRS (General Packet Radio Service) and showed that students mainly used the PDAs as electronic reference tools rather than for data recording and communication. The wireless features of PDAs were used limitedly and not actually incorporated into students’ learning process in the two studies.

Our present study is an attempt to respond to the challenges advocated by Billings, and to respond to the needs of nursing education practitioners. When placed in clinical settings, nursing students usually have limited opportunities to access computers, as most nursing sites allow computer access to “employees only.” They also have limited opportunities to interact with their nursing instructors. The aim of this study was to implement a wireless handheld learning environment, which took advantages of the mobility and wireless connectivity of handhelds, to support students’ learning in the clinical setting and to facilitate increased interactions among students and teachers. This paper reports our implementation and evaluation of such an environment in support of students enrolled in a Clinical Nursing Practicum course, an essential course for helping student link theory and practice during their final stage of professional training.
The wireless handheld learning environment

The environment was developed through a close collaboration with an experienced nursing instructor. The instructor has taught psychiatric nursing clinical practicum at a junior nursing college for over 20 years. She had participated in a study of using handhelds in a classroom environment and had modest knowledge and skills with PDAs. We identified the various needs of an effective clinical practicum by conducting several rounds of interviews with the instructor and observing the actual practices of her clinical learning sessions in a local hospital. Functionalities of the environment which would facilitate the instructor’s utilizing handhelds in the practicum were identified and implemented. The handheld learning environment was designed to support nursing students in recording information, organizing ideas, assessing patients, collaborating with peers, and interacting with the instructor during an on-site clinical practicum.

Figure 1 shows the hardware components of our wireless handheld learning environment. Students who enrolled in the course were provided with a PDA equipped with IR, Bluetooth, and WiFi wireless capabilities. They could use IR to share files with each other, Bluetooth to submit (download) materials to or from the instructor’s notebook, or WiFi to access the Internet. The instructor’s notebook and PDA were also equipped with wireless capabilities. A course server was hosted under the campus LAN (Local Area Network) of the nursing college. The instructor and students could use their PDA to access the Internet and the course server from the hospital, by ways of wireless access points (APs), which then connected to the hospital LAN and to the campus LAN.

Figure 2. The menu page of the handheld learning system
We have developed a series of application tools to meet the instructor’s needs, of which some were general-purpose tools (such as message board and discussion forum), while others were designated for psychiatric nursing practice (e.g., symptom assessment scales, nursing dictionary). Figure 2 shows the menu page which appeared when students logged into the course server using the Pocket PC IE browser. The Announcement feature shows course information and instructor’s messages to students. Patient’s ABC and Symptom assessment are online tools to support students’ observations of patients’ mental status. Reflective journals was the place where students could reflect upon what occurred during their daily practicum. Students could check the meaning of a nursing term and then listen to its pronunciation in the nursing Dictionary. Students could locate helpful documents such as PDA usage tips, nursing assessment procedures, and other useful resource links in the Resources section. Online discussions with peers were facilitated by using the Forum icon. Students could use the Interview tool to keep records of audio conversations with patients. The Portfolio feature allowed students to organize homework and individual learning materials. A concept map application tool and an audio-recording program were also included in the PDAs for student use. We describe several of the tools in detail below.

**Patient’s ABC**

A mental patient’s Appearance, Behavior, and Conversation (ABC) with others are indicators of the patient’s mental state. One of the nursing student’s daily tasks is to observe and record their patient’s ABC’s and submit them to the instructor at the end of the day. It also serves as a basis for students’ clinical interviews with patients. This task is traditionally conducted by using a paper-based observation sheet. We incorporated the traditional observation sheet into our handheld system so that students could denote their observations on a patient using their PDA and submit it to the instructor online. Figure 3 shows a student (Kelly) who was editing a patient’s ABC page. The [Hint] buttons (when tapped) provide students with suggestions for observing a patient. All students’ completed patients’ ABC records were kept in the course server and could be accessed by instructors and students when needed.

![Patient’s ABC page](image)

**Symptom assessment**

Several important symptom assessment scales were introduced to students in their theory courses before entering the practicum session, including the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1983) and its complementary instrument the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984). When conducting a standard clinical interview, students need access to the two scales in order to evaluate patients’ symptoms. In the past, the instructor did not carry out the symptom assessment activities during the practicum due to
the lengthiness of the scales, which is inconvenient to carry around when exercising nursing tasks. The SANS scale has 25 assessment items, each with 5 to 6 symptom descriptions, while the SAPS scale consists of 34 items, each with 6 symptom descriptions. We have incorporated the two scales in our system so that students could conduct assessment tasks using their PDA. Figure 4 shows a page from the SANS scale which lists assessment categories of hallucination. Figure 5 is the page which follows when students select the “Auditory Hallucination” category. Students would then need to check one of the symptom descriptions which best suited the patient. Once students finished the assessment and tapped the “submit” button, the assessment results would be transmitted to the database in the course server.

Figure 4. Symptom assessment for hallucination

Figure 5. Checking a patient’s symptom of auditory hallucination
Reflective journals

Students were required to write daily reflective journals in the hospital to improve awareness of their own learning. Reflective journals allowed students to better assess their needs, and at the same time, allowed the instructor to provide them guidance and support. Figure 6 is a reflective journal written by a student, which read as follows:

“… Although there were nurses and the therapist around, we felt scared because of being surrounded by so many patients. I felt very uncomfortable. …. I know I shouldn’t feel this way. I should try to get along with the patients -- no matter how seriously ill they are. …”

In the following mood check-boxes, one student indicated the feeling of being terrified and helpless. After seeing the journal, the instructor immediately wrote words of encouragement on the online feedback form, and then provided additional support for the student the following day. With the PDA-based reflective journals, the instructor could know students’ conditions promptly and provide them with timely feedback, which was just not possible with previous paper-written journals.

Figure 6. The reflective journals

Figure 7. The psychiatric nursing glossary
Nursing dictionary

We constructed a glossary of frequently used psychiatric nursing terms (see Figure 7) as part of our handheld system that students could use during their practicum. Students could check the meaning of a nursing term (in Chinese) and listen to its pronunciation (in English) via their PDA. The dictionary tool is especially helpful when students read a patient’s medical history and when for discussion with the instructor or peers (where the use of English terms was essential).

Discussion forum

To promote interactions among students, the instructor, and the nurse mentors, a discussion forum was created on our handheld system. Students could either respond to the instructor’s postings or initiate their own topics of interest by using their PDA in the hospital or via the PC at home. The nurse mentors could use PCs to join in discussions where ideas and experiences from the field could be shared. Each week during the practicum, the instructor initiated a discussion topic regarding issues involving students’ clinical practice. For example, in Figure 8, the instructor posted the question “Should patients call nurses by their first name?” Calling a person’s first name is acceptable in most western cultures but is uncommon in Taiwanese culture. In Taiwanese culture, calling a person by their first name is done only with very close friends or family members. The posting, consequently, generated much discussion among the students and with nurse mentors.

![Figure 8. The discussion forum](image)

Concept maps

The instructor encouraged students to construct concept maps with their PDAs to help students organize their ideas on assessing patients’ mental problems, performing nursing diagnosis, and outlining nursing care plans. Traditionally, students used paper and pencil methods to draw concept maps. In this study we offered students a PDA-based concept mapping tool, PicoMap, developed by the labs of Highly Interactive Computing in Education (Hi-CE) of the University of Michigan (http://hice.org/pocketpc). The PicoMap tool allows students to create and edit concept maps and to exchange such maps via IR beaming. Figure 9 is an example concept map in which a student depicted her assessment of a patient’s condition from physical, emotional, intellectual, social, and spiritual perspectives.
Methodology

To evaluate the effectiveness of the handheld learning environment, we incorporated the use of PDAs in a three-week nursing clinical practicum session conducted by the nursing instructor (who collaborated with us in developing the environment) in a clinical setting. The practicum session was designed to provide students with clinical experiences in the area of psychiatric nursing. We were interested in knowing the benefits and problems of implementing such an environment in an authentic clinical setting. Our evaluation focused on answering the following three questions: (1) what were students’ feedback on the features of the PDA-based environment in terms of perceived advantages and disadvantages?, (2) what were the benefits of the PDA-supported clinical practicum as compared to the previous ones?, and (3) what issues might occur when implementing such a PDA-based practicum session? Various qualitative data resources were collected in order to address the three questions. Although it might be preferable to implement a control-group experimental design in our evaluation, the small enrollment size of the practicum session made this approach infeasible. Alternatively, we asked the participants to compare their experiences of the previously attended traditional practicum sessions with the present one so that the effectiveness of the handheld learning environment could be identified.

Participants and settings

Six female students enrolled in the practicum session were each provided a PDA to use throughout the practicum. The PDA used was a HP iPAQ 2210 model powered with a Microsoft Windows CE operating system, 64MB memory, and an 802.11b wireless network card. The PDA came with many built-in application tools such as Pocket Word, Pocket PC IE, and Media Player. Students were trained in using PDAs prior to attending the clinical practicum. Two training sessions were held at the nursing college; each session lasted one and a half hours. Training activities included operations on data input, file management, software installation, wireless settings, Internet surfing, file transmission via IR and Bluetooth, synchronization of the PDA with PCs, PicoMap, and trial use of the handheld learning system.

The clinical practicum setting in the study was a private middle-size mental hospital located in central Taiwan. Two experienced nurses from the hospital served as mentors to guide students’ clinical practices. All nursing stations in the hospital had computers which were designated for medical management and students were not allowed to use...
Learning activities

Students’ primary daily activities in the hospital included attending to the hospital’s morning report, practicing morning care and routine nursing, participating in patients’ team therapeutics, and participation in a discussion session held by the instructor. The six students were paired into three teams and each team was responsible for caring for an assigned patient. The team members collaborated on most learning tasks and were encouraged to share and discuss their individual assignments with all other students. Students’ assignments included patient’s ABC observations, symptom assessment, clinical interviews, drawing concept maps, reflective journals, and in the end, outlining a nursing care plan. Students usually did the assignments during their free time in the hospital. While at the discussion session, the instructor closely monitored students’ daily progress, and the students used their PDA to submit homework and to download learning materials from the instructor’s notebook (via Bluetooth). Students often wrote their daily reflective journals and joined in the online discussion forum during their free time in the hospital or after returning home.

Data collection and analysis

Various data sources were collected for analysis, which included: students’ reflective journals; open-ended questionnaires answered by the students, the instructor, and the four nurses (including two nurse mentors) who worked in the same ward with the students; interviews with the students and the instructor after the practicum; and our field observation journals. Due to the small size of the participants, our analysis of data was primarily qualitative rather than quantitative. The authors first examined the collected data and identified contents which could be placed into the three evaluation question categories. Similar contents which addressed the same theme were then combined. Finally, arguments in support of various themes were justified. The results discussed in the following session were based primarily on the questionnaire data, along with the support from other data resources (such as interview data and field observation journals).

Results and discussion

Feedbacks on the PDA learning environment

In the questionnaire, we asked students to give their comments on advantages and disadvantages of the PDA-based features in the learning environment. Table 1 is the summary of the number of students reporting on each feature. Almost all students reported that they had benefited from the features provided by the PDA environment. Some problems were encountered, however, when using some of the features. Students indicated that the use of Patient’s ABC had saved a great amount of time because they could record patient’s data on site and upload it for instructor’s feedback soon afterwards. Some students did complain, however, about the occasionally slowness or connecting problems with the network (via WiFi). Students considered the Symptom assessment feature very convenient because they could revise the assessment scales whenever they observed that the conditions had changed for their patient. One student commented: "By allowing us to revise the assessment scales, we could see not only the progress of a patient but also of ourselves." Some suggested that it would be good to allow viewing the history of previously filled assessment scales, instead of only the latest version, so that the progress of their patient could be compared and analyzed. One student also indicated problems on connecting to the network. In using the Reflective journals, students appreciated the convenience when reflecting upon the day’s learning. This also provided a means for getting prompt feedback from the instructor. One student said that a disadvantage of the PDAs was the lack of “emotional icons” when writing the journal. Another student indicated her occasionally difficulty on accessing the Internet when
using the feature. All students considered the Discussion forum a good place for sharing experience and gaining alternative viewpoints.

As to using Bluetooth and IR, students loved its “quickness” and proved to be less troublesome because files could be transmitted directly between two PDAs without first connecting to a computer network. The only problem students mentioned was that the instructor had to remember to enable her Bluetooth before the transmission. All students enjoyed sharing files with peers via IR and thought it was convenient in facilitating discussion and learning from others. In using PicoMap software to draw concept maps, the benefits students reported were as follows: easy and fast to draw, easy to comprehend and memorize, ready for use when a thought is generated, and the ease of creating “clean” concept maps. Students found it inconvenient, however, in that the PicoMap only allowed 12 characters per node with only 32 nodes per concept map. This limitation was due to the small size of the PDA screen, which also prevented students from viewing an entire concept map (which consisted of more than 8 to 10 nodes).

Table 1. Number of students reported advantages and disadvantages on the features of the handheld learning environment

<table>
<thead>
<tr>
<th>Feature</th>
<th>advantages</th>
<th>disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient’s ABC*</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Symptom assessment*</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Reflective journals*</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Discussion forum*</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Homework submission via Bluetooth</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Sharing data via IR</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Concept maps (PicoMap)</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Note. * Accessed via WiFi network

A major problem raised by the students in using the PDA environment was the “connectedness” of the WiFi network. Due to the somewhat dated network infrastructure of the hospital, we could only manage to set up two APs in the hospital: one near the ward and the nursing station, and one in the meeting room. Students found it inconvenient when they tried to access the network at other locations in the hospital. Some students attempted working in the yard of the hospital, and found the wireless signals to be weak or even non-existent. Students also indicated occasional slowness of the network while accessing the Internet or the course server, especially when there were many individuals vying for access. The sluggishness of the network was attributed to the hospital’s small bandwidth ADSL connection with the Internet, with a 768 Kb downloading and 128 Kb uploading rate. We attempted to resolve the situation by modifying some of our application tools. For example, the Symptom assessment scales could originally only be used when connecting to the Internet. We added the offline viewing/editing feature so that students could also use it even without network access. Students could submit their work later when they had access to the network. These modifications provided students with more flexibility and mobility in using the PDA tool. This also reduced network traffic which, in turn, improved network access speed. Nevertheless, a broader network bandwidth and wider wireless coverage would still be desirable for future implementations.

Comparison with previous practicum sessions

Questionnaires and after-practicum interviews were used to ask students, the instructor, and the four nurses about their observed differences between the present PDA-using practicum session and the previous non-PDA-using practicums. We summarize their comments (below) in terms of the learning process, interactions among peers and instructor, and course management.

Learning process

The most cited differences which students stated were in regard to taking notes and recording conversations with patients. These two features provided by the PDAs spared students from memorizing excruciating details and made their learning more efficient. Students pointed out:
“The major difference was that I could now take notes whenever I wished. Without the PDA, we had to try very hard to recall what was done that day.”

“It saved me lots of time in doing homework. Whenever I had a thought or found useful information, I wrote it in my PDA. After going home, I did my homework by referring to what I had written. It now only took me one or two hours to complete the homework -- compared to four or five hours before.”

“The difference was that we had opportunities to use the high-tech product. It was convenient to take to anywhere, record information, and learn at anytime. We indeed gained a lot from using it.”

We have found that students used PDAs to do homework, take notes, record talks, and look up terms wherever needed (at the hospital, on the bus, or at home). Similar observations were also described in White et al. (2005).

The four nurses were asked about students’ performance in the PDA-enhanced clinical practicum versus the older, more traditional method. The nurses agreed with the students about the benefits of taking notes and recording conversations using the PDAs. In addition, they thought the use of PDAs to access reference materials in the PDA or resources in the Internet had greatly helped the students. The nurses pointed out that “PDAs help students find out a patient’s problem promptly.” and “students could use PDAs to search for answers anytime.” Another noted difference stated by the nurses was that students had integrated more theories into practices, which was due to the abundant materials and useful tools provided in the PDAs. One of them wrote:

“... This time the students could review theories first [by using their PDAs] and use PDAs on the spot, which saved both time and effort. Drawing concept maps gave them a clear picture of the concepts to be learned, and the maps could also serve as a review summary.”

Koeniger-Donohue (2008) also found that the use of PDAs not only saved students’ time, but students also spent more time with patients because it was not necessary to leave the patient’s room to look up information. By drawing concept maps on their assessment of a patient’s conditions, students would clearly organize their ideas and see what information they already had or were missing and yet to be collected. The nurses expressed that the students of the present session were more organized and analytical in learning.

The instructor observed the differences from more pedagogical viewpoints. She considered her students became engaged and self-directed in learning, attained better theory knowledge, and had stronger self-confidence. She believed that the symptom assessment scales and the concept mapping tool had played a crucial role in the improvement of students’ learning. When assessing a patient’s problem, students had to follow the assessment scales item by item to determine one’s symptoms. In doing so, they clearly learned the definitions of symptoms and the procedures for doing assessment. Thus, the PDA-based tools acted as learning scaffolds to support students’ assessing patients, organizing ideas, and allowed for reflection of students’ learning. The benefits indicated by the instructor were consistent with previous studies. Miller et al. (2005) found that students decreased their reliance on textbooks and clinical faculty after substantial use of the PDAs. PDAs also supported students’ formulation of questions associated with clinical situations. Ndiwane (2005) found students were confidence in providing patient care independently while using PDAs.

In summary, students enjoyed the convenience and time-saving aspects that the PDA tools had brought to the learning process, which were confirmed by the nurses’ observations. Nurses in the clinical sites indicated that students integrated more theories into practices and became better organized in learning. The course instructor, believed the PDAs served as a scaffolding and support system during students’ clinical practices, which resulted in engaged, self-directed, and more confident students.

Interactions among peers and instructor

In the practicum, students were paired in teams and had to collaborate on several activities and assignments, such as taking care of an assigned patient, conducting clinical interviews, and writing a nursing caring plan. All students commented that they had better interaction with peers when compared to their previous non-PDA-using practicum sessions. They attributed the effects to the immediacy and the file transmission features offered by the PDAs. With the PDA in hand, students could finish tasks far sooner than before and could easily exchange each other’s work (via IR) for discussion. The results were in accordance with our previous study (Lai & Wu, 2006) in which PDAs were used to support cooperative learning activities. Student commented on the questionnaires about how they had interacted with others:
“... This time we discussed and finished tasks such as the Patient’s ABC, interviews, and observations on the spot which greatly enhanced our interactions.”

“We usually did reports together, first drawing maps individually, and then having discussion; or we divided the work so that we could finish it sooner, and then transmitted the maps to each other.”

“This time we could prepare the data on site and then transmitted it to each other during discussion. We would reach conclusions very soon. In our previous sessions, we prepared the data after going home and emailed to others. We could hardly have time for face-to-face discussion.”

The instructor expressed that she was now able to identify students’ problems sooner and provide more individualized instruction when PDAs were incorporated into the clinical practicum. She attributed her improved interactions with students primarily due to the inclusion of online reflective journals. She also enjoyed the experience of beaming files (using IR) to/from students and considered her relationship with students “greatly enhanced” by the high-tech use. Students, additionally, admitted that their interactions with their instructor were increased due to the use of PDAs. With the PDA handy, they could easily show their work to the instructor, or even transmit it to the instructor’s PDA or notebook, during practicum discussions. As a result, they could get timely feedback from the instructor, which in turn increased the frequency of student-instructor interactions. Students’ increased interactions with the instructor were supported by the following statements:

“We had more interactions because I could transmit my written materials to the teacher and discuss my problems soon after. …”

“The teacher received everyone’s work immediately [via Bluetooth] during discussion sessions. She knew clearly how much we had learned and gave us appropriate feedback accordingly.”

“It used to take a while for receiving the teacher’s feedback after we emailed our work. Now, we got the result right away via Bluetooth.”

Course management

The instructor conceded that the use of PDAs had spared her from trivial course management and minimized many time-consuming tasks that had existed with the more traditional method of administration. For example, by logging into the system, she would know who had turned in the assignments on time and who had not, which had freed her from the frustration of constantly having to remind students the deadlines. She appreciated the use of online tools such as Patient’s ABC, Symptom assessment scales, and Reflective journals which provided students with clear instructions and formats to complete assignments and allowed her to give students timely feedback. The instructor said she now had more time to guide students with in-depth discussions during the practicum session. As pointed out by Huffstutler et al. (2002), additional time would be needed for class preparation using PDAs as a teaching strategy, the instructor admitted that it was challenging for her to learn all the “technology stuff” and to manage students’ use of the PDAs at the beginning. However, she considered all those “sweet burdens.”

Implementation issues

Several implementation issues arose during the course of our study, which included the following: the capacity of PDAs, internet resources, and students’ preconceptions about using PCs.

PDA capacity

Several studies (e.g., Guerrero, Ochoa, & Pino, 2006; Garrett & Jackson, 2006; Lai & Wu, 2006) have identified limitations of using PDAs in learning: data input, screen size, memory capacity, processing power, and battery life. Several problems raised by our students in the study were due to some of these limitations. The most frequently mentioned problem was the fear of losing their data when the PDA’s battery was low. The PDA model used in our study had an approximate 8-hour battery life when fully charged, but the battery life is far shorter when wireless functions such as IR, Bluetooth, or WiFi were enabled or when the PDA screen was left on. The loss of battery power was very inconvenient for students. When batteries failed, the PDA would return to its default settings. This meant that the data in the PDA would be erased and would require reinstallation of software. When batteries were running low, students would interrupt their work either to charge the battery or to borrow a back-up PDA from the instructor. Students were reminded to fully charge their PDA before coming to the hospital; however, even the instructor sometimes would forget to charge the PDA. The instructor also complained about battery problems in the
questionnaire. We do not consider the battery problems encountered here to be a major obstacle for the incorporation of PDAs into clinical practicum training. One could easily resolve this by replacing PDAs with more current models, which have longer battery lives and which do not erase data and programs even when battery failure occurs. Another problem occurred in our study was some unexpected system halts of the PDAs. The unexpected system halts sometimes resulted from students’ improper use of the PDA. When an operation was not immediately responding (e.g., opening a link on a web page), some students would keep tapping on the PDA, causing system halts. Yet, some system halts may have been possibly due to the PDA’s processing power or memory capacity, or the stability of the software application. We believe the mature of mobile technology would eventually resolve the problems.

Internet resources

Although abundant nursing resources are available via the Internet, few such resources are designed or adapted for use with PDAs, especially those for Chinese web sites. Students often had trouble viewing Internet PC-based web pages with their small PDAs. Viewing conditions would undoubtedly be worsened if the web was designed with various character settings, many pictures, or animations. In the worst cases, some web pages would result in system halts with the PDA. We could only accommodate the situations by offering important resources with PDA-displays formatted on our course server. There are growing numbers of Chinese Internet subscription resources which can be accessed and viewed via PDA software -- such as AvantGo (http://www.avantgo.com/). The incorporation of such software in future implementations should enable students to access and to view more Internet resources via their PDA.

Preconceptions about using PCs

We have found that the participants’ conceptions about using PDAs were influenced by prior use of PCs. The PDAs in our study used a Pocket “Windows” operating system and had similar PC Windows tools such as File Manager, Internet Explorer (IE), Pocket Word, and Pocket Excel. The look and feel of the Pocket Windows system was similar to that of a PC except for its small size screen. All our participants were experienced PC users. This facilitated the easily transfer of PC skills to that of the PDA. We observed that PC-based skills worked well most of the time, but sometimes students became confused. Closing a PC application window, for example, usually results in terminating the application. In the PDA system, however, it only closes the window (but does not terminating the application program). The application program can only be terminated by entering a system setting page to explicitly stop it. If there are too many applications running at the same time, the performance of a PDA will become slow. Students often found their PDA system responded very slowly, or even halted, without noticing the cause of the problem -- because they thought they had “closed” (terminated) the application program.

Students also had to get used to the arrangement of menu application buttons in the PDA, which appeared at the bottom of the screen rather than at the top of a PC screen (e.g., Pocket Word vs. Word). Students were sometimes frustrated when failing to find a feature which was not available in the PDA or had different operating procedures than the PC (e.g., setting of character size in IE). Participants who had used a PC based system expected similar features to appear in the PDA. PDA based systems are, in fact, only “pocket-size” versions of those in a PC in terms of the hardware and its applications. The practicum instructor was even subject to having to adjust from the use of PC based applications to that of the PDA. The instructor used to grade students’ assignment in the PC Word and was very disappointed to find the “Revision Marks” feature absent from Pocket Word. Some of our difficulties in communicating with the instructor when developing the handheld system and conducting the experiment were due to her unintentionally mapping of the PC based models to the PDA. We believe that more training hours may be warranted to allow both the instructor and students to learn the limitations of a PDA and overcome conflicting models that result from prior use of a PC based system.

Other issues

We had initial concerns about disturbing regular nursing practices when introducing PDAs into the clinical site. After conducting the study, our initial concerns were proved to be unnecessary. All of the nurses involved in the study were very positive about students’ use of PDAs and did not consider it a problem to the nursing practices. The nurses responded in the questionnaire:

“It simplified the task procedures and caused no troubles to the hospital and the patients.”
“It helped us understand the future trend of using technology in nursing practices and we acquired new information. Using PDAs is convenient for keeping records of nursing processes.”

The nurses envied students’ access to PDAs and expressed the desire to use PDAs in their future practice. The findings correspond to Farrell’s (2008) study in which the nurses encouraged students to use the PDA for accessing information in real time. Farrell also pointed out that some students in her study thought using PDAs in front of patients seemed rude. Our observation showed that the concern was not necessary. We found patients’ initial reaction to the PDAs was often that of curiosity towards the use of this technological device. Patients quickly got used to the students’ use of PDAs, eventually viewing them as almost routine tools of the clinical staff. Students had expressed few non-technical concerns about using PDAs in the clinical site. One concern which was voiced by almost every student was fear of losing their PDA in the hospital. PDAs were loaned to students for the three-week practicum period. They were reminded that the device was expensive and had to be well taken care. Due to its pocket-small size, students thought it would be easy to lose or to be stolen. Students felt a little insecure about carrying such an expensive technologically advanced device around the hospital throughout the practicum.

Conclusions

In the instructor’s final comments (expressed in the questionnaire), she wished that “one day all the clinical sites will become wirelessly connected.” The instructor stated that the handheld environment was a great “breakthrough” in what had been years of difficulties in conducting effective clinical practicum. The wireless handheld environment not only provided students with tools for taking notes, recording information, and accessing resources but also facilitated their collaboration with peers and improved interactions with the instructor. Ultimately, the PDAs acted as cognitive scaffolding tools to support student learning in the clinical site. Our findings were different from Garrett and Jacksons’ (2006) study, in which they found the PDAs mainly used as electronic reference tools, rather than data recording and communication devices. We believe the differences were mainly due to the design of the PDA-based tools, which did not closely integrate into the learning activities. Our implementation was a more integrative one where we incorporated several key clinical learning instruments into the PDA system. We have addressed the utilization of handheld devices in terms of pedagogical underpinnings by building the learning environment based on collaborative, contextual, and constructivist principles (as indicated in Patten, Sánchez, & Tangney, 2006). Our findings showed that the wireless handheld environment in a clinical nursing practicum was both technologically appropriate and pedagogically sound.

As we have discussed, there were limitations in our present approach such as sluggishness of the Internet access in the clinical site, short battery life of the PDAs, constraints on the PDA software (i.e., PicoMap), limited PDA-adapted web resources, and students’ confusion about the interface of PDAs with their previous PC using experience. However, the limitations were far overshadowed by the potential benefits. Slowness or inconvenience of hospital network access was mainly due to older network infrastructures which could be easily remedied. Issues such as short battery life could be resolved with newer model PDAs. We believe that some of the limitations and confusion about PDA based systems could be resolved with additional training. Some of the hardware and software problems would be alleviated by the maturation of future mobile technologies. PDAs have the potential of revolutionizing and transforming the way clinical practicum is conducted in nursing training settings. Finally, the results of the present study were primarily based on qualitative data collected from small number of participants, future study may want to recruit more participants and implement an experimental design to explore the effectiveness of the handheld learning environment from quantitative aspects.

References


A Coursework Support System for Offering Challenges and Assistance by Analyzing Students’ Web Portfolios

Liang-Yi Li and Gwo-Dong Chen
Department of Computer Science and Information Engineering, National Central University, Taiwan // Tel: +886-3-4227151 ext. 35327 // Fax: +886-3-4273485 // liang@db.csie.ncu.edu.tw // chen@csie.ncu.edu.tw

ABSTRACT
Students can practice skills and acquire knowledge by doing coursework. However, in conventional coursework activities, each student is assigned the same exercises, without considering learners’ diversity. Moreover, students typically have difficulty in receiving assistance for completing their exercises after class. Therefore, some students cannot learn effectively when doing their coursework. This work presents a Web Coursework Support System (WCSS) to assist students doing coursework within their Zone of Proximal Development (ZPD). This system applies decision tree analysis methodology for selecting programming exercises at a level of difficulty suited to each student. The assigned exercises thus challenge students. To assist students in completing these assigned exercises, this system allows students to access distributed online learning materials related to a programming concept through links on a single web page and motivates students to answer posted questions in a discussion forum. Experimental results show that the exercises-completed rate and the rate of the assigned exercises completed with assistance were increased. They indicate that WCSS can increase likelihood that students do coursework within their ZPD by offering challenges and assistance. Furthermore, some students’ responses were insightful in understanding the benefits and limitations of this system.

Keywords
Coursework, Concept map, Decision tree, Zone of Proximal Development

Introduction
Coursework is a significant learning activity to supplement classroom instruction (Cosden, Morrison, Albanese, & Macias, 2001). In conventional coursework activity, the teacher first assigns several exercises to all students. The students then do the exercises and hand them in. Finally, the teacher comments on each student’s work and credits each student with a grade. In the process, students can construct knowledge by reading textbooks, discussing topics with peers, and making inquiries through online resources (Cooper & Valentine, 2001; Epstein & Van Voorhis, 2001). A teacher can also evaluate the learning performance of students and understand their learning status based on the quality of their coursework.

However, the teacher assigns the same exercises to a diverse group of students who have varying learning statuses. Excellent students may feel that the exercises are too easy to teach them anything, while below-average students may feel that the exercises are too hard to allow them to learn (Corno, 2000). Additionally, learning aids, such as libraries and capable classroom peers, are located in different places. Students generally have difficulty receiving these aids to complete their exercises after class (Glazer & Williams, 2001). In these situations, students often fail to finish their coursework, or, occasionally, plagiarize.

Recently, some online assignment systems have been designed to support students and teachers in a conventional coursework activity. For instance, some systems provide assistance for teachers and students to manage the process of the conventional coursework activities, such as automatic assignment submission, assessment, and feedback (Collis, De Boer, & Slotman, 2001; Dawson-Howe, 1996; Lee & Heyworth, 2000; Saikkonen, Malmi, & Korhonen, 2001). The systems can help teachers manage the process of an assignment, and so reduce teachers’ workloads. However, they do not provide support for the assigning of appropriate exercises for each student or for students’ completion of these assigned exercises in the coursework activity.

Some systems provide personal tutoring that assigns adaptive questions for students and then guides students of varied abilities to correct their own assignment errors (Lilley, Barker, & Britton, 2004; Murray & Arroyo, 2002; Syang & Dale, 1993). These systems usually are applied in the Computerized Adaptive Test (CAT) domain to select the most appropriate questions based on Item Response Theory (IRT) (Lord, 1980; Rogers, Swaminathan, & Hambleton, 1991). However, in order to achieve reliable results, these systems require substantial interaction.
between a user and the system. Additionally, IRT is most suited to multiple choice or fill-in-the-blank questions and is less effective for open-ended questions, like programming exercises that need students to write a program to solve a problem (e.g., sum the odd integers between 1 and 99).

This study develops a Web Coursework Support System (WCSS), which is designed for supporting students who are learning the Java programming language in the coursework activity. It is composed of three subsystems: a Personalized Assignment Dispatching System (PADS), an e-dictionary system, and a peer recommendation system. The design of WCSS is based on the work of Vygotsky, who proposed Zone of Proximal Development (ZPD) as “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978, p. 86). Tharp and Gallimore (1988) proposed a four-stage ZPD model. In their model, ZPD is created when assistance is offered by more capable others or by the self-guidance. After learners progress through the ZPD, the skills and knowledge used within ZPD are internalized and fully developed. Therefore, teaching and learning occur only when assistance is provided to learners in their ZPD. Such assistance consists of six means of assistance: modelling, contingency managing, feeding back, instructing, questioning, and cognitive structuring (Tharp & Gallimore, 1988).

The WCSS provides three ways of helping students performing coursework within their ZPD. First, the programming exercises selected by PADS must be at the appropriate level of difficulty (this study views an exercise as being at the appropriate level of difficulty for a student only if it can be completed by the student under the assistance provided). The PADS analyzes students’ web portfolios to understand their knowledge status. According to the result of the analysis, the PADS select programming exercises that students potentially have difficulty in completing. Thus, these selected exercises can be seen as questioning assistance that asks students to apply and synthesize some programming concepts that they are not proficient now to complete their exercises. Second, the e-dictionary system that organizes useful learning materials with a concept map is designed to provide support for students to learn programming concepts related to these exercises. The concept map is used to help students understand the overview of Java programming and the relationship between different concepts before they select detailed elements to study. The learning materials associated with a concept provide a detailed explanation and related examples. Students can understand the meaning of a concept through its explanation and observe and imitate the related examples to learn how to apply this concept in a program (Astrachan & Reed, 1995; Schworm & Renkl, 2006). Thus, the e-dictionary system can be treated as a modeling and cognitive structuring assistance to help students completing these selected exercises independently. If students still cannot complete these exercises, they can post their questions on the discussion forum. Third, the peer recommendation system designates two capable classroom peers for each posted question and employs some strategies to motivate them to resolve the questions with the questioner. The peer recommendation system can be seen as a contingency managing assistance that encourages peers to resolve the posted questions. In addition, the recommended peers can be seen as a providing assistance by supporting students completing these programming exercises collaboratively. Ideally, students should do their coursework within their ZPD through challenging them and scaffolding them. Table 1 presents an overview of the design issues for WCSS. The first and second columns of Table 1 list the requirements and related reasons at the time that the system was designed. The third column of Table 1 indicates how to meet those requirements in the proposed system.

<table>
<thead>
<tr>
<th>Requirements</th>
<th>Reasons</th>
<th>Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assigning programming exercises that have a level of difficulty suited to each student</td>
<td>Assistance provided for learning through challenging engagement</td>
<td>PADS</td>
</tr>
<tr>
<td>Accessing web learning materials through links on a single web page</td>
<td>Assistance provided for completing exercises independently</td>
<td>E-dictionary system</td>
</tr>
<tr>
<td>Motivating capable classroom peers to answer assigned questions</td>
<td>Assistance provided for completing exercises collaboratively</td>
<td>Peer recommendation system</td>
</tr>
</tbody>
</table>

System Overview

In this study, students learned in a blended learning environment. That is, a teacher lectured for three hours weekly, and students took an exam every two weeks in the classroom. Students also learned via a web-based learning system after class. In the web-based learning system, students could discuss items on a discussion forum, submit coursework
assignments on the WCSS, access learning materials, including completed assignments and lecture slides, and perform keyword self-assessments. Some tasks, such as weekly keyword self-assessments, completing coursework assignments every two weeks, and taking an exam every two weeks, must be performed periodically.

All students’ web portfolios can be categorized as three types: learning products, learning behavior, and learning performance (Table 2). An e-dictionary system links these learning products with a concept map to allow students access learning materials related to a concept through links on a single web page. Additionally, to generate pedagogical rules for selecting appropriate programming exercises and suitable capable classroom peers, learning behavior and learning performance were employed to construct a student model. The student model includes students’ learning status for every programming concept in the course and learning performance in different learning activities. Based on the student model, the PADS uses a decision tree analysis methodology to generate a decision model that can determine whether a programming exercise is at the appropriate level of difficulty for a student. Moreover, the peer recommendation system is used to select two capable classroom peers for each posted question based on the student model. Figure 1 illustrates the system architecture of the web-based learning system.

<table>
<thead>
<tr>
<th>Category</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning product</td>
<td>Discussion threads, completed assignments</td>
</tr>
<tr>
<td>Learning behaviour</td>
<td>User logins, read /post /reply to a discussion topic, and submit /review /read completed assignments.</td>
</tr>
<tr>
<td>Learning performance</td>
<td>Exam grades, results of keyword self-assessments and coursework grades.</td>
</tr>
</tbody>
</table>

Table 2. Three web portfolio types

Concept mapping is a technique that is widely used to represent knowledge in a graphical format (Jonassen, Beissner, & Yacci, 1993; Novak & Gowin, 1984). A concept map is a node-link diagram, which represents a concept using a node and a relationship (for example, “is-a,” “related-to,” or “part-of”) using a link (Novak, 1991). Figure 2 shows a part of a concept map drawn by a teacher, representing partial object-oriented programming knowledge in a Java programming course.

Figure 1. System architecture of the web-based learning system
Concept mapping has been used for a variety of educational purposes, such as assessing learner’s knowledge status (Hoefl et al., 2003; Turns, Atman, & Adams, 2000), sharing meaning between people (Chiu, 2004; Hughes & Hay, 2001), planning the process of problem-solving (Chang, Sung, & Lee, 2003), organizing knowledge (Chang, 2002; Lee, 2004; Triantafillou, Pomportsis, Demetriadis, & Georgiadou, 2004), and representing a learner’s knowledge structure (Kuo, Lien, Chang, & Heh, 2004). In this study, a concept map drawn by teachers was applied to present students’ knowledge structure in Java programming and to organize learning products.

Each student has his/her own concept map to present the student’s learning status for every programming concept in the course. Each concept node in a student’s concept map has a value representing the student’s proficiency level in the concept. The value for each concept node is generated from the results of exams and keyword self-assessments using a formula. Additionally, learning products are connected to related concepts in a concept map to provide learning scaffolds. Three tables are used to store the concept map (table “Concept_Map” in Figure 3), students’ learning status for every programming concept (table “Learning_Status” in Figure 3), and the relationship between the concept map and those learning products (table “Learning_Material” in Figure 3). Figure 3 presents the three table schema and their relationship.

**Operations and interface of WCSS**

The teacher and Teacher Assistants (TAs) design ten candidate programming exercises based on the syllabus every two weeks. The PADS selects three exercises from these candidate exercises for each student, primarily based on the decision model in the PADS. A web page displaying three hyperlinks representing the three selected exercises is
shown when a student logs into the WCSS. The student can click one of the three hyperlinks to link to a respective exercise web page.

An exercise web page (Figure 4) consists of four sections: problem description, concept, discussion, and submission. A student receiving an exercise first must understand the programming problem and then plan the program based on the problem description. If the student does not know how to solve the problem, then the concepts shown on the concept section can remind the student which concepts should be adopted to solve the problem. Students who are not familiar with these concepts can learn about them on the e-dictionary system through the hyperlinks in the concept section. Students who still cannot solve the problem after using the e-dictionary system can post questions on the discussion forum via the hyperlink in the discussion section. Meanwhile, two capable classroom peers chosen for each question by the peer recommendation system are notified to resolve the assigned questions together with the questioner. Ideally, the student should be able to complete the exercise through this process. Furthermore, the exercise web page presents hyperlinks to the e-dictionary and peer recommendation system, giving students access to assistance, including learning products distributed in different web pages and online capable classroom peers, through links on a single web page.

![Figure 4. The layout of an exercise web page](image)

**Personalized assignment dispatching system**

The PADS applies decision tree analysis methodology to construct a decision model that can determine whether a programming exercise has an appropriate difficulty level for a student. The decision tree analysis originates from the machine learning discipline, which can induce rules for accurately predicting a target attribute from feature attributes.

C5.0 is a machine learning tool for generating decision trees and was developed by Quinlan (Quinlan, 1996). A trial version is available on the Internet at http://www.rulequest.com. The induction performed by C5.0 is based on the value of entropy (Mitchell, 1997). For instance, programming exercises can be clustered into a set of groups based on the values of feature attributes. If programming exercises in a cluster have the same value of the target attribute, then the cluster has the lowest entropy, while if every programming exercise in a cluster has a different value of the target attribute, then the class has the highest entropy. The decision tree tool is used to generate a tree with the minimum entropy.
C5.0 can help teachers to generate a decision tree from the observed data. The procedure used by C5.0 in generating a decision model comprises several serial steps. First, the feature attributes and the target attribute must be selected. Second, teachers must determine how to measure each attribute. Third, the teachers collect observed data composed of a set of instances, described by a fixed set of attributes and their values. Finally, the observed data are input into C5.0, and a decision model is generated. The detailed steps follow.

Determine attributes and measure the level of the attributes

The first step in generating a decision model is to determine the feature attributes and a target attribute. The feature attributes adopted in the decision tree analysis methodology should be associated with the target attribute. The target attribute in this case is whether a programming exercise is suitably difficult for a student. Specific attributes are applied to evaluate the difficulty level of a problem in some studies. For example, Lee and Heyworth (2000) evaluated the complexity of a logarithm problem to a student according to the number of steps needed by the computer system to solve the problem, the number of operators (+, −, *, /, etc.) in the problem expression and the degree of familiarity of the problem to the student; Beck et al. (1997) used the proficiency of sub-concepts in a problem to determine its difficulty level. Similarly, to solve a programming problem, a student must be familiar with the related programming concepts and possess an appropriate level of ability in algorithm analysis. Thus, a programming exercise that is suitably difficult for a student may be of a higher algorithm or programming concept complexity level than one that the student can complete without any assistance, but it can be completed with assistance. Therefore, this work selected five attributes: the student’s proficiency level in main concepts used for completing the exercise, the student’s proficiency level in sub-concepts used for completing the exercise, the complexity level in the algorithm of the exercise, the number of lines in program code used for completing the exercise, and the student’s proficiency level in algorithm analysis, as feature attributes. In addition, two other attributes, the coursework grade of the student’s last assignment and login times in the web-based learning system during the last two weeks, were also selected as feature attributes. These two attributes were chosen because the prior coursework performance and the learning motivation (We assume that students performing more login times may associate with that they have a higher learning motivation in learning) may be both related to the target attribute. For example, students having greater learning motivation are willing to spend more time and efforts to solve a more difficult problem. Additionally, students with high prior coursework grade may understand more strategies and knowledge in solving a difficult programming problem. Therefore, this work used seven feature attributes to generate a decision model.

Nine experts (TAs who had learned Java programming for two years and worked as TA for over one year) identified concepts for each exercise. In an exercise, the main concepts are those concepts that are central to the exercise, and sub-concepts are those concepts that have a secondary relationship to it. Item-Objective Consistency (IOC) (Rovinelli & Hambleton, 1977) was used to identify the main concepts and sub-concepts. Figure 5 presents an example that demonstrates how IOC is applied. The X-axis represents the nine experts; the Y-axis represents all concepts that the teacher has instructed in the last two weeks, and the Z-axis represents the ten candidate exercises. Each square contains one of three values, +1, 0, and −1. The value +1 indicates that the expert explicitly confirms that the concept is required by the exercise. The value 0 indicates that the expert non-explicitly confirm the concept demanded by the exercise. The value −1 indicates that the expert can explicitly confirm that the concept is not required by the exercise. The nine experts calculated the IOC for each concept for each exercise after filling in the squares. A concept is a candidate for main concepts when IOC ≥ 0.8 or a candidate for sub-concepts when 0.5 ≤ IOC < 0.8. Finally, a panel discussion was conducted by the nine experts to reach a consensus on what were the main concepts and what were the sub-concepts of each exercise. If one of the experts does not agree that a concept is a main concept in an exercise, he/she explains why he/she does not agree and suggests removing it from the candidates for main concepts. If the other experts find the explanation acceptable, they will either remove the concept from the candidates for main concepts, reclassify it as a sub-concept, or discard it entirely.

After identifying the main concepts and sub-concepts for each exercise, a database trigger was then executed to generate the values of two attributes, the student’s proficiency level in main concepts used for completing the exercise and the student’s proficiency level in sub-concepts used for completing the exercise, as one of three values (1 as high, 2 as middle, or 3 as low) for each student for each exercise and to generate the values of three other attributes, the student’s proficiency level in algorithm analysis, the student’s coursework grade in the last assignment, and login times in the web-based learning system in the last two weeks, as one of three values (1 as high,
As middle, or 3 as low) for each student according to the student model. The original data type of the five attributes is continuous. Because the decision tree analysis method must sort the data at each step, it may require a large amount of time at sorting the continuous data. In addition, the generated rules may be less meaningful and difficult to interpret. Therefore, we reduce the number of values for each of the five continuous attributes by dividing the range of the values into three intervals based on the method of natural partitioning (Han & Kamber, 2000). The nine experts also measured two attributes, the complexity level in the algorithm of the exercise and the number of lines in program code used for completing the exercise, as one of three values (1 as high, 2 as middle, or 3 as low) for each exercise. The experts measured the value of the complexity level in the algorithm of a exercise according to the criteria: (1) I can design the algorithm for this exercise in one minute; (2) I can design the algorithm for this exercise in five minutes; (3) I have to spend more than five minutes to design the algorithm for this exercise. In addition, the number of lines in program code in each exercise is estimated based on the experts’ experience.

Figure 5. An example IOC

Besides these feature attributes, the value of the target attribute was determined from the answer to a question. Students were required to answer the following question after submitting an assigned programming exercise: “How do you complete the exercise?” The student could choose the following answers: (1) I completed this exercise without any assistance; (2) I completed the exercise with the assistance of related learning materials; (3) I completed the exercise through collaboration with peers; (4) The exercise was far beyond my ability. If the student chose answers 1 or 4, then the system assigned a value of 0 as the target attribute, indicating that the exercise is inappropriate for this student. If the student selected answers 2 or 3, then the system assigned a value of 1 as the target attribute, indicating that the exercise was appropriate for that student.

Decision model construction

C5.0 can help teachers to generate a decision tree from the observed data. Observed data from coursework assignments 1–4 were collected to generate a decision model. The data consisted of 230 records. Each record comprised values of eight feature attributes and one target attribute. These records were input into Quinlan’s C5.0 software to generate a decision model. In total, 173 (75%) of the records were used for training, and 57 (25%) were used for testing. The training error rate was 13.3%, and the testing error rate was 21.1%, indicating a 79% probability
that a selected programming exercise was at the appropriate level of difficulty when using the generated decision rules. The decision rules were then programmed. According to the programmed rules, the PADS selected three programming exercises at the appropriate level of difficulty for each student.

**E-dictionary system**

In addition to creating plans, solving problems and making decisions, students learning programming need to understand programming concepts (Jakovljevic, 2003). When computer science students apply their programming knowledge to solve a problem or design a program, programming concepts must be well organized and integrated into their knowledge structures (Murphy et al., 2005). Therefore, the first step in attempting to solve programming problems is understanding programming concepts and organizing these concepts into a knowledge structure.

The World Wide Web (WWW) has become a vast resource for students to acquire knowledge, solve problems, and complete tasks that use web information. Moreover, students can share their learning products and have after-class discussions with peers on the web. The shared learning products present knowledge in a variety of forms. Students can learn programming concepts from different perspectives through these learning products. For example, a discussion thread denotes a process through which students identify and elaborate on ideas and thoughts, debate or interpret their own statements, and provide feedback to other students (Pena-Shaffa & Nicholls, 2004). Similarly, a programming assignment may require a student to provide a narrative description of the required process to complete a programming exercise, the program code of the programming exercise, and explanations on the program code. These learning products are useful for learning, but are distributed among different web pages. Therefore, students have difficulty finding helpful and reliable information (Hill & Hannafin, 2001) and are easily disoriented during the information-seeking process (Marchionini, 1995).

The e-dictionary system organizes useful learning products with a concept map. Each learning product is linked with related programming concepts. Therefore, students just need to find a concept that they want to learn in the e-dictionary system, and then the learning products related to this concept are shown.

In addition, the learning products included in the e-dictionary system were gathered and evaluated by a peer assessment process. The peer assessment process is modelled on the actual journal publication process of an academic society and was conducted last year in the same course (Chang, Chen, & Li, 2008). Every student is endowed with a role (reader, author, reviewer, and editor) according his/her learning performance. Students can submit on any of the following six topics: answers to the assignment, what he/she has learned, summary of discussion, problems, useful websites, and miscellaneousness. Then each submission is evaluated by two reviewers. Finally, an editor summarizes the opinions of all reviewers and determines whether or not the submission will be accepted. If a submission is accepted by editors, it can be included in the e-dictionary system.

The e-dictionary system is designed in two parts: a conceptual structure that denotes how learning products link with a concept map and a user interface that demonstrates how students interact with the e-dictionary system. The conceptual structure is represented as a network structure in which learning products are linked with a concept map. Nodes in the structure are of two types, concept and material. A concept node represents a programming concept in the course, and a material node represents a learning product in the web-based learning system. Each material node can be linked to one or more concept nodes. To link a learning product with the concept map, students must identify which concepts are related to the learning product when submitting the learning product. For example, a student posting a discussion topic or submitting a completed assignment must identify the programming concepts related to the discussion topic or the completed assignment in a listbox. The database schema of the conceptual structure comprises the two tables “Concept_Map” and “Learning_Materials” (Figure 3).

The interface to the e-dictionary system is a collection of web pages accessed by students to learn programming concepts. Figure 6 presents the interface of the e-dictionary system. This interface is divided into four sections. Section 1 is a full-text search engine in which students can search keywords. After a student enters a keyword, Section 2 shows a list of related keywords. If the search engine cannot find this keyword, then the list shows eleven keywords with similar spelling to that keyword. If the search engine can find the keyword, then this keyword is arranged in the middle of this list; and related keywords are arranged in the periphery (see section 2 of Figure 6).
Section 3 has five hyperlinks with this keyword. Table 3 lists the five hyperlinks in detail. The student can click one of the five hyperlinks, and the respective learning materials are then shown in Section 4.

![Figure 6. The user interface of the e-dictionary](image)

### Table 3. The five topics and detailed information

<table>
<thead>
<tr>
<th>Topic</th>
<th>Detailed information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanation</td>
<td>The explanation of the subject keyword that is elucidated by teachers</td>
</tr>
<tr>
<td>Example</td>
<td>The completed assignments associated with the subject keyword</td>
</tr>
<tr>
<td>Discussion Thread</td>
<td>The discussion threads associated with the subject keyword</td>
</tr>
<tr>
<td>Related Concept</td>
<td>The keywords linked with the subject keyword in a concept map</td>
</tr>
<tr>
<td>Web search</td>
<td>The search results of the subject keyword from a search engine</td>
</tr>
</tbody>
</table>

**Peer recommendation system**

According to social constructivist theory, people learn in a process of knowledge construction (Reigeluth, 1999). Knowledge is socially constructed by exchanging ideas, negotiating meaning, and sharing objectives and dialogue (Pea, 1993). Accordingly, the learning environments should provide activities and opportunities for students to articulate and reflect on material, and negotiate meaning with other students (Pena-Shaffa & Nicholls, 2004).

Online asynchronous discussion forums provide students with an appropriate context for meaningful learning. In such forums, students can collaborate, share and negotiate knowledge without the need to meet physically in person, or to work simultaneously. Therefore, these forums allow instructors and students to interact after class (Rossman, 1999).

However, most students do not post or reply to questions in discussion forums. A discussion forum fails if few or no messages are posted to it (Preece, Nonnecke, & Andrews, 2004). Therefore, teachers always have to act as facilitators to encourage students to discuss and guide them to produce beneficial dialogue in discussion forums (Mazzolini & Maddison, 2003). However, to create a successful forum, teachers have to spend considerable time and effort. One method for decreasing the load on teachers is to promote learners helping each other. Studies have developed several systems for locating the help of capable peers, such as Answer Garden 2 (Ackerman & McDonald, 1996), Intranet Peer Help Desk (Greer et al., 1998), and Expert Finder (Vivacqua & Lieberman, 2000). The
The differences between our system and these systems are both that the peer recommendation system not only finds capable peers to answer posted questions, but to motivate students to participate in discussion and that the peer recommendation system incorporates peer recommendations in the context of coursework activity.

The peer recommendation system selects two capable classroom peers to resolve a proposed question together with the questioner on a discussion forum. The reason for selecting two peers for each question is that we hope each question can be replied to by at least one student. If there is a question that a recommended peer can not answer or forgets to answer, then there is still the other peer to answer. To motivate those two peers to participate in the discussion, this system contains two assumptions: the students are interested in answering a posted question that relates to a programming exercise on which they are working, and students answer questions if they are requested. Based on the assumptions and students’ learning status, two capable peers for a posted question are selected in the following order: (1) students who receive and have previously submitted the same programming exercise as the questioner; (2) students who receive the same programming exercise as the one in which the question is raised, and who are also proficient in the main concepts required to complete the exercise; (3) students who have received the same programming exercise as the one in which the question is raised, and (4) students who are proficient in the main concepts required to complete the programming exercise in which the question is raised. This system can select two capable peers according to these rules. After selecting two peers to answer a question, the system sends an e-mail to these peers asking them to answer the question. The e-mail includes a hyperlink to the discussion thread, and some statements telling these peers that their comments are welcome.

**Experiment**

An experiment was conducted to evaluate the WCSS. The experiment was performed at the National Central University in Taiwan, and the subjects were 50 first-year undergraduate students. These students were all computer science majoring, enrolled in a mandatory course entitled “Basic Computer Concepts”, in which they studied Java programming (including structured programming, object-oriented programming, and graphical user interfaces).

The students learned in a blended learning environment. That is a teacher lectured for three hours weekly in the classroom and the students also learned via the web-based learning system after class. Seven coursework assignments were set over a 14-week period. Students were allocated two weeks to complete each assignment. During the first three coursework assignments (1-3), the students learned in the web-based learning system without the WCSS support. And each assignment consisted of three programming exercises, ranked as low, middle and high levels of difficulty, which were assigned to all students. The teacher determines the difficulty level of an exercise according to three rules. (1) an exercise that requires students to apply a programming concept with a complexity level of algorithm analysis that the teacher can design within one minute is ranked as low level of difficulty; (2) an exercise that requires students to synthesize two programming concepts with a complexity level of algorithm analysis that the teacher can design within five minute or an exercise that requires students to apply a programming concept with a complexity level of algorithm analysis that the teacher has to spend more than one minute on to design is ranked as middle level of difficulty; and (3) an exercise that requires students to synthesize more than two programming concepts or an exercise that requires students to synthesize more than one programming concept with a complexity level of algorithm analysis that the teacher has to spend more than five minutes on to design. During assignment 4, the students were able to use the e-dictionary and the peer recommendation system to solve their exercise problems. That means the students learned in the web-based learning system with the e-dictionary and the peer recommendation system but without PADS. During the last three coursework assignments (5–7), the students learned in the web-based learning system with the WCSS support. And three programming exercises selected by PADS were assigned for each student in each coursework assignment.

The proposed WCSS was evaluated using the following quantitative and qualitative data: (1) a set of web logs, which recorded all student interactions with the web-based learning system; (2) a questionnaire, which students completed at the course end; (3) notes from semi-structured interviews. These data were collected to answer the following questions:

- Whether the WCSS can increase likelihood that students do coursework within their ZPD,
- Whether the support provided by WCSS can help students complete their programming exercises,
- Whether the e-dictionary system is an effective tool in helping students complete programming exercises, and
- Whether the peers selected by the peer recommendation system were willing and able to answer the assigned questions.

**Results**

**Experimental results about the WCSS**

After submitting an assigned programming exercise, the students were required to answer a question, “How do you complete the exercise?” Figure 7 presents the percentages of each answer in each assignment. The sum of the percentages of answers 2 and 3 in each assignment from assignments 1–3 (without WCSS) were obviously lower than the sum of the percentages of answers 2 and 3 from assignments 5–7 (with WCSS). This result may reveal that students with the support of the WCSS have higher likelihood of doing their coursework within their ZPD than without the support of WCSS.

![Figure 7. Percentage of each answer of the question "How do you complete the programming exercise"](image)

In order to investigate whether the result was statistically significant, a paired $t$-test was conducted. Because in the assignment 4, the students were able to use the e-dictionary and the peer recommendation system, but the three exercises were not selected by PADS. Therefore, the paired $t$-test compares the average number of answers 2 and 3 in an assignment during assignment 1-3 (without WCSS) with the average number of answers 2 and 3 in an assignment during assignment 5-7 (with WCSS). As shown in Table 4, the result was significant ($t$=-11.366, $p=0.00<0.05$), revealing that there were more exercises (means=1.81) that were completed with assistance in an assignment when using the web-based learning system with WCSS than without WCSS (mean=0.81). This result indicates that students with the support of the WCSS have higher likelihood of doing their coursework within their ZPD than without the support of WCSS.

To investigate whether WCSS can support the students in completing their exercises, a paired $t$-test was conducted by comparing the average number of answer 4 in an assignment during assignments 1-3 (without WCSS) with the average number of answer 4 in an assignment during assignment 5-7 (with WCSS). As shown in Table 4, the result was significant ($t$=9.376, $p=0.00<0.05$), revealing that there were fewer exercises (means=0.55) that the students could not complete before the deadline of an assignment when using the web-based learning system with WCSS than without WCSS (mean=1.59). In other words, there were more exercises that the students could complete before the deadline of an assignment when using the web-based learning system with WCSS than without WCSS.

There might be two reasons to explain why more exercises could be completed with the support of WCSS. One is that the e-dictionary and peer recommendation systems can provide effective support for students, so students can complete more difficult exercises. The other is that the PADS assigned easier exercises than the conventional method where the teacher assigns three programming exercises to all students. Because the average number of answer
1 (students can complete without any assistance) in an assignment during assignment 5-7 (mean=0.64) is approximately the same with the average number of answer 1 in an assignment during assignments 1-3 (mean=0.60), so the PADS in some degree did not assign easier exercises than the conventional method. Therefore, we tend to believe that the e-dictionary and peer recommendation systems could support the students in completing their exercises.

Table 4. Comparing students’ responses when doing coursework without WCSS and with WCSS

<table>
<thead>
<tr>
<th></th>
<th>Without WCSS</th>
<th>With WCSS</th>
<th>t-statistic</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer1 (completed without any assistance)</td>
<td>0.60 0.31</td>
<td>0.64 0.48</td>
<td>-0.447</td>
<td>49</td>
</tr>
<tr>
<td>Answer2 and 3 (completed with assistance)</td>
<td>0.81 0.49</td>
<td>1.81 0.39</td>
<td>-11.366**</td>
<td>49</td>
</tr>
<tr>
<td>Answer4 (far beyond my ability)</td>
<td>1.59 0.70</td>
<td>0.55 0.44</td>
<td>9.376 **</td>
<td>49</td>
</tr>
</tbody>
</table>

*p-value < .05  ** p-value < .01

Experimental results about the e-dictionary system

The e-dictionary system was evaluated according to students’ perceptions. In general, an information system is more highly used and accepted when users have a more positive attitude toward it. Thus, a questionnaire, used in Davis’s technology acceptance model, was chosen as an evaluation tool to measure the effects (Davis, 1989). The questionnaire using 7-point Likert scales (from “1” which means “strongly disagree” to “7,” “strongly agree”) was devised with twelve items to investigate whether the students perceived that the e-dictionary system was useful and easy to use for completing their programming exercises. The Cronbach alpha reliability for perceived usefulness is 0.94, and for perceived ease of use is 0.94. The total Cronbach alpha reliability is 0.97; this implies the questionnaire data have a high reliability. Analytical results indicated that students’ attitude of perceived usefulness (mean=5.10, SD=1.35) and perceived ease of use (mean=5.20, SD=1.13) were all high, and thus the students perceived that the e-dictionary system was useful and easy to use for completing their programming exercises.

Experimental results about peer recommendation system

Table 5 lists the statistics of the students’ discussions in the discussion forum for each coursework assignment. The first column, “Question,” represents the number of questions posted on the discussion forum. The second column, “Reply,” represents the number of replies for each posted question. The third column, “Question solved,” represents the number of posted questions that were solved as determined by the teacher.

The peer recommendation system was able to be used after coursework assignment three was completed. This system selected two students for each posted question. Accordingly, the reply rate (Reply /Question) increased once the peer recommendation system was employed. The average reply rate from assignments 1–3 (without the peer recommendation system) was 0.43, which was lower than the average reply rate from assignments 4–7 (with the peer recommendation system), which was 1.14. The analysis may indicate that the peer recommendation system can motivate students to participate in discussion. The number of questions posted and answered declined in assignment 7, since the final exam was taken during this time, causing the students to spend less time on the web-based learning system.

Additionally, the questions-solved rate (Questions solved / Question) increased after assignment three. The average questions-solved rate (0.82) in assignments 4–7 (with the peer recommendation system) was higher than the average questions-solved rate (0.21) from assignments 1–3 (without the peer recommendation system). The peers selected by the peer recommendation system successfully solved 82% of the posted questions. These results may represent that most of the selected peers are willing to answer the assigned questions and are able to help questioners solve the questions. However, there were still 18% of the posted questions that can not be solved by the selected peers and the reply rate for each coursework session during assignment 4-7 was not anticipated. Theoretically, if the peer recommendation system requested two students to answer a question, then each question would be answered by at least two students. The reasons for this unanticipated reply rate and unsolved question rate may be: (1) some questions were too difficult for the selected students to answer, and (2) one student answered the question and, consequently, the second student did not feel obliged to answer the question.
Table 5. The statistics of the students’ discussions related to coursework assignments

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Question</th>
<th>Reply</th>
<th>Reply /Question</th>
<th>Question solved</th>
<th>Question solved / Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignment1</td>
<td>4</td>
<td>1</td>
<td>0.25</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>Assignment2</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Assignment3</td>
<td>10</td>
<td>5</td>
<td>0.50</td>
<td>2</td>
<td>0.40</td>
</tr>
<tr>
<td>Assignment4</td>
<td>27</td>
<td>25</td>
<td>0.93</td>
<td>20</td>
<td>0.74</td>
</tr>
<tr>
<td>Assignment5</td>
<td>27</td>
<td>39</td>
<td>1.44</td>
<td>25</td>
<td>0.93</td>
</tr>
<tr>
<td>Assignment6</td>
<td>17</td>
<td>19</td>
<td>1.12</td>
<td>15</td>
<td>0.88</td>
</tr>
<tr>
<td>Assignment7</td>
<td>8</td>
<td>7</td>
<td>0.88</td>
<td>5</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Comments from interview with the students

Three students, SA, SB, and SC, were chosen randomly from three grade-based groups for in-depth interviews. The teacher interviewed with the three students after the course end. Some of the students’ comments that are useful for improving WCSS are summarized as follows:

- The suggested concepts listed on an exercise web page can hint which concepts I can use in completing the exercise. If I don’t know how to complete it, the hint can help me. But sometimes it may limit my creativity in using other concepts to complete the exercise. (responded by SB)
- Since each student received different programming exercises, it was not easy to find someone who received the same exercises as me. It forced me to complete the assigned exercises by myself. (responded by SC)
- Most of the assigned exercises are very difficult. I needed to spend considerable time completing each assigned exercise. I feel that it is not fair to give me difficult exercises. (responded by SA)
- The e-dictionary system was useful in learning programming concepts since it provided related learning materials, such as explanations, discussion threads, and programming examples. However, it did not help me analyze the algorithm of the assigned programming exercises. (responded by SB and SC)
- If I did not know which concepts I needed to complete a programming exercise, then the e-dictionary system was not useful in helping me complete the programming exercises. (responded by SB)
- The learning materials in the e-dictionary system were too simple and did not cover enough materials. It can not advance my skills. I prefer finding learning materials on the Internet for solving problems and learning. (responded by SA)

Discussions

The designed WCSS system, including PADS, e-dictionary, and peer recommendation system, provide three ways of assistance to support students in performing coursework within their ZPD. A result is statistically significant revealing that students with the support of the WCSS in a coursework activity have higher likelihood of doing their coursework within their ZPD than without the support of WCSS. In addition, the number of exercises that can be completed before deadline by students with the support of WCSS is significantly more than those done by students without the support of WCSS. It partly prove that the WCSS can provide effective support for students to complete their exercises. According to these results, we believe that WCSS can improve learning in the coursework activity by providing challenge and assistance.

The exercise web page that presents hyperlinks to the e-dictionary and peer recommendation system reminds students which ways of assistance can be used for completing their exercises. Moreover, the programming concepts listed in the concept section suggest which concepts can be used for completing the exercise. The WCSS providing explicit guidance and task-related assistance in the exercise web page can direct students to complete their exercises. It also reduce students’ cognitive load in help seeking activities (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003). However, the explicit guidance may be ineffective where self-regulation is needed (Hill & Hannafin, 2001) and sometimes may hinder students from using different programming concepts for solving problems (SB responded this). The explicit guidance has both positive and negative effects for learning. Therefore instructional designers should consider this factor to provide learners appropriate support.
The PADS sometimes assigned very difficult exercises for advanced students (SA responded this). These exercises required more time to complete than the simple exercises. Hence, the advanced students needed to spend more time completing the assigned exercises than other students. They thought it was not fair. Although we did not focus on the assessment mechanism in this study, it remains an important issue to have a fair assessment mechanism when each student is assigned with the exercises of different levels of difficulty. This issue is also found in the researches related to CAT (Lilley, Barker, & Britton, 2004). In addition, to reduce the advanced students’ workload, the rules for assigning programming exercises may use the decision model and also consider students’ workload. Thus, if a student received very difficult exercises, then PADS can automatically reduce the number of assigned exercises from 3 to 2 or 1.

E-dictionary integrates distributed learning materials in a concept map structure. Hence, students can easily find the learning materials related to a programming concept and learn the concept from its explanation and related examples. The experimental results showed that the students highly accepted and used the e-dictionary. It may represent that the e-dictionary system is an effective tool in helping students complete programming exercises. However, some students complained that the e-dictionary system contained too little material and that some of it was too vague. Most of the learning materials within the e-dictionary were gathered from the web-based learning system and were categorized into explanation, example, and discussion. Because of the limitation of sources and categories of the learning materials, the e-dictionary may be more useful in helping students perform low-level cognitive processes, such as remembering, understanding, and applying programming concepts, than in helping students perform high-level cognitive processes, such as analyzing, evaluating, and creating (Anderson & Krathwohl, 2001). Future work can redesign the e-dictionary system so that students can collectively insert and edit the content in the e-dictionary system by themselves. The editing process is similar to that used in a wiki system, such as wikipedia (http://www.wikipedia.org/). Moreover, the e-dictionary also can provide personal note-taking tools that can associate collective knowledge representation with personal knowledge representation (Chen & Hung, 2002).

The reply and questions-solved rates were increased after the students were able to use the peer recommendation system. These results may represent that most of the selected peers are willing to answer the assigned questions and are able to help questioners solve the questions. A student who is notified to answer a question may recognize that he/she has a responsibility to answer it. Thus, the reply and questions-solved rate were increased. The students who just passively receive answers will not learn better than the students who actively participate in discussion. Therefore, peer recommendation can be thought of as a strategy for motivating students to participate in discussion. However, the number of questions posted was very few. The reason may be that we did not encourage students to post their questions in the discussion board in the beginning of the course (Preece, Nonnecke, & Andrews, 2004).

Limitations

In the proposed WCSS, the teacher and TAs were required to perform certain tasks such as identifying the concepts needed in a programming exercise, measuring the complexity level of algorithm in each programming exercise, constructing a concept map for this course, and extracting concept explanations from textbooks. These tasks require a significant amount of time. Therefore, nine TAs were employed to help the course teacher complete the tasks. Although these tasks required a significant time commitment, most were performed once, including constructing a concept map and extracting concept explanations from the textbook, and some tasks were completed by students, such as identifying the concepts used in a submitted assignment. Only two tasks, entering the examination grade and the coursework grade into the database, were repeatedly performed by the teachers and TAs. Therefore, the workload of teachers and TAs was only minimally increased.

The nine experts sometimes had some difficulty in consistently identifying the concepts required for completing an exercise. That is students can adopt different programming concepts to complete an exercise. For example, when solving the programming problem “Sum the integers between 1 and 100”, a student may use “for”, “while” or “recursive method” to control the loop. Thus, some experts may choose “for” and others choose “recursive method” as the main concept. The learning objective of the course was to train students in understanding and applying the basic programming concepts, but not in learning complex programming skills, such as algorithm analysis or data structure. Hence, the concepts used in each exercise must be clearly identified. For example, the above problem can be stated as “Sum the integers between 1 and 99, using recursive method.”
Conclusion

To increase the likelihood that students do their coursework within their ZPD, this paper has presented a Web Coursework Support System based on the notion that teaching and learning occurs only when assistance is provided to learners in their ZPD. Therefore, to assist students in performing coursework within their ZPD, WCSS provides three subsystems: PADS, which applies decision tree analysis methodology to select programming exercises with a level of difficulty suited to each student; an e-dictionary system that allow students to access distributed learning products related to a programming concept from a single web page; and a peer recommendation system that selects two capable classroom peers and motivates them to answer assigned questions. Experimental results show that the exercises-completed rate and the rate of the assigned exercises completed with assistance were increased. The findings indicate that WCSS can increase the likelihood that students do coursework within their ZPD by providing challenge and assistance. In addition, some suggestions for improving the coursework activity with WCSS support are proposed. Those suggestions consist of (1) considering the positive and negative effects of explicit guidance for providing learners appropriate coursework support, (2) considering the fairness of assessment when each student is assigned with the exercises of different levels of difficulty, (3) providing support for students to collectively construct knowledge in the e-dictionary system, (4) providing personal note-tasking tools for associating learning materials within the e-dictionary system with personal knowledge structure, and (5) conducting peer recommendation in the beginning of a course to promote interaction in the discussion forum.

References


Intra-action, Interaction and Outeraction in blended learning environments

Wu-Yuin Hwang¹, Jung-Lung Hsu², Alexei Tretiakov³, Huey-Wen Chou⁴ and Ching-Yuan Lee¹

¹Graduate School of Network Learning Technology, NCU, Taiwan // ²Department of Information Management, Kainan University, Taiwan // ³Department of Information System, Massey University, New Zealand // ⁴Department of Information Management, NCU, Taiwan // wyhwang@cc.ncu.edu.tw // srl@mail.knu.edu.tw // A.Tretiakov@massey.ac.nz // hwchou@mgt.ncu.edu.tw // smallboss@lst.ncu.edu.tw

ABSTRACT

Theory of distributed cognition unveils the answer to what constitute distributed cognition. However, how the distributed cognition in web-based learning environment spreads out still remains a black box. This study sought to deepen our understanding on how learners’ knowledge disseminates online by examining the impacts of three kinds of communication, namely intra-action, interaction, and outeraction. This paper conducted a quasi-experimental study and invited 135 Taiwanese college students to help us explore our attempted research question. The results clearly demonstrated that intra-action has stronger relationship with learning outcomes than interaction or outeraction. In addition, considerable correlation was detected between intra-action and interaction. Subsequent interviews with learners indicated that considerable part of the positive relationship between intra-action and learning achievement may be due to cause-effect, namely, intra-action activities are perceived by learners as contributing to positive learning outcomes. In sum, this study contributes to the literature because (a) we extend previous research on distributed cognition by examining the impacts of three kinds of communication; (b) we suggest it is worthwhile for researchers to further study the impact of intra-action; (c) by distinguishing the three communications, we introduce one way that can measure status of the constituent elements of distributed cognition that exists in a distance learning environment.

Keywords

Communication, Intra-Action, Interaction, Outeraction, Annotation, Instant messenger, Distributed cognition

Introduction

Undoubtedly, people communicate to understand and to cooperate with each other in order to accomplish goals together. Similarly, individuals learn and solve problems through various communications by exchanging knowledge and by combining different perspectives at a knowledge or problem domain (Garrison & Shale, 1990; Laurillard, 1997; Lipman, 1991; Nardi, Whittaker, & Bradner; 2000; Wagner, 1994; Wenger, 2001). As a result, many new real-time applications that incorporate novel features seem to have potential to benefit learning performance. For example, Instant Messenger (from AOL) and Windows Messenger (from Microsoft) provide learners instant feedback and more immediate emotional connection (Nardi et al., 2000; Ron, 2003), which believe to have complemented asynchronous communication. In this regard, fostering new construction of knowledge for learners are more like the way of distributed cognition, particularly when web-based learning setting integrates both ways of communication.

Theory of distributed cognition (TDC) refers to process with two properties (Hutchins, 1995; McClelland, Rumelhart, & the POP Research Group, 1986). First, they are cognitive, i.e. they involve forming certain representations of the world. Second, they are not performed by a single person, but are distributed across multiple individuals. Considering the former, McClelland and his associates (1986) believe that people do the cognitive processing required by creating and manipulating external representations. Therefore, this process involves an external representation consisting of written symbols. As to the latter property, Hutchins (1995) considers that no one could physically do all the things that must be done to fulfill the cognitive task. Accordingly, we are equipped with knowledge derived from others’ external representation.

Accepting the idea that cognition may be distributed throughout a system comprising both individuals and artifacts, this study believes that learners in web-base settings come to know the knowledge by being part of the cognitive system. However, the work by Hutchins and McClelland and his associates only help us understand what constitute distributed cognition. The work how these components function still remains a black box (Bell & Winn, 2000). This study therefore sought to deepen our understanding of how the distributed cognition in web-based learning environment spreads out. Seeing that research should reflect the ways that cognition is socially enabled and
distributed through communication (Hutchins, 1995), we first referred to Nardi and his colleagues’ (2000) definitions of two distinct kinds of communication: outeraction and interaction. Outeraction, as they define, is “a set of communicative processes outside of information exchange, in which people reach out to others patently in social ways to enable information exchange.” Alternatively, interaction is “the actual exchange of information directly relevant to knowledge sharing or problem solving.”

Two reasons tempt us to borrow their ideas to shape our research framework. First, in order to effectively collaborate and function as a virtual learning community, learners need to be aware of other peers, where they are located (demographic awareness), what others know (knowledge awareness) and what they are able to do (capability awareness) (Daniel, Zapata-Rivera, & McCalla, 2003; Gutwin & Greenberg, 1998). In light of the viewpoint, outeraction serves as a mechanism that creates connections to others, and consequently providing learners with the needed awareness, which is consistent to distributed cognition assumption that cognition is distributed across multiple individuals. Second, knowledge is mutually constructed through continuous interactions in which learners interpret information and knowledge that they try to share and exchange to one another. Thus, interaction inevitably leaves certain external representations, which is another requisite of distributed cognition.

Through the lens of TDC, much research considers a process is noncognitive simply because it happens in a brain (Fazio & Olson, 2003; Semin, 2007; Semin, de Montes, & Valencia, 2003). Therefore researchers drew their attention on the exchange of external representation across multiple individuals. However, according to Social Development Theory (SDT), Vygotsky (1978) states that “learning occurs in social or interpersonal context prior to its becoming internalized or individualized within an intra-psychological category.” While learning in interpersonal context occurs via communication between different individuals (interaction assisted by outeraction), internalization of learning may also be assisted by communication, but in this case the discourse would be directed not at others, but at the learner himself/herself. People benefit from this kind of communication-with-self by taking different perspectives at a problem at different points in time. Accordingly, this study considers that SDT is potential to make TDC more complete by highlighting the importance of internal communication. Based on the terminology used in Vygotsky’s theory, we call this kind of internal communication “intra-action.” Comparing with interaction focusing on the exchange of external representation across multiple learners, intra-action focuses on a single student and his own representation of knowledge. We regard this kind of representation as a single student’s inner representation to discern from external representation that is for distributing across multiple learners. In this study, we argue that intra-action is as important as outeraction and interaction in promoting positive learning outcomes in blended learning environments, and present experimental data in support of our point of view. In fact, we believe it is the intra-action that plays an intermediate role between interaction and outeraction for distributed cognition to spread out.

We note that although some blended learning environments do provide tools supporting communication with oneself (e.g., WebCT, one of the most popular commercial blended learning systems, provides bookmarking and note-taking tools, while Moodle provides a journal tool that can be used for note-taking), in actual pedagogical practice interaction and outeraction tools such as discussion forums or chat rooms receive much greater emphasis. Moreover, before working out how the distributed cognition spreads out, we need to appreciate the relationships among the three types of communications. Hence, it is highly important to assess the impact of intra-action on learning outcomes. If, indeed, intra-action activities have a strong positive impact, blended learning could increase return on their investment in blended learning infrastructure just by shifting emphasis from encouraging interaction only to encouraging both interaction and intra-action.

**Literature review**

**The significance of communication in learning**

In early 1980s, peer-to-peer communication was not emphasized as part of distance education due to the limited availability of communication technology (Holmberg, 1989). When Internet became widely available in mid 1990s, it led to an eventual change of emphasis as social constructivist views already firmly established in traditional education became relevant to distance education due to enhanced technical capabilities enabling communication.

At present, Web-based communication facilities are widely used to enhance learning in blended learning environments. Communication is recognized to serve a variety of functions in the learning process. Sims (1999) has
listed these functions as allowing learners to control their learning pace, facilitating adaptive learning based on learner traits, and acting as an aid to meaningful learning. Lipman (1991) and Wenger (2001) asserted that peer-to-peer communication is fundamental for fostering learning. Valuable ideas can be obtained, elaborated and constructed through communication, which is a key component in constructivist theory (Jonassen, 1991). Garrison and Shale (1990) asserted that all forms of learning can be reduced to the process of communication among students and teachers. Laurillard (1997) introduced a conversational learning model in which communication plays a central role. In his work on collaborative learning, (Slavin, 1995) demonstrated that peer-to-peer communication leads to better performance in cognitive learning tasks, as well as increases completion rates and facilitates acquisition of critical social skills. Damon’s (1984) study illustrated the communication benefits to both instructors and learners resulting from a variety of forms of reciprocal teaching. (Wenger, McDermott, & Snyder, 2002) found peer-to-peer communication to be critical to the development of learning communities: it enables learners to develop interpersonal skills and to access tacit knowledge shared by community members as well as knowledge prescribed by the formal curriculum.

**Interaction**

Interaction is a kind of communication involving more than one person communicating. Parker (1999) was the first to study “interaction” as an activity in blended learning, which he defined as a reciprocal communication among senders and receivers under a specific topic. At about the same time, Berge (1999) defined the notion of interaction in blended learning context as follows: “Interaction is the two-way communication among two or more people with the purposes of problem solving, teaching or social relationship building.”

Generally speaking, interaction may have a variety of purposes, such as information sharing, problem solving, or even social exchange. However, Nardi and his colleagues (2000) suggested that the scope of the notion of interaction should be narrowed down to distinguish it from the notion of communication in general. They proposed that the term interaction should only apply when communication between individuals is directly relevant to problem solving or information exchange. Henceforth in this paper we employ the term “interaction” in this narrow, more specific meaning.

Both traditional, instructivist view of learning and the now widely accepted social constructivist view emphasize interaction (between a teacher and a student, and between students and members of a community of practice, respectively). In this study, we do not draw a distinction between traditional and constructivist views, but just accept that interaction promotes learning.

**Outeraction**

According to Nardi and his colleagues’ (2000) definition, outeraction is a set of communicative processes outside information exchange that enable people to connect to others in patently social ways to enable information exchange. Based on their study of the use of instant messaging by office workers they listed the following types in outeraction exchanges: (a) negotiating conversational availability; (b) preambles; (c) communication zone in an intermittent conversation; (d) awareness moments; (e) managing conversational progress. In this study, we use the notion of outeraction as it was defined by Nardi et al. (2000), although we apply it to blended learning. In our study, conducted over a shorter period of time and in a different context (blended learning), we were able to clearly distinguish only two categories of outeraction: preambles and awareness moments.

The reason we are concerned with outeraction in this study is that outeraction can be viewed as a prerequisite for successful interaction, which is also the precondition of TDC, multiple individuals. Although outeraction does not promote the knowledge of the subject matter in the target domain directly, it facilitates the relevant interaction, which, in its turn, results in learning. The benefit of adopting a narrow definition of interaction and considering outeraction as a distinct phenomenon is a more refined and complete account of the role of communication in learning.
Intra-action

Although interaction and outeraction can be used to explain most of the communicative intentions, there are certain types of discourse playing an important role in learning that can not be described as either interaction or outeraction. For example, note-taking, annotation, diary-keeping and bookmarking all involve repeated passing of meaningful messages, both informative and emotional, and hence could be regarded as communication. In particular, when a person adds annotation or digital bookmark to a page of multimedia content, it can be regarded as sending a message to himself. When, at a later time, the same person encounters the annotation or bookmark he created earlier, it can be regarded as receiving a message. This kind of communication differs from interaction or outeraction in a number of respects: it is from a person to himself/herself and time separation is essential (as space separation is not applicable). As the individual receiving the message is (due to time separation) in a different state from the moment when the message was issued, the message is likely to contain something “new” and hence, valuable to the receiver, something that is not in his/her immediate field of attention, and yet, pertinent to his/her overall goals. We call such communication-with-self in context of a blended learning environment “intra-action”.

Vygotsky (1978) introduced the concept of “intra” in the psychological studies and stated: “Every function in the child’s cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (interpsychology) and then inside the child (intrapsychology).” Intrapsychology refers to the change of existing internalized knowledge or internalization of a new knowledge.

The difference between intrapsychology and intra-action is that intrapsychology is a mental process while intra-action is one of its external representations. On one hand, intra-action is an output of intrapsychology, so one could investigate the constant change inherent to intrapsychology by studying intra-action. On the other hand, intra-action is not just a mere reflection of intrapsychology: both externalization of knowledge as a message, and the subsequent interpretation of a message-from-oneself may lead to evolution and refinement of internal mental models as contradictions and omissions are discovered and resolved, leading to a more complete state of knowledge. Intra-action effectively expands the state space of the intrapsychology mental process by allowing knowledge captured on media used for intra-action to effectively extend the internal mental state. Hence, better understanding attained via intra-action can be viewed as the outcome of searching for an optimum over a wider set of possibilities.

Some evidence on the relevance and importance of intra-action for learning is available in the literature (Howe, 1997; Brown & Smiley, 1978; Hwang, Wang, & Sharples, 2007). For example, Howe (1997) demonstrated that writing notes leads to better learning outcomes, especially when notes are written in the learner’s own words. This result is consistent with the findings by Brown and Smiley (1978), which showed that the more underlining or notes are taken the higher are the learning achievements. Marshall (1997) proposed a division of annotations into two types: inexplicit and explicit. Explicit annotations (such as text) convey more meaning than inexplicit ones (such as highlighting or drawing graphical symbols). In a web-based annotation system, Hwang and his associates (2007) studied the effects of explicit and inexplicit annotations on learning performance. Their study found that explicit annotations have much more impact on learning than inexplicit ones. Moreover, because annotation sharing can be easily applied in the web-based environment, sharing annotation scenarios were employed, thus, combining intra-action and interaction. The results indicated that the use of shared annotation can improve students’ learning achievements more significantly compared to individual ones.

Bransford and his associates (2000) asserted that teachers need to use formative assessment to develop students’ ability of self-assessment, to enable students to assess their own situation. They demonstrated that self-assessment correlates positively with learning achievements. We observe that self-assessment, as discourse with oneself, is related to intra-action.

Interplay between the three kinds of communication

Outeraction, interaction and intra-action do not appear independently. They are all related to each other, and one kind of communication may be opening way to another. For example, communication may start by a simple greeting (outeraction). Once rapport is established, one may ask a substantive question (interaction). Once the answer is clarified, one might write a note listing most important points of the answer as a message to oneself at a later moment in time (intra-action). All three contribute to the new information being internalized/learned.
In the research presented in this article, we study the relationships between interaction, outeraction and intra-action in context of a distance learning blended learning course by considering their relationships with each other and with learning outcomes. Some (not all) of the possible causes of relationships between interaction, outeraction and intra-action are illustrated in Figure 1 below. First, outeraction provide learners with awareness of what peers know and what they are able to do (Daniel, Zapata-Rivera, & McCalla, 2003; Gutwin & Greenberg, 1998). In this regard, students benefit from outeraction in that it establishes connections to their distant peers and facilitates the process in which they can quickly find out correct ones to conduct meaningful interactions by exchanging external representations. In addition to the knowledge awareness, outeraction also facilitates to negotiate one another’s available time and space before interaction takes place as well. Alternatively, quite often learners willing to share their knowledge seek to attain a balance between donating and collecting knowledge. This implies that individuals share their own knowledge because they expect others to contribute as well (Nahapiet & Ghoshal, 1998). Along with this reasoning, an attempt to interact with one another in the future motivates learners to outeract with others for future knowledge construction.

Conceptually, learners acquire new knowledge by being evolved into various perspectives from others. And a learner may not come out a notion groundlessly; instead, he should have an in-depth cognitive processing in order for him to clearly elaborate it to others. In this cognitive processing learners may come out some representation which is meaningful for himself. In this study, we consider this whole process as an intra-action. Similarly, after interactions, learners may have to summarize what they learn before the bridge from his extant knowledge repository to a comprehensive understanding fades out. In addition, this study also considered intra-action would motivate learners to outeract with the one who has the needed expertise. For instance, we may all experience the case in which we find something elusive while reading course material. Thus, we may put more effort to make things clear. Or it may motivate us to ask the competent learners to reserve available time in their convenience, and then to help us work out the things in a short time. Although there are bidirectional influences of communication types drawn in Figure 1, this study considered that there was one-way arrow having major influence in any pair of the bidirectional influences. While the one-way arrows stand for a major way of influence, it does not mean that the opposite influence of any one-way arrow does not exist. Instead, the opposite direction also may influence the target communication type. For example, in a classroom instructors spent most of the time to teach students, however, it does not mean that students cannot deliver their thinking to the instructors. In such a case, we consider instructor-to-students a major way of influence, whereas students-to-instructor a minor way of influence.
Research design

In this study, we presented a quasi-experimental design in which we explored how distributed cognition spreads out by uncovering relationships that are consistent with a view that intra-action is as important as interaction and outer-action in promoting learning. The study is conducted in an blended learning environment supporting interaction and outer-action via discussion forum and instant messaging facilities, and intra-action - via enabling students to make annotations. Using our own environment, rather than a product such as WebCT, makes it possible for us to log learner activities in sufficient detail. Snapshots of the environment were provided in Figure 2. On the right-hand side of Figure 2, the system showed material for learners to study. In the Navigator section of Figure 2, learners could select any chapter in their attempts and once click one of the material, it would trigger the system to show corresponding content on the material section. If learners felt uncomfortable while reading the material because of the relatively small screen size, they were allowed to hide the JMSN interface as shown in Figure 2 by just clicking on the Hiding button. Of course they could make JMSN interface visible whenever they needed it by once again clicking on the Hiding button.

![Figure 2. Main learning environment window consisting of JMSN messenger (left), web content (top right), and discussion forum (bottom right) panes. Text highlighted with Vpen and textual annotations are seen in the web content pane.](image)

Participants

Three classes, totaling 135 college students in Taiwan, participated in the experiment, which was conducted in context of a college course. Of the participants, 56 were male and 79 were female. They all were not major in computer related department. Instead, 45 of them were major in Early Childhood Care and Education; others were major in Foreign Languages and Applied Linguistics. The title of the course was “Basic theory of computer” and it ran for 3 months from September 2005 to November 2005. Content of the course included “Fundamental of computer architecture”, “Word”, “EXCEL” and “PowerPoint”. During the 3 months, instructors would teach not only theoretical knowledge but also operating skills of the applications. The course activities were structured as follows: each week, a three hours face-to-face lecture/tutorial in computer classroom was followed by individual self-paced problem solving, with problems given as homework. Individual work was supported by providing relevant web-based materials and web-based communication tools. Usage logs for web-based facilities provided us data reflecting interaction, outer-action and intra-action patterns.
Research Tools

Web-based learning environment

We provided four communication tools to enable and to capture intra-action, interaction and outeraction. Figure 2 and Figure 3 show screen snapshots for these communication tools.

![Image of communication tools]

**Virtual pen system (Vpen):** Allows students to make annotations of web-based materials. As when making annotations on paper, students can easily use Vpen to highlight, underline, or textually comment on the web-based materials. As in our system students are not provided a facility to share annotations.

**Discussion board:** Supports deferred communication: primarily used by students to post questions and request help, and to answer such queries. Some of the messages posted to the discussion forum are not directly related to the subject matter of the course, and involve encouraging, complaining, offering excuses etc.

**JMSN Messenger:** A real-time communication tool based on Java technology. Message board: Supports deferred communication. While discussion board is offered as a separate pane on the same window as web-based materials, and thus is most suitable for use in context of these materials, message board is offered in its own window and offers more direct access to message content, making it most suitable for sharing information pertaining to the course overall and information perceived to be particularly important.

**Interview**

An unstructured face-to-face interview with randomly selected 20 course participants was conducted after the experiment. The primary aim of the interview was to elicit student perceptions regarding the impact of interaction, outeraction and intra-action on their learning. This allowed us to have an insight into whether the relationships uncovered by the experiment are due to cause and effect. All interviews were recorded, transcribed, and later analyzed.
Experiment design

Due to ethical considerations, it was not possible to set up control groups and we designed the experiment as a correlational study. Over the duration of the course, we logged student usage of the blended learning system. Each instance of usage was later manually (based on researcher's judgment) classified as an instance of interaction, outeraction or intra-action.

At the beginning of the course, students were given homework to introduce themselves by uploading a self-introduction, as text and as a voice message, and a photograph to the discussion board. This way, students acquainted with each other while we were able to confirm that they can use the communication tools. After the initial two weeks devoted to learning to use communication tools and to self-introduction, the main learning topics were presented in sequence. There were four learning topics in the course. Each topic was two weeks long. For each topic, teacher released homework covering the topic via the discussion board, and offered some relevant self-assessment units. Students were instructed to complete homework and self-assessment on their own. If they had questions about learning materials, homework or self-assessment, they could communicate with peers by using JMSN, discussion board or message board. At the end of the course, students underwent a final examination. After course completion, we conducted unstructured interviews to deepen our understanding of learners’ detail perspectives about the learning experiences. In addition, student learning outcomes, measured as final examination scores, were available for the study.

Research variables

We formulate three purposes of our research as follows:
- to study the relationships between intra-action, interaction, and outeraction;
- to study the possible effect of intra-action, interaction, and outeraction on students’ learning achievements;
- to consider, which kind of communication: intra-action, interaction or outeraction, is likely to have the greatest impact on students’ learning.

Our research variables include intra-action, interaction, outeraction, and learning achievements. Figure 4 illustrates the research structure of the study. Pearson correlation was used to measure the strength of relationships between different kinds of communication (a, b, and c). Regression analysis was used to measure relationships between different kinds of communication and learning outcomes (d, e, and f). Table 1 provides operational definitions of research variables.

![Figure 4. The research structure of the study](image-url)
### Table 1: Operational definitions of research variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Example indicators</th>
</tr>
</thead>
</table>
| Intra-action    | A set of self communicative processes that occur when individuals employ any form of symbols with meaning to externalize his own representation of knowledge. | • Drawing conclusions on the material  
• Writing meaningful symbols by annotation to support one’s learning  
• Highlighting words, lines or other elements on the material to remind oneself |
| Interaction     | A set of communicative processes that support joint problem solving, coordination and social learning (Nardi et al., 2000). | • Asking a course relevant question  
• Seeking out specialized information  
• Request clarification of one’s idea  
• Raising doubts and querying |
| Outeraction     | A set of communicative processes outside of information exchange, in which individuals reach out to others in patently social ways to enable information exchange (Nardi et al., 2000). | • Using words to express one’s feelings  
• Writing jokes  
• Presenting greetings  
• Self-introduction  
• Presenting closure |
| Learning        | Learning achievement refers to the extent the learners understand content of the course in this study. | • Correctly distinguish functions of computer units.  
• Correctly operate computers to use the applications to finish a given task |

Based on the definition mentioned above, two graduated student who had been told the difference between interaction and outeraction was responsible for the classification of each post in discussion board, message board, and JMSN in terms of the content of the sentences. A sentence would be labeled as “interaction”, if the content was mainly related to the course material. Alternatively, it would be classified into “outeraction”, if the sentence was relevant to socially interact, or to solve problems that outrange the course. As a result, unit of analysis in this study was the actual meaning of the sentence in JMSN. In particular, a conversation conducted via the use of JMSN usually contained several sentences, which means each sentence of the conversation might have different communication type. Therefore, JMSN would save each sentence and regard it as a record in the database. Furthermore, intra-action could be easily identified by using the VPEN and differentiated from the other two communications. And the definitions of the two kinds of communication were also obvious, thereby reducing learners’ messages to material-specific content and social-specific content, the coder was able to clearly distinguish which one was an interaction, whereas another was not. Similar to the case in JMSN which employed the meaning of the sentence as our unit of analysis, learners’ posts were coded in terms of their meaning of the messages. As a result, whether a post was an interaction or outeration depended mainly on the messages the posters actually addressed. Two coders were involved in the coding process. Both of them coded the messages independently and the percentage of agreement was found to be 82%.

Unlike interaction and outeraction, intra-action was much easier to identify because the system automatically recorded each learners’ intra-action. As a learner used the annotation tool to mark, underline, or comment the material, the system counted the number of annotation and further recorded that whether it was a mark, underline, or comment. This information was then used in the following analysis and helped us distinguish between explicit annotations and inexplicit annotations. In this study, we only focus on those intra-actions conducted by Vpen, but did not consider those conducted by discussion board or JMSN. The main reason was that although students could use any tools to conduct intra-action in their attempts, it was seldom happened for participants to post messages reminding themselves in discussion board or JMSN. Particularly, when one has come out ideas with the material, it is intuitive for one to highlight the words or the lines, or to make comments directly on the material. This was also the reason why we integrated Vpen into the online material.

Then, we are able to compare counts for different communication kinds directly to each other, and it is meaningful to directly compare regression coefficients for learning achievement predicted via different communication kinds. Finally, learners were asked to conduct a post test to indicate their learning achievements. In the post test they were asked to answer several quizzes associated with the course material and were all told that the test score would be
weighted as their final grade. Sample questions of the quizzes were “What’r e the differences between EXCEL and ACCESS”, “What the five major components of a computer are and what their functions are while comparing them to a human body”. Besides, learners also were asked to operate the computer to fulfill the requirements proposed by the instructors. The requirement was, for example, “Please use pivotal table to find out the average salaries of managers in different industry.” The total score of learning achievement was 100 and the score was given according to the answers and efforts the participants actually exhibited.

Results and discussion

Distinguishing explicit and inexplicit annotation

On inspecting the raw data, we found that there was a significant difference in the quantities of underlining and highlighting among students. The reason was that some students were used to underlining or highlighting the materials without consciousness or purpose. Underlining or highlighting learning materials is just their reading habit. To get rid of the disturbance caused by annotation with no meaning, we considered an alternative operational definition for intra-action, based on counting the number of comments (explicit annotation) only, and disregarding all highlights and underlines (inexplicit annotation). Here, the distinction between explicit and inexplicit annotation follows Marshall’s (1997) classification.

Correlations between intra-action, interaction, and outeraction

The results of the Pearson correlation analysis indicated that intra-action, interaction, and outeraction were strongly related to each other. Relationships between intra-action and interaction and between interaction and outeraction were particularly strong. A likely reason of this outcome is that different kinds of communication are triggering each other (as discussed in the section 2.2.4 above). Numerical values are shown in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Intra-action (Total counts)</th>
<th>Intra-action (Explicit)</th>
<th>Interaction</th>
<th>Outeraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-action (Total counts)</td>
<td></td>
<td></td>
<td>.307(**)</td>
<td>.171(*)</td>
</tr>
<tr>
<td>Intra-action (Explicit)</td>
<td>.497(**)</td>
<td></td>
<td>.225(**)</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>.307(**)</td>
<td>.497(**)</td>
<td></td>
<td>.322(**)</td>
</tr>
<tr>
<td>Outeraction</td>
<td>.171(*)</td>
<td>.225(**)</td>
<td>.322(**)</td>
<td></td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed)

Relationships between each of the communication kinds and learning achievements

Correlation

Pearson correlation was also utilized in this analysis to examine the strength of the relationship of each communication kind with students’ learning achievement. The results showed that all communication kinds (intra-
action, interaction, and outeraction) were significantly associated with students’ learning achievement. Table 3 shows the results. The more communication students exhibited, the higher learning achievement they obtained. The relationship was especially strong between explicit annotation and learning achievement.

Table 3. Pearson correlation between learning achievement and each communication kind

<table>
<thead>
<tr>
<th></th>
<th>Intra-action</th>
<th>Interaction</th>
<th>Outeraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>learning achievement</td>
<td>.217(*)</td>
<td>.411(**)</td>
<td>.294(**)</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed)

Regression

Regression analysis was utilized to estimate the significant coefficients for each communication kind. First, we used simple regression analysis to test how well one can use each of the three communication kinds to predict learning achievement. We can know predictability of each variable from the value of R-square and β. The value of R-square stands for the percentage of a dependent variable explained by predictors, and the value of β stands for the slope of regression. The value of β is obtained from B after standardizing. Therefore, we can compare predictability of learning achievement by different predictors. The larger is the modulus of β, the larger part of learning achievement variation is accounted for by the corresponding predictor. Secondly, to gain further insight, we used multiple regression analysis to examine the predictability of learning achievement by all communication kinds considered simultaneously.

Simple regression

Predictability of learning achievement by each communication kind is shown in Table 4. Because R-square and β for each communication kind were significant, it meant that each communication kind was suitable to predict learning achievement. The β values for intra-action, interaction, and outeraction are 0.217, 0.234, and 0.294, respectively. The sign of β is such that higher level of each communication kind corresponded to higher level of learning achievement. Specifically, intra-action estimated via explicit annotation, with β value of 0.411, related much stronger to learning achievement than intra-action estimated via all annotation types (β value of 0.217). This result is consistent with the view that students often make inexplicit annotations with no meaning. Hence, count of explicit annotations is a better operational definition of intra-action than count of all annotations, explicit and inexplicit. For student population we had in the course, inexplicit annotation activity is, effectively, not an intra-action activity.

Table 4. Predictability of learning achievement by each communication

<table>
<thead>
<tr>
<th>Variables</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Estimate of B</th>
<th>Standardized-β</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-action(total count)</td>
<td>.047</td>
<td>.040</td>
<td>.092</td>
<td>.217</td>
<td>2.568*</td>
</tr>
<tr>
<td>Intra-action(Explicit)</td>
<td>.169</td>
<td>.163</td>
<td>.412</td>
<td>.411</td>
<td>5.197**</td>
</tr>
<tr>
<td>Interaction</td>
<td>.087</td>
<td>.080</td>
<td>1.247</td>
<td>.294</td>
<td>3.553**</td>
</tr>
<tr>
<td>Outeraction</td>
<td>.055</td>
<td>.048</td>
<td>2.815</td>
<td>.234</td>
<td>2.772**</td>
</tr>
</tbody>
</table>

Multiple regression

To gain further insight into which communication kind has the most significant effect on learning achievement, we used multiple regression analysis. As simple regression results indicated that inexplicit annotation does not measure intra-action (as discussed in the previous section), for multiple regression analysis we considered explicit annotations only.

The result shown in Table 6 indicated that the R-square increased from 0.169 to 0.180 after adding interaction in model 2. In turn, the R-square increased by 0.015 after adding outeraction in model 3. According to model 3, the β value of intra-action is 0.34, which is considerably larger than that of outeraction (β = 0.13) and interaction (β
The results showed that the relationship of intra-action with learning achievement was stronger than the relationship with learning achievement of interaction or outeraction.

It’s worthy to emphasize that interaction had the smallest predictive power according to multiple regression analysis. However, the results shown in simple regression analysis indicated that outeraction should be the one has the smallest predictability. The reason for this contradiction was that the correlation between interaction and intra-action was so high that the predictive power of interaction decreased when we put all variables together in the multiple regressions analysis.

To verify whether interaction and intra-action are two independent variables, the variance inflationary factor (VIF) was employed to measure the collinearity of our independent variables. In statistics, VIF measures how much the variance of the standardized regression coefficient $\beta$ is inflated by collinearity. If VIF value is less than 10, it is customary to assume that there is no collinearity between the predictor variables (Neter, Wasserman, & Kutner, 1985). In Table 5, all VIF values are between 1.116 and 1.416. It means that there is no collinearity among the three kinds of communications. That is, each kind of communication is to a considerable degree an independent predictor of learning outcomes, and, if cause and effect is assumed, each communication kind affects learning in its own way.

<table>
<thead>
<tr>
<th>Model</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Variables</th>
<th>Estimate of B</th>
<th>Standardized-$\beta$</th>
<th>$t$</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.169</td>
<td>.163</td>
<td>Intra-action</td>
<td>.412</td>
<td>.411</td>
<td>5.197**</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>.180</td>
<td>.167</td>
<td>Intra-action</td>
<td>.379</td>
<td>.377</td>
<td>4.694**</td>
<td>1.328</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Outeraction</td>
<td>1.794</td>
<td>.149</td>
<td>1.853</td>
<td>1.116</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Outeraction</td>
<td>1.572</td>
<td>.130</td>
<td>1.571</td>
<td>1.123</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interaction</td>
<td>.353</td>
<td>.083</td>
<td>.893</td>
<td>1.416</td>
</tr>
</tbody>
</table>

### Table 5. Predictability of learning outcomes by multiple communication kinds simultaneously in multiple regression

**Interview analysis about interaction, outeraction, and intra-action**

In the interview, we asked students their opinion about the three kinds of communication. Regarding intra-action, students generally thought that it did provide chances for them to reflect on what they learned. For example, one student thought that using Vpen let her learn more. More specifically, when she had a problem doing exercises, she mentioned, in the past she used to search for an answer by looking for data in the library or on the Internet. Now, she would rely to a large degree on her own annotations to help her to come up with her own answer. In her case, we found that intra-action may increase motivation to reason in terms of the target domain, rather than to just to search for a ready answer. The following is the content when interviewing her:

“The annotation is really helpful for doing some exercises. I usually went to library or found classmates to help doing exercises. In the beginning, I was not good at using computers. Nowadays, my computer skill become better and better and my own annotations done in class give vital clues to solve exercises”

“Writing notes on the web material makes me convenient because I don’t have to bring my notebook every time I go to the class”

Other students mentioned that they tended to shift from intra-action to interaction (e.g. highlighting a problematic chunk of material, and then discussing it with peers).

The following is one student’s opinion in the interview:

“If having any questions when I study by myself, I can easily find other classmates online and ask for their help. It is very convenient and efficient for further study”

“At the same time, the website also helps me interact with other students. So I can upload my homework to the instructor and even watch others’ effort in a short time. This function gives us opportunities to review different comments or solutions proposed by our classmate. So we can learn from one another, I think this is exactly what our teacher wants us to do”

Students mentioned that the ability to connect instantly was essential for socializing with other students (and thus, for outeraction). In addition, many students thought the most important requirement to learning-related communication
for them was getting the answer as soon as possible when they had encountered a problem. Hence, for some of the students, the ability to conduct both outeraction and interaction in real time or almost in real time (e.g., via JMSN) is essential.

Discussion

In their work, Hutchins (1995) and McClelland his associates (1986) unveil the answer to what constitute distributed cognition. Although they successfully convey the idea that cognition may be distributed throughout a system comprising both individuals and artifacts, they leave the process of how the cognition spreads out blank. Thus, the answer to this question has remained elusive. In this regard, we believe our study makes several contributions. First, Giere (2002) argues that collaborative practices, involve distributed cognition, as these practices are “situation(s) in which one or more individuals reach a cognitive outcome either by combing individual knowledge not initially shared with the others or by interacting with artifacts organized in an appropriate way (or both)” (2002, p. 641). More specifically, this study suggests it is the interplay among intra-action, interaction and outeraction that result in the artifacts to be transformed into a mental process. Drawing on an online course and providing communication facilities to assist all students to conduct the three communications, this study recorded most of the three communications happened in the blended learning system. Students in this study were asked to participate in the online course activities, indicating a heavy dependence on the system for them. Apparently, participants have to make use of the system often to have a communication with others. In addition, if students come out an idea with the material, they could exploit Vpen, in most cases, to make meaningful representation of their knowledge. Although the quantity of explicit and inexplicit annotations, two kinds of intra-action in this study, may not fully reflect the relative significance of the content, it has widely documented that writing notes indeed convey more meaning and impact than highlight the words or lines (Howe, 1997; Hwang et al., 2007; Marshall, 1997). Along with the line, results of this study not only consistently appreciate the significance of explicit annotation, but also support our suppose that intra-action plays an important role in blended learning settings. More specifically, before individuals can effectively collaborate with one another, they need to be aware of other peers (Daniel et al., 2003; Gutwin & Greenberg, 1998). Outeraction, therefore, provides connections to multiple individuals for the future shift of cognition. Whereas interaction mainly focuses on the exchange of external representation across multiple individuals, intra-action is specific to an individual’s cognition process of forming inner representations of knowledge.

Second, distributed cognition includes not only cases where a cognitive task is distributed across multiple individuals, but also cases where such a task is distributed between a single individual and his artifacts, such as annotations. While researchers often focus on the other two proprieties of distributed cognition, this study introduce the notion of intra-action to emphasize the importance of distributed cognition between a single learners and his external representations. Our results documented that intra-action with only explicit annotation accounts for most of the variance of students’ learning outcomes, indicating its significance in the process of distributed cognition in blended learning settings. In fact, both in terms of developmental process of the human species and that of each individual, meanings are generally first represented as artifacts in one’s attempt. Seeing that the relationship between explicit annotation and interaction is high, in line with the principles of Vygotsky, this study emphasizes that knowledge should have been externalized by individuals as meaningful symbols, before it can be socially constructed in interactions among learners. Although this study only employed learners’ annotations to refer to meaningful symbols in their attempts, we believe these annotations function, partly, like artifacts required for later knowledge construction. In this case, intra-action clearly promotes learning and interaction. Figure 5 illustrates that intra-action invokes or evolves into interaction. When one student studied learning material or conducted some self-evaluation test, he took note beside the learning materials since having questions. Afterwards, he posted his question in the forum and lots of discussions were invoked among peers.

Third, we introduced one way that can measure status of the constituent elements of distributed cognition. Multiple individuals was measured by the outeraction communication, external representation was estimated by interaction. In addition to the two traditional constituent elements, we extended our understanding of the distributed cognition by including the notion of intra-action, which was evaluated by explicit and inexplicit annotations. In order to measure objective quantity of learners’ distributed cognition in blended learning settings, studies may consider the amount of
the three communications. On the contrary, to measure the quality of learners’ distributed cognition, they may consider content of the communication.

Finally, we learned it from this study that the production of intra-action, interaction, and outeraction might be influenced by technological functions or the features of environment learners were facing. In short, Vpen in this study mainly supported learners’ intra-action. While the discussion board primarily supported learners’ interaction, it also supported outeraction. On the contrary, learners usually employed JMSN for their outeraction, but occasionally for interaction. Or learners might use the message board for outeraction, but sometimes for interaction. Accordingly, this study summarized what this study learned in Table 6.

<table>
<thead>
<tr>
<th>Table 6. Research tools–three kinds of communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-action</td>
</tr>
<tr>
<td>Vpen</td>
</tr>
<tr>
<td>Discussion board</td>
</tr>
<tr>
<td>JMSN</td>
</tr>
<tr>
<td>Message board</td>
</tr>
</tbody>
</table>

Figure 5. Illustration of how intra-action invokes or evolves into interaction

Lots of interactions were invoked among peers

Intra-action invokes interaction (From the upper figure, one student had a question when he conducted a self-evaluation test. Then he posted the test item in the discussion forum)
Table 6 shows the learners’ preferences of exploiting different technological functions or features to support corresponding communication types. The relationships between the three communication types and the tools mentioned above. We implemented these tools by making use of a number of free, open source components and technologies, and seamlessly integrated them into a single blended learning environment. As we had full control and understanding of the source code, it was possible to log user actions at sufficient level of detail.

Implications

Based on our findings, several implications raised for practitioners and academics. First, research (Brown & Smiley, 1978; Howe, 1997; Hwang et al, 2007) has acknowledged that writing notes leads to a better performance. More specifically, this study documented that explicit annotation is more important than inexplicit annotation. Seeing that previous research has addressed little comparison of explicit and inexplicit annotation, one of the purposes of this study thus was on the understanding of intra-action. This study suggests both of which are measurable components of intra-action, thereby the relationships among interaction and outeraction are of interest. As a result, to work out intra-action, comparison of these two components seems provide more information for us to deepen our understanding of intra-action. Along with this finding, we suggest researchers of interest to this domain may consider on the influence of cognitive style because cognitive style, originally proposed by Allport (1937), refers to an individual’s habitual way of perceiving, remembering and thinking. Recently, Riding and Rayner (1998) refer it to “an individual’s preferred and habitual approach to organizing and representing information”. As a result, it is worthwhile for us to investigate whether learners’ cognitive style affects their willingness to conduct an explicit annotation because of the preferences to organize and represent information. In line with the reasoning, future research may test this relationship to see if it is true or not.

Second, research on knowledge management focuses on exchanging of explicit knowledge and tacit knowledge among multiple persons. In fact, we may conclude that many studies on this domain deal with one of the three communications, namely interaction, a kind of communication involving more than one person. Unlike intra-action emphasizing the importance of explicit annotation; Interaction, in line with findings from knowledge management, implies that tacit knowledge is more influential. Future research may design apprenticeship activities and then examine whether tacit knowledge plays a critical role in explain variation of students’ learning performance. In addition, research advocating the significance of explicit knowledge consistently invited only students, who mainly dealt with relatively structured knowledge. On the contrary, initiators of tacit knowledge focus on employees, who addressed relatively unstructured knowledge because of the complexity of business context. In this regard, it seems that the degree to which knowledge is structured determines the importance of tacit knowledge and explicit annotation. Future research may also take into account the moderating effect of this construct.

Third, formative assessment is especially helpful for students to assess their own situation, and for teachers to adjust lesson planning and activities (Bransford et al., 2000). In line with TDC, much attention has drawn on external representations among students. While agreeing the importance of evaluating this information among learners, this study considers that a learner’s self representation should not be ignored. The evanescent nature of ephemeral evidence of learners’ self representations makes it compulsory to be recoded immediately, especially when the impact on learning performance has been corroborated. In this regard, when instructors are willing to employ the message content in blended learning system to conduct a formative assessment, they may wish to task into account learners’ intra-action as well. This implies that blended learning developers need to shift part of their attention on the implementation of mechanism that supports learners’ self inner representation, namely intra-action.

Finally, a cognitive barrier is, for example, the difficulty to bridge the distance between expert and novice (Hinds & Pfeffer, 2003). Our study shows that intra-action, interaction, and outeraction fundamentally shape cognition across multiple learners in blended learning setting, making cognition truly social. As a result, it uncovers a good news for instructors and educators since the cognitive barrier results from electronic networks may be overcame. Thus, a pedagogical activity that helps learners be used to explicitly annotate material as inner representations is supposed to be important, because it in turn makes students easily evolve into a meaningful interaction. And then a student may summarize the external representations, a prerequisite of distributed cognition, derived from the interaction into his inner representations. Alternatively, activities that promote learners’ social connections are valuable as well, particularly when learners are dispersed in physical location. Outeraction makes learners with knowledge awareness (Daniel et al., 2003; Gutwin & Greenberg, 1998) and availability awareness, which facilitates the negotiation of
conducting an interaction. On the contrary, for the future interaction, learners are also motivated to participate in outeraction, leading to multiple individuals available, another prerequisite of distributed cognition.

Limitations

While there are interesting implications for research and practice to our findings, there are also some methodological limitations inherent in our approach. While the findings are consistent with our suggestion of the interplay among the three kinds of communication, the crosssectional nature of our data limits the extent to which our logical explanations can be conclusively supported by the data. Additional research, ideally involving longitudinal data, is needed to fully address the issue of the interplay direction. Another methodological limitation concerns the generalizability of our findings. Specifically, there are issues of generalizability that relate to the restrictions that a student sample brings. The idiosyncratic nature of the sample is unavoidable in most forms of research, but we are encouraged by age of our respondents. Because the participants were all the freshmen of university, their age are so close, which could result in a restriction of range. Replication of this study in samples with a broader range of age would be a good way to address this limitation of the current study. Finally, the limitation of using a particular blended learning system developed in this study may have affected the way in which the participants learned from the course material, leading to the variation of learning performance.

Conclusions

In this study, in context of an blended learning course, we demonstrated that intra-action, interaction, and outeraction are positively correlated. More importantly, in statistically significant way, higher incidence of communication corresponded to better learning achievement. The tendency of higher level of intra-action activity to correspond to higher learning achievement is particularly clear. This is consistent with the widely held view that intra-action activities, such as note-taking, promote learning. Unstructured interviews conducted with students supported the view that the positive relationship between high level of intra-action and high learning outcomes is to a large degree a cause-effect one.

For annotation as intra-action activity, our research indicated that only explicit annotations can be viewed as a reliable measure of intra-action, as inexplicit annotation are often applied as motor activity with no cognitive purpose. This result has implications for future blended learning systems design, as well as for the way existing systems are configured by teachers and blended learning material designers - enabling explicit annotation is likely to have more positive effect on learning outcomes than enabling inexplicit annotation.

In a wider context, our results suggest that designers of blended learning courses should reconsider their practices, and put more emphasis on offering intra-action tools to students. We notice that while interaction and outeraction carry risks, such as bulling among students or inappropriate and unfair criticism of the course, intra-action tools are virtually risk free as their inappropriate use is hardly possible.

Our study underlines the importance of intra-action in blended learning environments. While the study is, due to ethics considerations, of correlational nature, by showing the importance of intra-action it opens way to further studies of intra-action, which eventually may justify experimental designs demonstrating cause-effect between intra-action and learning outcomes in a more direct manner, but at the expense of certain inconveniences for some of the human subjects involved.

Acknowledgement

This study was supported by the National Science Council under the grant number: NSC-96-2524-S-008-002.

References


Hinds, P. J., & Pfeffer, J. (2003). Why organizations don’t know what they know: Cognitive and motivational factors affecting the transfer of expertise. In M. Ackerman, V. Pipek & V. Wulf (Eds.), Beyond knowledge management: Sharing expertise (pp. 3-26), Cambridge, MA: MIT Press.


CogSkillnet: An Ontology-Based Representation of Cognitive Skills

Petek Askar and Arif Altun
Hacettepe University Dept. of Computer Education and Instructional Technologies // paskar@hacettepe.edu.tr // altunar@hacettepe.edu.tr

ABSTRACT

A number of studies emphasized the need to capture learners’ interaction patterns in order to personalize their learning process as they study through learning objects. In education context, learning materials are designed based on pre-determined expectations and learners are evaluated to what extent they master these expectations. Representation of these expectations in learning and assessment objects, on the other hand, is a new challenge for e-learner providers. In order to address this challenge, POLEonto (Personalized Ontological Learning Environment) proposes a new method to separate these expectations by determining domain concepts (ConceptNet) and cognitive skills (CogSkillNet) for expectations via creating cognitive skill and concept ontology for K-12 education. In this paper, we report only the development and design processes of CogSkillNet within POLEonto environment; and, we further discuss how CogSkillNet can be modeled in the e-learning domain. We also describe how ontological representations play a role in creating personalized navigational guidance for allowing visualization of cognitive skills and providing useful navigational feedback to learners.

Keywords
Ontology, Ontology Development, Cognitive Skills, Educational Expectations, Instructional Design

Introduction

Semantic web opened a new paradigm to design adaptive e-learning systems. By using the capabilities of semantic web approach, World Wide Web led the interchange of information about data (i.e., metadata) as well as documents. Such capabilities also indicated a new kind of challenge for e-learning providers to design a common framework that allows learning objects to be shared and reused within and across applications.

An ontology is an explicit specification of a conceptualization (Gruber, 1995) or a model (Musen, 1998), which is used for structuring and modeling of a particular domain that is shared by a group of people in an organization (O’Leary, 1998). Domain ontologies provide explicit and formal descriptions of concepts in a domain of discourse, their properties, relationships among concepts and axioms (Guarino, 1995).

Ontologies are being developed in order to make learning objects reusable (Qin & Hernández, 2006), interoperable (Biletskiy & Hirtle, 2006), and to increase the reliability of using computational (intelligent) agents in e-learning environments (Koper & Tattersall, 2005). Consequently, an ontology design is needed to equip the learning objects with the ability of semantic relations among different types of objects in educational settings.

In order to model the semantic relationships, various models have been designed by various researchers and committees. In their model, for example, Bailey, et. al. (2006) use the nugget metaphor and focus on the need for assisting teachers in the thought processes involved in selecting appropriate methods, tools, student activities and assessments to suit the required learning objectives. Similarly, IMS Learning Design (LD) model proposes an application of a pedagogical model for a specific learning objective, target group, and a specific context or knowledge domain (See, Koper and Olivier, 2004). Another model is proposed by IEEE Learning Technology Standards Committee (LTSC). In this model, learning object metadata covers areas of learning technology, digital rights, metadata, and structured definitions related to instruction.

These attempts to model the learning processes invited several criticisms. For example, Paquette (2007) points out that the actors’ competencies in IMS-LD are indirectly defined in learning units or activities only if educational objectives are associated. Such unstructured model for competencies makes the system unable to provide personalized learning activities for its users, since this association between learning objects within a unit of learning is a key concept for semantic web applications (Berners-Lee, Hendler, & Lassilla, 2001). Waldman, Hopkins, & Burgess (2004), on the other hand, claim that the use of metadata in IEEE-LTSC model is based on metadata schemas, which are user-defined; therefore, reusability and interoperability of learning objects and the quality of user...
created metadata became questionable. In addition, Murray (1998), in his model for distributed curriculum on the web, draws attention to certain pragmatic problems as well, among which are agreement on vocabulary for topics, terms (pedagogical aspects of the topics), subjective meaning for a term, differing context and limits.

In order to reach a more refined mechanism of matching what learners need and making inferences from users’ interactions, some form of domain ontology (e.i., Davies, van Harmelan, & Fensel, 2002; Breuker, Muntjewerff, & Bredewej, 1999) and levels of mastering scales (e.i., Paquette, 2007) is needed. One of the approaches to determine the levels of mastering scales is to use existing taxonomies, such as Bloom (1956)’s taxonomy and apply the design axioms to make inferences from them. Another approach that is being proposed in this paper is to design an ontology representing the cognitive skills in the domain of education and describes the development and design processes of skill ontology, called CogSkillNet, which is an embedded agent in an e-learning environment.

Using an ontology to represent and model the cognitive skills has several advantages over taxonomies. A taxonomy is a hierarchy and each skill is usually assigned a code within the hierarchy. In an e-learning environment, this code also encodes its path. Let’s assume that in an educational expectation, a learner is requested to “analyze interpersonal communication”. Certain cognitive skills can be grouped within the taxonomy of “to analyze” (in Bloom’s taxonomy, to analyze covers to compare, to contrast, to deconstructs, and some others). By relating the learning objects with a cognitive skill in this taxonomy, it is possible to give this skill a certain amount of semantics. For example, when a learning object uses the skill “to compare” in order to describe its semantics, we can understand that this learning object is about “to analyze”. Therefore, when a learner looking for a learning object to study how to “compare” can search for “to analyze” to obtain all related skills within the taxonomy. Note however that these skills might be referred in other skills within the taxonomy. For example, “to compare” as a skill is also included in “to evaluate” in Bloom’s taxonomy. There can be several such properties like these skills within expectations and hence the learner has to go through all skills found to manually pick the skill that satisfies the expectations (or educational standards). In short, taxonomies do not help in this respect.

In this paper, we discuss the advantages of modeling the cognitive skills through ontology. We further note that, once the cognitive skills ontology is defined, it is also necessary to relate the defined ontology with concepts embedded in educational expectations within a learning space. Therefore, we describe how this can be achieved in an e-learning environment.

The paper is structured as follows: First, a background about using ontology in education is summarized. Second, the need for ontology in order to represent cognitive skills will be argued. Then, CogSkillNet ontology development process will be explained. Fourth, ontological representation of cognitive skills will be contextualized within K-12 educational expectation space. Finally, further recommendations for e-learning providers and researchers are stated.

**Background**

There are several ontologies being developed in the field of education. One of them is EduOnto. EduOnto is based on the metadata schemes for The Gateway to Educational Materials (http://www.thegateway.org/) and its controlled vocabulary. The class types include reusable classes (Person, Organization, and Contact), resource object classes (instructional, informational, research), and vocabulary classes (subject categories and terms) (Qin & Hernandes, 2006). Another ontology is Personalized Education Ontology (PEOnto). PEOnto claims to provide learners relevant learning objects based on their individual needs. In PEOnto, five interrelated educational ontologies (curriculum ontology, subject domain ontology, pedagogy ontology, people ontology, and personalized education agents) are being employed (Fok, 2006).

Another project is ELENA, which uses the metaphor of smart space for learning. ELENA claims to provide a personalized space for learners to meet learners’ individual needs with available resources in educational meta repositories and services, such as brokering, learning, service evaluation and assessment services. ELENA includes a number of Personal Learning Assistants. These assistants use learner profiles to search for, select and negotiate with suitable learning services (Klobucar, Senicar, & Jerman-Blazic, 2004).

EML (Educational Modelling Language) is another system developed at the Open University of the Netherlands. EML focuses on representing: (a) the content of a unit of study (e.g., texts, tasks, assignments) and (b) “the roles,
relations, interactions and activities of students and teachers”. EML goes beyond standards such as IEEE LOM to model the social context of education, and the learning, unit of study, domain, and learning theory models which form the pedagogic meta-model can be construed as a set of ontologies (Koper, 2001).

POLEonto is another ontology, which is still under development within POLE (Personalized Ontological Learning Environment) e-learning platform specifically for K-12 education. POLE is being conceptualized by the learning space metaphor (Altun & Askar, 2008), which represents the spatial dimension where learners navigate through the links in order to reach a destination in a given period of time. This learning space includes a combination of various skills (any mental skills that are used in the process of acquiring knowledge) and concepts (contextualized terms within a domain and more specific than knowledge) in defining (or modeling) an expectation or an education standard.

An educational expectation or a standard, set by either the board of education or instructors themselves, connotes something external, extrinsic, and explicit. Moreover, when these statements are written, internal ideas, objectives, and goals are put into external forms so that they can be shared with others (Seels & Glasgow, 1998). It is important to emphasize that instructors are expected to reach and to perceive the same meaning from these expectations and standards. In order to achieve this end, we propose a method to separate these expectations by deconstructing them as concepts and skills for each expectation via creating cognitive skills ontology (hereafter CogSkillNet) and concepts ontology (ConceptNet) separately in order to represent an expectation space from an ontic perspective (See Figure 1).

In POLEonto, learning processes were defined as a set of cognitive skills, which were embedded in the curriculum and requested by instructors. In POLEonto context, skill is defined as the interaction and any processes between persons and concepts. For example, the concept of “square” is envisioned in one’s mind; yet, they can define it, they can extend square into some other thing (i.e., a table or a flower-stand), which is creative thinking. The square can be manipulated to approach a problem by using its types and functions, which requires problem solving. In the following section, arguments for ontological representations for cognitive skills will be provoked and the methodology of CogSkillNet ontology development will be explained in detail.
Why is a need for ontology to represent cognitive skills?

CogSkillNet ontology has been proposed to address two major needs. First, the ontology design in an expectation-based education context includes concept space and skills space (See Figure 2). The elements in an expectation represent a conceptual model, which formally constraints the semantics of the concept taxonomy based on explanations formulated in natural language. Skills are dynamic in nature. Therefore, such dynamic relations are needed to be included in an ontology in order to show the execution of a skill within a cognitive process. Therefore, there is a need to define skills and their relational axioms in a dynamic approach.

The second one is related to the architectural specifications. From the proposed specifications, the IMS Learning Design (IMS LD) (2003) has emerged as the de facto standard that addresses the ontological modeling of any specification for learning design. XML-Schema language is suggested and used when modeling the IMS LD specification in order to facilitate the interoperability between software systems.

CogSkillNet presents a new idea of modeling skills for educational context. In their comparison of conceptual modeling methods in databases (DB), knowledge representation formalisms (KB), and Semantic Web ontology languages (SW-onto), Ding, Kolari, Ding, Avancha, Finin, and Joshi (2006) assert that the Semantic Web significantly improves visibility and extensibility aspects of knowledge sharing in comparison with the previous approaches (See Figure 3). Its URI-based vocabulary and XML-based grammar are key enablers to web-scale knowledge management and sharing. Thus, by using semantic web ontology languages, CogSkillNet is aimed to increase usage, sharing and management of learning and assessment objects in educational contexts.
Method

It is a well-accepted fact that there is not a single way of developing ontology (Noy & McGuinness, 2001). Building object-oriented applications with reusable learning objects based on semantic webbing techniques require a sequence of steps to build such an application (Askar, Altun, Kalinyazgan, & Pekince, 2007). In order to reach CogSkillNet ontology, a five-step ontology development process, with an iterative approach to ontology development, has been applied (adapted from McGuiness, 1999). These steps were:
1. Enumerate important terms in the ontology
2. Define the classes and the class hierarchy
3. Define the properties/relations of classes
4. Create Instances
5. Create axioms/rules

When designing the ontology, Protégé was used as an ontology editor. The visualizations are captured by TGViz and Jambalaya plug-ins embedded in Protégé. In the following section, these steps are explained in detail.

CogSkillNet Development Process

Enumerating Important Terms

A term in CogSkillNet refers to an entity, which represents a textual representation of a cognitive skill. Skills are conceptual representations of cognitive processes in an action. At the first step, important terms regarding the cognitive skills in K-12 educational curricula have been determined by using bootstrapping method. Bootstrapping is a technique that enables the designers to pull out their corpus from domain-specific corpora. These terms were clustered according to their functional and semantic relations based on pre-determined expectations in K-12 educational curricula. These clusters help disambiguate the repeated skills within each subject matter; for example, does the action word “state” in a document refer to an idea (i.e., liberal arts), a calculation result (i.e., mathematics), or a measurement (i.e., an interdisciplinary critical thinking activity). Consequently, when a single skill is repeated several times, semantic relations for each skill become evident.
Defining the classes and Class Hierarchy

The clusters obtained from bootstrapping represent the taxonomical classification of cognitive skills in an educational curriculum. Each learner, having completed the stated expectations successfully, is expected to reach a higher order thinking skill via planned instruction. In CogSkillNet, cognitive skills, or actions, are clustered in three hierarchical classes. These action classes are labeled as base actions, encapsulated actions and integrated actions. In CogSkillNet context, integrated actions represent higher order thinking skills such as problem solving, scientific thinking, and critical thinking. The following schema (Figure 4) represents the classes and class hierarchy for cognitive skills in education domain.

Properties and Relations

For the next step, properties for these classes were defined by using implicit semantic techniques in order to find connections and patterns in the data (Askar, Altun, Kalinyazgan, & Pekince, 2007). This process was mainly used to generate and understand the relationships between entities in the selected context. In CogSkillNet, six different types of relations (with their inverses) were identified. These are:
- Y: is an instance of
- X: is a class of
- C: is a superClass of
- A: is a subClass of
- K: is a process_component of
- T: has process_component of

It should be noted that these relationships are dynamic in that they could be increased or limited in use depending on the content area within a curriculum.

Creating Instances

Skills in CogSkillNet include actions, processes (aka relations and relation types) and delegations as instances. Base actions are pre-defined, universal sets of functions. These are axiomatized actions, which do not necessarily require logics or explanations, since they are self-exploratory command-like acts (i.e., to identify, to show, to read). These actions both generate new actions from existing actions and validate the consistency of actions. These acts are processable, can be in the form of compound or simple forms; and, can be defined as visible or invisible.

Additionally, encapsulated actions are tacit processes in nature. These actions are organized from base actions in the form of input-process-output framework. Eventually, each action is designed to return an output. Consequently,
ontology will easily be extensible in an incremental manner by reusing as many existing actions as possible before creating a new action from scratch.

Integrated (Higher order) actions, on the other hand, are designed as tropes. They include certain encapsulated actions as well as base actions in order to fully accomplish the related process. These actions are processed depending on the related acts. Integrated Actions can be in Public or Private State. Public state includes all the process collection and includes the steps in the form of input-process-output framework, whereas Private states refer to attitudinal states, such as to like, to enjoy, etc., which are personal experiences in nature. Private state can be either delegated or non-delegated. Delegated actions are triggered when certain conditions are met. Non-delegated actions do not necessarily require certain conditions to be processed. The following figure represents the conceptual framework of instances in CogSkillNet (See Figure 5).

![Diagram of conceptual framework of instances](image)

**Figure 5. Conceptual Framework of Instances**

**Setting Axioms/Rules**

As a next step, a series of axioms were developed for CogSkillNet. The axioms for the model take an action to be processed. In referring to figure 4, an action contains an act. Each act starts with an initiator as input, processes it, and terminates with an output. The process may have one or more sub-actions. Actions in the CogSkillNet apply the following axioms:

1. Each action can be acted upon.
2. Each action can contain actions (sub-actions).
3. Each action can call all actions’ acts while acting.
4. Each act operates on input and returns output.
5. A returned output in each process is transferred to the next coming act as an input.
6. The output from the last act in the process is the input of the corresponding act.
7. The inputs and outputs can be Null, single, or multiple.
The algorithm for this process can be modeled as in Figure 6.

![Figure 6. Initiator-process-terminator (IPO) framework](image)

IPO framework is necessary to contextualize the cognitive skills in an expectation space. All interaction in this framework represents actions, which are in the form of acts. These acts require an initiator and a terminator. Terminators provide input for any corresponding act; thus, they are transferred to the next coming act as input. Acts are processed until all related sub-actions are completed. An educational expectation requires learners process certain skills based on pre-determined initiators, such as a reading material or a list of statements. Moreover, learners are required to communicate their terminated process either with their peers or with instructors, such as writing a report or delivering an oral presentation. This combination of initiators and terminators with corresponding cognitive skills (let’s say to summarize) in the ontology constitute the context of each educational expectation (or goal).

**Visual Representation of Cognitive Skills and CogSkillNet Algorithm**

By taking into account all the considerations described above, ontology for the cognitive skills has been designed. The main parts of the ontology include actions and concepts as two different classes. Action class represents basic, encapsulated and integrated skills; whereas, concept class includes instances for initiators (input) and terminators (output) to be processed. In addition, depending on the stated relationship, Boolean and Question types were included within the ontology to activate rules and axioms in order to make further inferences. The following diagram displays a quasi-algorithm of cognitive skills within CogSkillNet (See Figure 7).
An essential part of the ontology is the representation of skills in an expectation. In order to determine students’ cognitive performances and to provide them meaningful learning paths, expectations are tied with assessment and learning objects. Thus, CogSkillNet represents and operates on the cognitive skills in line with assessment and learning objects in an e-learning environment (See Figure 8).

In an e-learning environment, each expectation may be associated with learning and assessment objects. Assessment objects represent what learners have accomplished whereas learning objects indicate units of learning materials. By incorporating CogSkillNet, the system will be able to identify the types of and the context for actions for learners. Types of actions refer to action classes (base, encapsulated, or integrated). Furthermore, by means of the declared axioms, the system would provide suggestions to learners which paths to continue. The context for actions, on the other hand, represents the concepts when to initiate and terminate the actions. For example, let’s consider the skill “generalization”. The initiation for generalization can be a “proposition” or a “list” as input depending on the content of a learning or assessment object. Additionally, generalization leads to produce an output, such as a “statement” or an “argumentation”. These concepts indicate a termination of generalization. Altogether, (generalization, statements, and propositions) these concept and cognitive skill classes create a context of learning in an educational domain. The numbers of initiators and terminators can be numerous and provide different contextual information for each cognitive skill.

The action algorithm for basic actions is self-exploratory. In other words, if the input includes a base action, the system will provide either a learning or assessment object related to a single base action. Action algorithm for encapsulated and integrated actions, however, is as follows:

**Action Algorithm for encapsulated and integrated actions**

Input: An Educational Expectation
Output: Suggestion of a learning path

Step1: Determine the action
Step 2: Determine the type of the action
Step 3: If encapsulated or integrated
Then get the Sub-actions
Repeat until all sub-actions are processed
Loop all actions until reaches to base actions
Else
Base actions

Furthermore, the system further contextualizes the educational expectations (See Figure 9) by incorporating the nature of actions into the process as shown in the following algorithm.

**Input:** Contextualize the expectation  
**Output:** Contextualized e-learning environment

Step 1: Determine the action  
Step 2: Determine the nature of actions  
Step 3: Get the initiators  
Step 4: Get the terminators
Step 5: Use Boolean to combine the initiators and terminators with actions
Step 5: Determine the context
Step 6: If context is multiple
    Then provide pathways accordingly

![Figure 9. The structure of action processes](image)

![Figure 10. A Sample View from CogSkillNet](image)

Let’s illustrate this with an example. Let’s assume a student is given a task (either as in an LO or as an AO) to *make a generalization* either in an assessment object or in a learning object to study. Generalization as a cognitive skill is not self-exploratory but tacit by itself. Since it is not a base action, it includes sub-processes and it is difficult for a student to analyze this process by himself/herself. If the system provides the learners sub-processes of generalization, which are information gathering, determining similarities and determining differences, the task will be more explicit for the learner to process as in the following algorithm.

Input: An Expectation (i.e., students learn to generalize the principles of sliding frictions)
Output: Suggest learning paths based on learners’ progress and responses
Step 1: Check the action in the expectation
Step 2: Determine its class
Step 3: If encapsulated or integrated
Then get the instances of each action within the class

Figure 10 visualizes the graphical representation of generalization.

As visualized above, generalization has process_component of (a) Information Gathering, (b) Determine_similarity, and (c) Determine_Difference in CogSkillNet. Moreover, the system can further suggest the learner other cognitive skills by using the CogSkillNet relations and axioms. Determining differences, for example, will be needed to do comparisons, which is a process_component of decision making. In addition, the actions “Generalize”, “Information Gathering”, “Determine Difference”, and “Comparing” may need a “Collection” as initiator and only “Information Gathering” returns a “collection” as terminator. The following figure (Figure 11) is a visual representation of this ontic relationship.

![Figure 11](image)

Figure 11. Determine Difference as an instance in CogSkillNet

Consequently, the system will be able to provide a personalized path for each learner based on his/her progress and/or responses.

**Conclusion**

This paper proposes an ontological representation of cognitive skills for K-12 education domain. CogSkillNet is one of the two ontologies -- the other one is concept ontology (ConceptNet) -- embedded in POLE e-learning system. Along with the concept ontology, CogSkillNet models the cognitive skills, which are associated with two repositories (learning object repository and assessment object repository) and are embedded in K-12 curriculum and requested by instructors.

On the basis of CogSkillNet, this study proposes the following contributions;
1) learners are provided navigation guidance for their learning path based on their progress through cognitive skills, which aim at scaffolding learning in a constructivist perspective.
2) CogSkillNet provides classroom instructors and instructional designers another tool when designing learning objects as well as other learning materials. These relations and representations of cognitive skills in an ontological space can further be used when making instructional decisions for classroom teaching.

3) CogSkillNet enables evaluators to base their assessment process and diagnose the deficiencies in students’ learning as far as cognitive skills are concerned.

Many researchers hold the belief that the learning specifications, such as IEE LOM, IMS-LD and SCORM are of limited use in that prescriptive metadata have various drawbacks (See, Jovanović, Gašević, Knight, & Richards, 2007 for detailed drawbacks). Jovanović et. al. (2007) proposes the idea of creating a learning object context (LOC) as a unique set of inter-related data that characterize a specific learning situation. They go further to add that their LOCO framework integrates several kinds of learning-related ontologies in order to capture specific context of use. We believe that separating expectations as concepts and skills contribute to further extend the notion of context in use –within or without LOCO framework. As a result, more circular flow of information in the learning process can be maintained.

The cognitive skills of the proposed ontology, targeted for use by automated systems, can be embedded in standard metadata registries of learning and assessment objects to facilitate the design of instructional materials. In defining or adding the new skills to CogSkillNet, we have based these skills on an Action upper-class, under which any action, regardless of having input and/or output, can be placed easily. Moreover, using the Concept class within CogSkillNet, any further concepts can easily be integrated or added to the existing ontology.

The success of any e-learning environment depends highly on instructors’ and learners’ interactions between the system and themselves. We hope that the proposed ontology in this paper leads to more work towards the formalization of the knowledge about cognitive skills and its interdependent use with concepts in e-learning environments. Therefore, future research should address a complete review of the skills in the proposed ontology by domain experts in the creation and design of learning objects in K-12 education. In addition, separating expectations as concepts and skills and combining them in domain ontology is a new perspective for education ontology design. More research is needed to test the efficiency and accuracy of using this design, especially when sequencing, searching, and contextualizing learning objects, for e-learning platforms.

Acknowledgement

We thank anonymous referees for improving the earlier version of the manuscript through their critical review. We also would like to extend our thank to Kagan Kalinyazgan and S. Serkan Pekince from Yuce Information Systems for their valuable support to this project.

References


Engaging students in multimedia-mediated Constructivist learning – Students’ perceptions

Mai Neo and Tse-Kian Neo
Faculty of Creative Multimedia, Multimedia University, Malaysia // Tel/Fax: +6-03-56357817 // kneo@pc.jaring.my

ABSTRACT
This study was conducted in the Faculty of Creative Multimedia, Multimedia University, Malaysia and investigated students’ perceptions while working on a multimedia project that was embedded within a constructivist-based learning environment. We studied the impact of using multimedia on students who have little experience with working in a problem-solving design environment. Students worked in groups and created an interactive multimedia application with Macromedia Director. They were responsible for all project development decisions during their learning process. A survey questionnaire administered at the end of the project captured their perceptions. The students showed positive attitudes towards the project with respect to their learning motivation and understanding, skills and their teamwork abilities. By incorporating multimedia into a constructivist learning environment, students learned to design multimedia, as well as to experience critical-thinking, creative, presentation and communication skills; enhanced motivation and understanding various levels of the subject domain. These skills would all be useful in their future undertakings. Our findings provide strong support and encouragement for Malaysian educators to incorporate multimedia technology and constructivist learning into the classrooms for the enhancement of teaching and learning.

Keywords
Constructivist learning, Multimedia projects, Student perceptions, Malaysian perspective, Multimedia authoring

Introduction
The infusion of Information Communication Technology (ICT) and, in particular, multimedia technology into education has created a significant impact on the instructional content development and the methods of communicating information to learners. This leads to the evolution of new concepts and innovative teaching techniques in the instruction-learning process, changing the way teachers teach and students learn. This changing landscape of education focuses on learning, rather than on teaching and pedagogy, curriculum and instruction. It seeks to create a generation of learners whose learning is defined as "the ability to retain, synthesize, and apply conceptually complex information in meaningful ways" (Lambert & McCombs, 1998). It also encourages better student learning through the learning objectives of project-based learning, or learning by doing (Schank, Berman and Macpherson, 1999), and to enable problem-solving, analysis, creativity and communication to take place in the classroom (Bates, 2000). In addition to these, multimedia technology has been shown to affect students’ motivation and self-esteem levels, as well as allow them to become creative and self-directed thinkers (Agnew, Kellerman & Meyer, 1996).

In Malaysia, the traditional mode of learning is used in many institutions of learning. However, in the context of introducing technology and multimedia in learning, the Malaysian Government is echoing this learner-centred learning initiative with a call for Malaysian institutions of higher learning to integrate ICT into their classrooms (Mat, 2000). In addition, there is a strong push by the Malaysian Government to develop creativity, communications skills, analytical and critical thinking, and problem-solving skills — skills that are significantly lacking in current graduates (Tan, 2000; Teo & Wong, 2000). This mismatch has prompted Malaysian educators to seek new ways to develop these appropriate skills and knowledge in students in order to meet the rising expectations of the knowledge society. World-wide research, as well as in Malaysia, has shown that using constructivism and multimedia technology has becoming increasingly important in teaching and learning in higher education in order to promote and enhance the teaching and learning process (Shaziti 2000; Wong, Kamariah & Tang, 2003) when set within authentic contexts (Herrington, Reeves, Oliver & Woo, 2004). This enables teachers to better transfer knowledge to their students in the classrooms (Wong et. al, 2003).

In the past, in the Malaysian region, many research studies tested ready-made CAI multimedia courseware on students (Norhayati & Siew, 2004; Mishra & Yadav, 2006), as well as the development of multimedia courseware by
educators for teaching purposes (Rozhan & Habibah, 2000; Suraya 2005). Research on constructivist learning environments and problem-solving activities have been confined to Internet-based learning environments where students’ perceptions of using Web-based learning materials were investigated (Low, Low & Koo, 2003; Lee & Tsai, 2005). The focus was to produce graduates with critical and creative thinking skills, oral and written presentation skills, and active learning skills. However, more focus should be placed on learning with ICT and multimedia technology within authentic, project-based, collaborative and multimedia-mediated learning environments.

This study was thus designed to investigate students’ perceptions in using a multimedia project that was embedded within a constructivist-based learning environment and to study the impact of using a multimedia project on students with little experience with working in a problem-solving environment. Students in a second-year multimedia course worked in groups to develop an interactive multimedia application using a high-end multimedia authoring tool, Macromedia Director. The students all played active roles in their project development processes, as well as in determining and reaching their learning goals. In this context, the focus was on the learning process rather than on the content: learning “how to learn” rather than on “how much is learned”. In this learning environment where technology was integral to their learning, the students develop critical thinking skills, problem-solving and team skills, experiential learning and inter-disciplinary knowledge (Oliver, 2000). This represented a move away from the traditional modes of tuition to one where the learners became active participants in the learning process (Oliver, 1998).

This study was intended to show that, through their perceptions and feedback on the project, the students would be able to reveal their ability to acquire skills integral to the demands of the workplace relevance: skills such as collaborative and teamwork, problem-solving, learning motivation, and critical thinking and understanding of a topic area.

The constructivist-based learning environment

Wilson (1995) defines a constructivist learning environment as “a place where learners may work together and support each other as they use a variety of tools and information resources in their pursuit of learning goals and problem-solving activities.” It is an environment that allows for learner-centred activities to take place where the teacher provides the students with experiences that allow them to develop problem-solving, critical-thinking and creative skills, and apply them in a meaningful manner. Here students work collaboratively to determine their learning goals, set their learning paths to reach their learning goals and monitor their own progress. The teacher is no longer perceived as the sole authority, but as the facilitator of learning, guiding and supporting learners in the process of constructing knowledge. Constructivism entails a strong belief that learning is a personal interpretation of the world; and that learners create interpretations of the world based on their past and current experiences and interpretations (Wilson, 1995; Jonassen 1994; Duffy & Cunningham, 1996; Jonassen & Henning, 1999). During constructivist learning, the emphasis is on learning and on the student-centric the learning environment. Students become active participants in their own learning processes and also learn to solve problems and work collaboratively.

Jonassen (1991, 1994, 1999) believes that constructivist learning environments can be designed to provide learners with meaningful, interesting and relevant problems to solve. Constructivist learning environments also demand a meaningful and authentic context for social and collaborative activities. Peers play an important role in encouraging learning, while supporting each other in a project-based curriculum as an innovative teaching and learning environment. In such an environment, students are engaged in their own learning process and become active learners. Jonassen (1999) proposed that the following components should be incorporated when designing a constructivist learning environment:

- **Conception of the problem.** A problem must first be conceived in order for the students to begin their learning development.
- **Interpretation.** Students interpret and develop solutions to their problems.
- **Information sources to support the understanding of the problem.** The learning environment provides the information that learners need to understand and solve problems. Appropriate information (text documents, graphics, sound, video and animation resources) can be accessed through the World Wide Web.
- **Cognitive tools.** Learners interpret and manipulate aspects of the problem through the World Wide Web as a cognitive tool.
• **Conversation and collaboration tools.** Learners form communities to negotiate and co-construct meaning for the problem through these tools. Students require a platform to share and exchange their ideas and create a community to solve their problem collaboratively and to facilitate and foster communities of learners. Examples are of collaborative tools are: email, chat, listserves, BBS (Bulletin Board Service), MUDs (Multi-user Dimensions), or MOOs (MUDs object-oriented).

Jonassen (1999) posited that an essential part of the learning problem is that it has to be interesting, engaging and appealing. It must also be authentic, personally relevant, challenging and interesting to learners, and provide a physical simulation of the real-world task environment. In addition, by collaborating with one another, students are exposed to multiple perspectives to their learning problems, enabling them to consider “…varying and discrepant points of view with which to consider the merits of his or her own mental models” (Oliver, 2000). In such an activity, student-centred learning can be cultivated because the students will engage in collaborative activities with their team members, as well as with the instructor, who acts as a facilitator and guide. Through working in groups, students will have to tap into their group’s skills and use a variety of activities to accomplish the overall project objectives. The group would be responsible for their goals and, thus, a collaborative learning experience can be gained.

Jonassen’s (1999) framework for designing a constructivist learning environment is adapted for this study in order to design the learning environment, and a multimedia project is embedded into this environment to form the core learning problem of the study, through which students will be able to become active participants in their learning processes and develop skills that would allow them to think critically, function as a team, develop collaborative abilities and deepen their understanding of their task, and improve their learning.

**The student learning process**

The study consists of of 53 students (N=53) in their second year of study. They comprised of students from the Faculty of Management, the Faculty of Information Technology and the Faculty of Engineering and were all enrolled for the Interactive Multimedia course. The objective of this course was to imbue students with multimedia project development skills over a 14-week trimester that culminated in an interactive group project that was multimedia and authored in Macromedia Director.

**The learning outcome**

The learning outcomes of the study were in the form of the students’ final interactive multimedia applications. These learning outcomes were evidence of their learning, as posited by Winnips and McLoughlin (2001). Students completed their final applications on time and presented them at the end of the trimester. These applications ranged from edutainment to the marketing of corporate applications, centred around the theme, “Malaysian Culture”. The applications demonstrated multiple solutions and perspectives of a learning problem. Figure 1 shows an example of a group’s application, an edutainment presentation of Hang Tuah, a legendary figure in Malaysian history. Figure 2 shows an example of the Malaysian martial arts, Silat. Both applications target children, but with different approaches.

In the Hang Tuah application (Figure 1), the group used a storytelling approach for their application, teaching children about the legendary figure through presentation and additional activities to support their learning process. After the initial splash screen (Image 1) of this application, the Menu page (Image 1) presents to the users the option to read the story (“Tell Me”, Image 3), or proceed to test their knowledge on the story via a quiz (“Test Me” and “Know Me”); a game (“Scramble Me”), a spelling contest (“Spell Me”) or a treasure hunt (“Find Me”) activity. Throughout the entire application, the interface is kept consistent with changes in background colour that depicts the various sections of the application that the user engages in.

Another group chose a more informational approach (Figure 2) to present their application, using an interactive room as their Menu (Image 1). Users can click to access various topics to access the various interactive areas. For example, if users click on the “General Info” image in the Menu, they will be presented with general information about the sport with embedded hyperlinks to keywords and definitions (Image 2). Similarly, users could view video clips on the performance of the sport (“Videos”) to get a more realistic presentation of activity. Users can also choose to test their knowledge on the sport by clicking on the “Quiz” component on the Menu and be asked questions on Silat
(Image 3). From these representations, it is clear that students were able to demonstrate multiple perspectives and solutions to their design problems in line with the constructivist pedagogy (Jonassen 1994; Herrington et al, 2004). By looking at the different approaches followed by the various groups in their applications, students were able to construct new knowledge on their own, resulting in an improved learning experience. Furthermore, the overall performance of the groups was good, as the students achieved mostly A’s and B’s for their projects.

Figure 1. A storytelling approach on Hang Tuah

The students followed a 6-step learning process (Jonassen 1999) to complete these assignments to develop the multimedia project.

1. **Group formation**

At the start of the project, students divided themselves into groups of 4-5 members and selected a group leader for each group. The decision of their group’s name, their leader and group tasks was taken collectively.
2. Problem identification

Groups discussed the problem and proposed multiple ideas for discussion. They brainstormed amongst and reached consensus for their final application.
3. **Project conceptualisation**

Group ideas were translated into concepts and storyboards. Members created sketches for the interfaces of their application’s screen and started to acquire the necessary media elements (e.g., graphics, sound, video, animation, and text) and created the visuals for their Director application.

4. **Project authoring**

The final application was authored. Each member of the team was responsible for at least one screen of the application. They proceeded to develop the application collectively and collaboratively.

5. **Presentation**

The groups made two presentations. The first was a work-in-progress presentation to give the class an overview of their work and solicit their comments. This provided them with preliminary experiences in presenting to an audience. They fine-tune their work, taking into account the comments from other students regarding their design, navigation, interactivity and overall presentation. The second presentation was more formal on completion of the applications. The students presented their work, discussed modifications made on their first presentation, and submitted their application to the lecturer.

6. **Reflection**

Students reflected on: their own individual progress, their relationships with the group, the team leader, the problems they encountered individually and collectively, how they overcame their problems, as well as on the comments made by the other groups on their presentations. These reflections were incorporated in their final group report and submitted to the lecturer.

The conceptual framework of their multimedia-mediated development process is illustrated below in Figure 3.
Methodology

At the end of the project, students were given a thirty-item survey, to measure their attitudes towards developing a multimedia project in their learning process, to fill out. The survey was adapted from Diamond (1998) and Irani (1998). The items measured on a 5-point Likert scale, and with 1 = Strongly Disagree, 2 = Disagree, 3 = Undecided, 4 = Agree and 5 = Strongly Agree. The survey aimed to gauge their motivation, critical-thinking and creativity skills, teamwork, presentation and communication skills, and overall attitudes towards learning with multimedia and developing a multimedia project.

Results

The collected data were analysed with SPSS version 11.5, and yielded a Cronbach Alpha coefficient of 0.9106, which satisfies the requirement of survey reliability (Lim, Khine, Hew, Wong, Shanti & Lim, 2003). Results showed that students had very positive attitudes towards the project and their use of multimedia technology in this learning environment. From their responses from the survey, open-ended questions and the interviews, students’ perceptions toward this study can be categorized into the following areas:

- Learning motivation
- Increased understanding and learning of subject domain
- Skills and real-world relevance
- Teamwork and collaboration

Tables 1-4 present the results of the questionnaire in the 4 areas, with the overall means (m) ranked for each items in these 4 areas, the percentage of responses (p) by students according to the 5-point Likert scale (i.e., SA = Strongly Agree, A= Agree, U = Undecided, D = Disagree, and SDA = Strongly Disagree) as well as their respective standard deviations (SD). In addition, student comments and feedback were also obtained from their journal reports and from the survey’s open-ended question section. Students provided feedback on their learning process as a result of doing this interactive multimedia project, which are also show in Tables 1-4 accordingly.

### Table 1. Results for students’ learning motivation

<table>
<thead>
<tr>
<th>Learning motivation Items in the survey</th>
<th>mean</th>
<th>SA (p)</th>
<th>A (p)</th>
<th>U (p)</th>
<th>D (p)</th>
<th>SDA (p)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Motivated to do project</td>
<td>4.25</td>
<td>30.2</td>
<td>64.2</td>
<td>5.7</td>
<td>0</td>
<td>0</td>
<td>0.55</td>
</tr>
<tr>
<td>2. Project made me want to do my best</td>
<td>4.25</td>
<td>32.1</td>
<td>60.4</td>
<td>7.5</td>
<td>0</td>
<td>0</td>
<td>0.59</td>
</tr>
<tr>
<td>3. I found the project to be challenging yet stimulating to do</td>
<td>4.19</td>
<td>30.2</td>
<td>58.5</td>
<td>11.3</td>
<td>0</td>
<td>0</td>
<td>0.62</td>
</tr>
<tr>
<td>4. I enjoyed working on a project like this</td>
<td>4.08</td>
<td>24.5</td>
<td>62.3</td>
<td>11.3</td>
<td>0</td>
<td>1.9</td>
<td>0.73</td>
</tr>
<tr>
<td>5. This course has given me confidence in my newly acquired skills</td>
<td>3.91</td>
<td>18.9</td>
<td>56.6</td>
<td>20.8</td>
<td>3.8</td>
<td>0</td>
<td>0.74</td>
</tr>
<tr>
<td>6. I was able to maintain contact with my lecturer</td>
<td>3.74</td>
<td>7.5</td>
<td>62.3</td>
<td>26.4</td>
<td>3.8</td>
<td>0</td>
<td>0.65</td>
</tr>
<tr>
<td>7. I am very satisfied with my contribution to the project</td>
<td>3.68</td>
<td>17.0</td>
<td>41.5</td>
<td>34.0</td>
<td>7.5</td>
<td>0</td>
<td>0.85</td>
</tr>
</tbody>
</table>

N=53

Students’ comments on Learning Motivation:

1. “Very motivated. This subject is very fun. I get the chance to come out with my own ideas and creation on an application.”
2. “I feel motivated from that cause I can learn some new skills that I can’t learn from my faculty”.
3. “I am so happy and motivated when do this project.”

Learning motivation

From Table 2, it can be seen that students’ motivation levels were high and their interest in doing the project was very much enhanced, as the three highest ranked items in the survey were motivation items. In particular, students felt highly motivated in completing their projects, and reported that the project encouraged them to work hard. Most students (94.3%) reported favourable motivation levels (m=4.25) and 92.5% of students reporting that the project made them do their best (m=4.25), making these two items the highest ranking items in the survey. They also indicated that although the project was challenging, it provided much stimulation in their learning process -- 88.7%
reporting in favour of the item (m = 4.19). Almost everyone in the class reported that they enjoyed working on such a project (m = 4.08, p = 98.1), whereas 75.5% of students reported that doing the project has given them confidence in the skills which they had acquired in the process (m = 3.91), and 69.8% were able to maintain contact with their lecturer during the project development period (m = 3.74). And as such, 92.5% of students reported that they were, on the overall, satisfied with their contribution to the project (m = 3.66).

Support for this favourable report can also be seen in their comments and feedback. Students reported that they enjoyed doing the project as it made them feel very motivated and satisfied with their contributions. Students reported in their journals (Table 1) that they were motivated because the project was something that was relevant to them and elicited their curiosity. They also expressed excitement with the innovative way information was presented as they used multimedia technologies to create something that was digital and interactive.

**Increased understanding and learning of subject domain**

This construct was incorporated to gauge students’ perceptions on whether doing the project resulted in an increased level of understanding for the subject matter and on their own learning process (Table 2).

Table 2. Results for students’ understanding of the subject domain

<table>
<thead>
<tr>
<th>Increased understanding Items in the survey</th>
<th>mean</th>
<th>SA (%)</th>
<th>A (%)</th>
<th>U (%)</th>
<th>D (%)</th>
<th>SDA (%)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The project enhanced my learning of interactive multimedia</td>
<td>4.13</td>
<td>26.4</td>
<td>62.3</td>
<td>9.4</td>
<td>1.9</td>
<td>0</td>
<td>0.65</td>
</tr>
<tr>
<td>2. The project increased my understanding on how to manage and develop an interactive project</td>
<td>4.13</td>
<td>24.5</td>
<td>67.9</td>
<td>3.8</td>
<td>3.8</td>
<td>0</td>
<td>0.65</td>
</tr>
<tr>
<td>3. I am now a better learner</td>
<td>4.02</td>
<td>20.8</td>
<td>64.2</td>
<td>11.3</td>
<td>0</td>
<td>0</td>
<td>0.69</td>
</tr>
</tbody>
</table>

N=53

Students’ comments on Increased Understanding:

1. “Before doing this project, I do not know much about Silat...but after working on this topic for past four months, I learnt Silat in detail. Now I am able to explain about Silat to my other friends”.
2. “I believe my understanding towards interactive multimedia has vastly improved after undergoing the project given to us.”
3. “It helps me more understand the subjects in multimedia.”

Results in Table 2 showed that students were favourable toward this construct. In the survey, students reported that in terms of their overall understanding of the project’s objective, 88.7% of students reported that the project enhanced their learning of interactive multimedia (m = 4.13) and that the project increased their understanding of how to manage and develop an interactive multimedia application (m = 4.13, p = 92.5). They also reported that they now have become better learners, (m=4.02, p=84.9). Again, these perceptions are well supported by their comments and feedback, which showed that students did perceive themselves to have increased understanding of the subject matter from the project. Many reported that they now understood what it meant to develop a project and that hands-on experience made them understand interactive multimedia development as well as on their chosen topic.

**Skills and real-world relevance**

This construct was to gauge students’ attitudes and perceptions on their acquisition and experience with skills such as critical-thinking and creativity skills, teamwork and group skills, communication and presentation skills, multimedia technology skills, and the ability to properly apply them (Table 3).

Results from their survey shed more light and support for this construct, as, in terms of acquired skills, 88.7% of students reported that they were now able to apply their skills in a more effective manner on future projects (m = 4.15), making it the highest ranking item in this category. They also reported being able to analyse, synthesise, and evaluate information (m = 4.08, p = 88.7). Critical-thinking skills were also enhanced, as 79.2% of students reported that they were now able to think critically about developing interactive multimedia applications, (m=3.94), as were presentation and communications skills, as 67.9% of students reported that the project allowed them to improve their
presentation skills (m = 3.72), and presenting their project well using multimedia (m=3.89, p=75.5). Students also showed increased perception of the relationship between their work and the work in real-life situations. 84.9% of students reported that after completing their project, they were now able to see the relevance between the project’s task and the course, with real-world situations (m = 4.04), allowing them to develop skills needed in the real-world (m = 3.98, p = 81.1).

Table 3. Results for students’ skills acquisition

<table>
<thead>
<tr>
<th>Skills and real-world relevance</th>
<th>mean</th>
<th>SA (%)</th>
<th>A (%)</th>
<th>U (%)</th>
<th>D (%)</th>
<th>SDA (%)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I am now able to apply my skills in a more effective manner on future projects</td>
<td>4.15</td>
<td>26.4</td>
<td>62.3</td>
<td>11.3</td>
<td>0</td>
<td>0</td>
<td>0.60</td>
</tr>
<tr>
<td>2. The project allowed me to analyse, synthesise and evaluate information</td>
<td>4.08</td>
<td>18.9</td>
<td>69.8</td>
<td>11.3</td>
<td>0</td>
<td>0</td>
<td>0.55</td>
</tr>
<tr>
<td>3. I saw the relevance between the course and real world situations</td>
<td>4.04</td>
<td>22.6</td>
<td>62.3</td>
<td>11.3</td>
<td>3.8</td>
<td>0</td>
<td>0.71</td>
</tr>
<tr>
<td>4. This project allows me to develop skills needed in the real-world</td>
<td>3.98</td>
<td>18.9</td>
<td>62.3</td>
<td>17.0</td>
<td>1.9</td>
<td>0</td>
<td>0.66</td>
</tr>
<tr>
<td>5. I am now able to think critically about developing interactive applications</td>
<td>3.94</td>
<td>17.0</td>
<td>62.3</td>
<td>18.9</td>
<td>1.9</td>
<td>0</td>
<td>0.66</td>
</tr>
<tr>
<td>6. We were able to present our project well using multimedia</td>
<td>3.89</td>
<td>18.9</td>
<td>56.6</td>
<td>20.8</td>
<td>1.9</td>
<td>1.9</td>
<td>0.80</td>
</tr>
<tr>
<td>7. The project allowed me to develop and improve my presentation skills</td>
<td>3.72</td>
<td>7.5</td>
<td>60.4</td>
<td>28.3</td>
<td>3.8</td>
<td>0</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Students’ comments on Skills and Real-World Relevance:
1. “...from the experience of working with my group members, I am prepared to face different people I might meet in the near future.”
2. “I get to know more about how to develop a good project in the future as if I have the chances to take on the multimedia task.”
3. “I learnt a lot of skills and knowledge...which enable me to understand and may apply to my future as well.”

Students’ comments and feedback further provided support to the results and showed that they were able to acquire and experience these skills and that they saw the relevance of these skills with that needed in the real-world. They also reported that the acquisition and experience of these skills increased their confidence levels in solving project- and team-related problems in the future (Table 3).

Teamwork and collaboration

The final construct in this study was to gauge students’ perception and attitudes towards working in a team and collaborating with their group members (Table 4).

This teamwork construct was well supported in the survey results as many students reported favourably on many of the teamwork items. In particular, 77.4% of students reported being able to get to know their team members well from doing this project (m = 3.94), 73.6% reported that they enjoyed working in a team (m = 3.83), 66% of students reported that they were able to interact well with their teammates (m = 3.66), 67.9% of students reported that they were able to contribute creative ideas in the group (m=3.75), that their group helped them to do their best in the project (m =3.72, p = 62.3%) and they were able to work well together to present their project (m=3.72, p=64.2). Problem-solving and conflict management within the team was also an item with was positively reported, as 73.6% of students reported that their group was able to solve their problems and conflicts in a positive manner (m = 3.77), and were supportive of each member’s problems and tried to help resolve them (m = 3.75, p = 67.9).
### Table 4. Results for students’ teamwork and collaboration skills

<table>
<thead>
<tr>
<th>Teamwork and collaboration Items in the survey</th>
<th>mean</th>
<th>SA (%)</th>
<th>A (%)</th>
<th>U (%)</th>
<th>D (%)</th>
<th>SDA (%)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Team collaboration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. I got to know my group members well</td>
<td>3.94</td>
<td>24.5</td>
<td>52.8</td>
<td>15.1</td>
<td>7.5</td>
<td>0</td>
<td>0.84</td>
</tr>
<tr>
<td>2. I enjoy working in a team</td>
<td>3.83</td>
<td>26.4</td>
<td>47.2</td>
<td>15.1</td>
<td>5.7</td>
<td>5.7</td>
<td>1.07</td>
</tr>
<tr>
<td>3. My group was able to solve our problems and conflicts in a positive manner</td>
<td>3.77</td>
<td>15.1</td>
<td>58.5</td>
<td>18.9</td>
<td>3.8</td>
<td>3.8</td>
<td>0.89</td>
</tr>
<tr>
<td>4. We were able to contribute our creative ideas in the group</td>
<td>3.75</td>
<td>24.5</td>
<td>43.4</td>
<td>17.0</td>
<td>13.2</td>
<td>1.9</td>
<td>1.04</td>
</tr>
<tr>
<td>5. My group was supportive of member's problems and helped resolved them</td>
<td>3.75</td>
<td>20.8</td>
<td>47.2</td>
<td>20.8</td>
<td>9.4</td>
<td>1.9</td>
<td>0.96</td>
</tr>
<tr>
<td>6. My group helped me do my best in the project</td>
<td>3.72</td>
<td>28.3</td>
<td>34.0</td>
<td>22.6</td>
<td>11.3</td>
<td>3.8</td>
<td>1.12</td>
</tr>
<tr>
<td>7. My group worked well together to present our project</td>
<td>3.72</td>
<td>20.8</td>
<td>43.4</td>
<td>22.6</td>
<td>13.2</td>
<td>0</td>
<td>0.95</td>
</tr>
<tr>
<td>8. I was able to interact well with my classmates</td>
<td>3.66</td>
<td>11.3</td>
<td>54.7</td>
<td>24.5</td>
<td>7.5</td>
<td>1.9</td>
<td>0.85</td>
</tr>
<tr>
<td><strong>B. Team dynamics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Our group encouraged positive contributions from each member</td>
<td>3.60</td>
<td>18.9</td>
<td>45.3</td>
<td>20.8</td>
<td>7.5</td>
<td>7.5</td>
<td>1.12</td>
</tr>
<tr>
<td>10. We were able to complete our tasks on time</td>
<td>3.51</td>
<td>13.2</td>
<td>41.5</td>
<td>32.1</td>
<td>9.4</td>
<td>3.8</td>
<td>0.97</td>
</tr>
<tr>
<td>11. There was a lot of unity in my group</td>
<td>3.51</td>
<td>17.0</td>
<td>37.7</td>
<td>30.2</td>
<td>9.4</td>
<td>5.7</td>
<td>1.07</td>
</tr>
<tr>
<td>12. We were able to organise our work effectively</td>
<td>3.36</td>
<td>9.4</td>
<td>35.8</td>
<td>41.5</td>
<td>7.5</td>
<td>5.7</td>
<td>0.96</td>
</tr>
<tr>
<td>13. My group was able to make and follow a set agenda</td>
<td>3.34</td>
<td>7.5</td>
<td>35.8</td>
<td>39.6</td>
<td>17.0</td>
<td>0</td>
<td>0.85</td>
</tr>
</tbody>
</table>

**N=53**

**Students’ comments on Teamwork and Collaboration**

1. “We can share out our ideas together and discuss everything together.”
2. “Great! My group member is really good as we came from different faculty….we really work as a team especially when I face problems, my group member will always give me a hand.”
3. “Can’t be denying that, sometimes we do have argument on certain idea, but we will find a positive manner to settle it.”
4. “I have learnt how to work in group, improve myself in working with someone new...”

Interestingly, 64.2% of students reported that their group encouraged positive contributions from each member (m=3.60). Only 54.7% reported being able to complete their tasks on time (m=3.51), and 54.7% reporting that there was unity in the group (m=3.51). In addition to that, only 45.3% of students reported being able to organise their work effectively (m=3.36) and only 43.4% reported that their group was able to make and follow a set agenda (m=3.34). This clearly shows that working as a group was a challenging task for these students and many were experiencing several obstacles to group management and group collaboration. This was also reiterated when asked for their comments. Many students reported that the biggest challenges to them were in having to solve the conflicting ideas of the groups, and the lack of time due to conflicting schedules. These problems did affect their motivation levels at times, but they managed to solve these problems through the ability to communicate using Web and telecommunication tools (that helped solve their communication problems), and by increasing their group efforts to solve their conflicts in ideas, as reported in the following comments:

1. “The most serious problem here is idea conflict among members... In order to solve this problem, we have record down each member unsatisfied problem in a paper and solve it one by one.”
2. “...we also spend some time to figure out each member ideas and vote for the best idea that most people prefer.”
3. “...we had faced problems especially communication problem…
4. “We had problems such as dividing the workload among us since we hardly knew each other that time and most of us lived outside the campus premise…
5. As it reaches the middle of the current progress of the project, all members started to tense out and the motivation level within myself began dropping but as when it goes on doing, each one of us give and help each other to give encouragement and this makes everyone to boost up their confidence and courage to go on complete all.
6. Can’t be denying that, sometimes we do have argument on certain idea, but we will find a positive manner to settle it.

7. The project was a good start at first but as it progress, we had some problems lie we stuck onto how is it going to implement to do in the application … For this, we might need some advices and lesson from other people to help us in completing through this project to be handled up on time.

8. It has been a lot of difficulty when we have trouble to find some information about the content and have one of our group members did not cooperate well but we manage to solve it wisely when we define, analyze, identify the problem and we decide what is the best solution.

From these comments, it can be seen that there were some serious hardships and challenges to the students with respect to their group work, which affected their motivation levels during the project’s development. However, by being able to solve their group problems, students were able to comment favourably on their overall attitudes toward teamwork at the end of the project development. Students reported that teamwork and group collaboration were indeed important to the success of the project, that they enjoyed working in a team and found the project a good opportunity to enhance their creativity and to share their workload in the project, and were happy that they did not give up in the middle of the project.

Discussion

The results show strong support for using a multimedia project in a constructivist-based learning environment. From the survey and their feedback, students’ demonstrated positive attitudes and perceptions to developing a multimedia project within this learning environment. They were able to reveal that they had acquired several key constructivist learning skills through doing the project, which would enable them to become better skilled workers in the IT industry in Malaysia.

Specifically, the development of this multimedia project within this learning environment allowed students to experience the following:

1. In this constructivist-based learning environment using multimedia, it can be seen that students experienced high levels of motivation and self-esteem when doing the multimedia project as shown in their comments and survey results, and is in line with Reeves’ (1998) perspective that, “multimedia can stimulate more than one sense at a time, and in doing so, may be more attention-getting and attention-holding.” More importantly, it enhanced their confidence levels in their newly acquired skills, knowing that they can use the same skills in their future undertakings.

2. Students showed increased understanding of the topic and being able to see the relevance of the project to real-life situations. This falls within the realm of authentic learning (Herrington et. al, 2004).

3. Students learnt that teamwork and collaboration became important factors to the successful completion of the project. As they had experienced difficulties in some of their group activities, students reported having had to learn about the importance of teamwork and cooperation, and to develop group and leadership skills. Although group management and teamwork were positively reported on the overall, students reported that they were still challenging skills to acquire and balance. As such this study was able to demonstrate that a technology-enhanced constructivist learning environment, as suggested by Jonassen (1994, 1999), and specifically through the incorporation of an interactive multimedia project, can be successful in allowing students to engage in meaningful, relevant problem-solving activities, experience active learning, and be more engaged in their learning process. This is in line with Herrington et. al’s (2004) position and with Winnips and McLoughlin’s (2001) position that the creation of applications is proof of student learning. As such, this constructivist learning environment has been successful in supporting project-based learning as an alternative to traditional classroom curriculum.

4. Multimedia technology was successfully integrated in this constructivist learning environment as an enabler and catalyst to support the students in successfully using the tools to demonstrate their creativity and problem-solving skills. By incorporating multimedia into this constructivist learning environment via their multimedia project, students were able to design using multimedia and experience critical-thinking, creative, presentation and communication skills, as well as enhancing their motivation and understanding levels. Using various combinations of media elements to illustrate their messages in the application gave students the opportunity to think critically about their messages and the flexibility to present them. Using Web communication tools also
gave them the opportunity to solve group meeting problems and to conduct any meetings or discussions, and any exchange of ideas amongst themselves whenever they had scheduling conflicts for face-to-face meetings.

5. Results of this research study provide good evidence for using multimedia technology and a project-based learning approach within a constructivist learning environment, as adapted from Jonassen’s (1999) framework, for Malaysian educators in tertiary education, who want more flexible options in their classroom teaching methods, and to inculcate their students with better real-world skills. More research can be done, however, in investigating further the role of multimedia in motivating students, and the group dynamics within such a learning environment.

Conclusion

The study was thus successful in providing students with the experience in problem-solving, critical-thinking and creativity skills, communication and reflection, and in improving their overall understanding of the project’s objective, as students reported satisfaction with their contributions. They also became active participants in their learning process. The project also allowed students to become more independent in their thinking and as a whole, enabled them to improve their learning process.

Adapting Jonassen’s (1999) proposition for building a constructivist learning environment using technology in this classroom environment created a learning environment where students were able to solve a problem, via the interactive multimedia project, that was authentic and relevant to them, and allowed them to collaborate and work together as a team to complete the project, and to claim responsibility and ownership of their development process. This multimedia-mediated learning environment allowed students to experience a constructivist-based approach in their learning process where they became active participants and constructed their own knowledge. Although the students reported that there were some challenges in teamwork and collaboration, it did not deter them from having an overall positive attitude towards the project. Therefore, this research study has shown that incorporating multimedia technology into a constructivist learning environment can lead to innovative teaching and learning methods for the improvement of classroom learning.

References


Comparing Learning Performance of Students Using Algorithm Visualizations Collaboratively on Different Engagement Levels

Mikko-Jussi Laakso, Niko Myller and Ari Korhonen

ABSTRACT
In this paper, two emerging learning and teaching methods have been studied: collaboration in concert with algorithm visualization. When visualizations have been employed in collaborative learning, collaboration introduces new challenges for the visualization tools. In addition, new theories are needed to guide the development and research of the visualization tools for collaborative learning. We present an empirical study, in which learning materials containing visualizations on different Extended Engagement Taxonomy levels were compared, when students were collaboratively learning concepts related to binary heap. In addition, the students’ activities during the controlled experimental study were also recorded utilizing a screen capturing software. Pre- and post-tests were used as the test instruments in the experiment. No statistically significant differences were found in the post-test between the randomized groups. However, screen capturing and voice recording revealed that despite the randomization and instructions given to the students, not all of the students performed on the engagement level, to which they were assigned. By regrouping the students based on the monitored behavior, statistically significant differences were found in the total and pair average of the post-test scores. This confirms some of the hypothesis presented in the (Extended) Engagement Taxonomy.

Keywords
Algorithm visualization, Algorithm simulation, collaborative learning, Engagement taxonomy

Introduction
Since its introduction, it has been hoped that Algorithm Visualization (AV) would solve problems related to learning of data structures and algorithms. However, empirical evaluations have yielded mixed results when determining the usefulness of such visualizations as teaching and learning aids over traditional methods (see the meta-analysis of the research on AV by Hundhausen et al. (2002)). Thus, researchers have sought explanations for the mixed results as well as better grounds to justify the use of visualizations in teaching. Hundhausen et al. (2002) concluded that the activities performed by the students are more important than the content of the visualization. This has led to the analysis of different engagement levels Naps et al. (2002) by ITiCSE Working Group that proposed Engagement Taxonomy (ET) to describe the various types of activities that students perform with visualizations and their effect on learning and Myller et al. (in press) have developed it further into Extended Engagement Taxonomy (EET).

Collaboration has become accepted and popular in Computer Science education. A good example is the benefits of pair programming (Nagappan et al., 2003; Williams et al., 2000; McDowell et al., 2003). Whilst visualizations are employed in collaborative learning, collaboration introduces new challenges for the visualization tools. For example, the exchange of experiences and ideas, and coordination of the joint work are needed when students are not working individually anymore (Suthers and Hundhausen, 2003). Furthermore, visualizations can provide a shared external memory that can initiate negotiations of meanings and act as a reference point when ideas are explained or misunderstandings are resolved (Suthers and Hundhausen, 2003). This implies that also new theories are needed to guide the development and research of the visualization tools for collaborative learning.

In this paper, the applicability of EET in collaborative use of visualizations has been studied. We test the impact of EET levels on the performance when visualizations are used in collaboration. We present an empirical study, in which learning materials containing visualizations on different EET levels were compared when student pairs were collaboratively learning concepts related to binary heap. The pairs had a mutual task to read through a tutorial including visualizations and answer questions related to the topic. Although, statistically significant differences were
not detected in a previous study, the results indicated that the engagement level of the visualizations has an effect on
the performance when students are working in pairs (Myller et al., 2007). Thus, we replicated that study in a different
institution, and improved the settings in such a way that the detection of the statistically significant differences would
be possible. In this paper, we report the results from the replication study conducted at the Helsinki University of
Technology in which two groups of students were randomized to the computer lab sessions. Each session was
randomly assigned to an EET level, either changing or controlled viewing (in the rest of the paper this can be also
shortened to viewing when we are discussing about the groups), with the limitation that parallel sessions belonged to
different conditions.

During the analysis of the screen and voice recordings collected in the study, it was detected that despite the
randomization and instructions given to the students, not all of the students performed their learning on the expected
EET level. This meant that although the tool allowed students to learn on a higher EET level, some of the students
choose not to do so, but worked on a lower engagement level. Fortunately, the screen capturing and voice recording
done during the students’ learning process provided us a tool for noticing this and taking it into account in the
analysis. Thus, in addition to the results from the study, we learned an important methodological lesson as well.
Screen capturing and voice recording should be a standard procedure, because otherwise we cannot know for sure if
the participants really do what we expect them to do.

In Chapter 2, we describe the relevant literature related to the engagement taxonomy and similar theories. In
addition, we give an overview of the learning tool used in the experiments. Chapter 3 describes the research setting,
i.e., the used pre- and post-tests, subjects, materials, and procedures. In Chapter 4, we report on the results. Finally,
in Chapters 5 and 6, we make conclusions and highlight some future directions.

**Previous Research**

**Visualizations and Engagement**

As an attempt to describe the mixed results of previous research in AV usage (cf. (Hundhausen et al., 2002)) in
learning and teaching of algorithms and data structures, Engagement Taxonomy (ET) was introduced by Naps et al.
(2002). The central idea of the taxonomy is that the higher the engagement between the learner and the visualization,
the higher the positive effects on learning outcomes. ET consists of six levels of engagement between the user and
the visualization:

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No viewing</td>
<td>There is no visualization to be viewed.</td>
</tr>
<tr>
<td>Viewing</td>
<td>The visualization is only looked at without any interaction.</td>
</tr>
<tr>
<td>Responding</td>
<td>Visualization is accompanied with questions, which are related to the content of the visualization.</td>
</tr>
<tr>
<td>Changing</td>
<td>Modification of the visualization is allowed, for example, by varying the input data set or algorithm simulation.</td>
</tr>
<tr>
<td>Constructing</td>
<td>Visualization of program or algorithm is created.</td>
</tr>
<tr>
<td>Presenting</td>
<td>Visualizations are presented to others for feedback and discussions.</td>
</tr>
</tbody>
</table>

ET has been used in the development of AV tools and several studies have utilized the framework and provided
further support for it (see, e.g., Grissom et al. (2003); Naps and Grissom (2002)). However, the time to study the
materials on different ET levels has commonly been an uncontrolled variable in the studies, meaning that students
have had freedom to use as little or as much time as they wanted to. Thus, those students who have been studying
with visualizations that are on the higher ET level have spent more time on the task. This, in turn, makes it
questionable if the reason for better performance in the post-test is due to the additional time spent on studying or the
higher ET level of the materials. In the experiment, which is presented in this paper, we controlled the time so that all
the students needed to spend exactly the same amount of time on learning the topic.

There are also other studies which have shown that visualizations improve learning, without actually utilizing the ET
framework in the design of the study (Ben-Bassat Levy et al., 2003). In addition to this, research in educational
psychology and multimedia learning had also had similar results (Evans and Gibbons, 2006).
Myller et al. (in press) have proposed an extension to the ET called the Extended Engagement Taxonomy (EET). The idea of this extension is to let the designers and researchers of visualizations to use finer granularity of engagement levels in their tools and experimental designs. They provide the following engagement levels to be used together with the original ones: controlled viewing, providing input, modification, and reviewing. In this study, we will utilize the controlled viewing level in order to make a difference between the visualizations that can only be viewed by the student (EET level: viewing, e.g., static visualizations or animations with only a playing option) compared to those which can be controlled (EET level: controlled viewing, e.g., animations with VCR-like controls in order to step and play the animation both forwards and backwards).

Visualizations and Collaboration

From a more general perspective, there are studies that analyze the use of visualizations in collaboration. For instance, Suthers and Hundhausen (2003) have performed research in the area of scientific inquiry. They compared the effects of different representations (i.e., matrix, graph, and text) when students were collecting and analyzing data, hypotheses and their evidential relations. Their research showed that the form of the visualization and what kinds of interactions it drives have an effect on the collaboration process by making certain data and their relations more explicit or implicit.

Roschelle (1996) studied pairs of students using the learning environment of Newtonian physics and analyzed their learning outcomes as well as the process that led to those outcomes. During the study, it was recognized that learning tools and especially visualizations used in collaboration should focus more on supporting communication rather than presenting the underlying model as accurately as possible. Furthermore, Roschelle (1996) tells as the last lesson in his paper that, “one should design activities, which actively engage students in doing and encountering meaningful experiential feedback as a consequence of their actions”. Scaife and Rogers (1996) also identified the analysis of the interactions between external presentation and its users as a key research area for the future. All these points of view seem to support the applicability of ET/EET in the context of collaborative learning.

Although several AV tools have been developed and empirical studies carried out, the collaborative use of AV tools is researched very little. Myller et al. (in press) have studied the applicability of EET to describe differences in the learning process when visualizations are used during collaborative learning. They pointed out that when students were using visualizations on lower EET levels the interaction/engagement between students also dropped, meaning that students communicated and collaborated more when they were using materials on higher EET levels.

The work of Hundhausen (2002) is related to the collaborative aspects of AV construction and presentation. This work led into the development of a visualization tool, ALVIS, which supports construction and presentation of AVs in small groups (Hundhausen and Brown, 2008). Their results also indicate that ET is applicable in the context of collaborative learning, although it is not directly tested. Furthermore, Hundhausen (2005) has proposed a communicative dimensions framework in order to analyze the aspects of visualizations that affect communication between end-users. Hübscher-Younger and Narayanan (2003) developed a web-based system that allows students to post their own algorithm representations (e.g., text, pictures, animation, or multimedia) and discuss them on the web. The research concluded that the students who actively participated in this activity achieved higher grades than the passive students who might have only viewed and commented on others’ presentations.

Other Algorithm Visualization Studies on Heap Data Structures

Stasko et al. (1993) utilized algorithm animations focusing on a pairing heap that was implemented as a binary tree. The results were disappointing: the animation group outperformed the control group but the differences were not high even on absolute scale, and the differences were not statistically significant. Moreover, they noted that using animations did not grant obvious learning benefits and they believe that algorithm animations benefit advanced students more than “novice students”.

In 1996, Byrne et al. (1996) conducted algorithm animation research on binomial heap. The results were not statistical significant, either, and their findings supported the view that the benefits of animations are not that obvious, and careful task analysis is essential to determine in which situations animation can be helpful. Also Kehoe
et al. (2001) studied the learning of binomial heap through animations in open lab sessions. They hypothesized that animations make complex algorithms more accessible and less intimidating and enhance students’ motivation, interaction and learning. Their study, however, was inconclusive (they made hypotheses), and further empirical studies were suggested.

There are some differences between these studies and ours. Our students were novices with little or no previous knowledge on the topic, but they were not novices in using the visualization tool but had previous knowledge on how to use the tool and how to make sense of its visualization. However, students needed to study in our experiment concepts related to binary heap, which might be easier to understand and more accessible for novices compared to the pairing heap or the binomial heap. Furthermore, we used fixed time limits for the learning session meaning that all students needed to use exactly the same time to learn the topic, and we monitored their learning process in order to detect how they were learning.

TRAKLA2 Overview

TRAKLA2 is a practicing environment for visual algorithm simulation exercises (Korhonen et al., 2004) that can be assessed automatically. The system distributes individually tailored tracing exercises to students and provides feedback about students’ solutions automatically. In visual algorithm simulation exercises, a student directly manipulates the visual representation of the underlying data structures (i.e., a student acts on the EET level changing). Thus, the student manipulates real data structures through GUI operations with the purpose of performing the same changes on the data structures the actual algorithm would do. An answer to an exercise is a sequence of discrete states of data structures, and the task is to perform the correct operations that will cause the transitions between each of the two consecutive states.

Each TRAKLA2 exercise page consists of a description of the exercise with links to other pages that introduce the theory and examples of the algorithm in question, instructions on how to interact with the GUI, code window, and an interactive Java applet. The current exercise set consists of over 40 assignments on basic data structures, sorting algorithms, search trees, hashing methods, and graph algorithms.

![Figure 1: TRAKLA2 exercise page. The student acts in EET level changing by solving the exercise in terms of swapping the data elements in the data structure(s)](image)

Let us consider the exercise in Figure 1. The student is supposed to manipulate the visual representation(s) of the Binary Heap data structure by invoking context-sensitive drag-and-drop operations. The idea is to simulate the
linear time \texttt{BuildHeap} algorithm. The manipulation can be done in either of the representations shown in the figure (i.e. the array or the binary tree representation). A key can be sifted up in terms of swap operations with its parent until the heap property is satisfied (the key at each node is smaller than or equal to the keys of its children). A single swap operation is performed by dragging and dropping a key in the heap on top of another key.

An exercise applet is initialized with \textit{randomized input data}. The \texttt{BuildHeap} exercise, for example, is initialized with 15 numeric keys that correspond to the priority values. The student can reset the exercise by pressing the \texttt{Reset} button at any time. As a result, the exercise is reinitialized with new random keys. When attempting to solve the exercise, the student can review the answer step by step using the Animator panel. Moreover, the student can submit the answer in which case the answer is assessed and immediate feedback is delivered. The feedback reports the number of correct steps out of the total number of steps in the exercise. This kind of automatic assessment is possible due to the fact that, again, the student is manipulating real data structures through the GUI. Thus, it is possible to implement the same algorithm the student is simulating, and execute it so that the algorithm manipulates the same data structures, but different instances, as the student just did. The assessment is based on comparison between these two different instances of data structures with each other.

An exercise can be submitted an unlimited number of times. However, a solution for a single instance of an exercise with certain input data can be submitted only once. In order to resubmit a solution to the exercise, the student has to reset the exercise and start over with new randomized input data. A student can also review a Model answer for each attempt. It is represented in a separate window as an algorithm animation accompanied with a pseudo code animation so that the execution of the algorithm is visualized step by step. The states of the model solution can be browsed back and forth using a similar animator panel as in the exercise. For obvious reasons — after opening the model solution — the student cannot submit a solution until the exercise has been reset and resolved with new random data.

\textsc{Previous Studies on TRAKLA2}

In 2001, the first intervention study Korhonen et al. (2002) with three randomized groups A, B, and C ($N_A = 372$, $N_B = 77$, $N_C = 101$) was performed. Students’ behavior was monitored over the second year course in data structures and algorithms (DSA) lasting twelve weeks. The examination results of students using the TRAKLA learning environment (predecessor of TRAKLA2) were compared with those in the traditional classroom sessions. The results showed that, if the exercises are the same, there is no significant difference in the final examination results between students exercising on the web (group A) or in the classroom (group B). In addition, the commitment to the course (low drop-out rates), is almost equal in both versions of the course. However, if the exercises are more challenging (group C), there is a significant difference in the examination results, but the drop-out rate is significantly higher as well.

Laakso et al. (2005a) reported on another whole semester study, in which TRAKLA2 was introduced at the University of Turku. The students’ learning results were compared between students, who used or did not use TRAKLA2, during a course on DSA. In addition, a survey-data ($N = 100$) was collected on the changes in students’ attitudes towards web-based learning environments. The results showed that TRAKLA2 considerably increased the positive attitudes towards web-based learning. According to students’ self-evaluations, the best learning results were achieved by combining traditional and web-based exercises. In addition, the overall student performance was clearly better than in 2003 when only in class pen-and-paper exercises were used.

In 2005, the 2001 and 2004 studies were repeated at the Helsinki University of Technology (HUT) and at the University of Turku (UTU) during the spring semester (Laakso et al., 2005b). The students ($N = 133 + 134$) were divided into two randomized exercise groups in both universities. The first group started their exercises on the web with the TRAKLA2 learning environment while the second group did their exercises in classroom sessions. In order to prevent the high drop-out rates (see, group C in 2001), however, the same learning experience were provided for
all the students. At the midpoint of the course, the treatment for the students was changed. The first group continued in the classroom and the second group on the web. Moreover, the same attitude survey, which carried out at UTU in 2004, was administered in both of the aforementioned universities.

The study concluded that it is good to introduce easy and guided exercises at the very beginning of the course. In addition to this, there is an emerging need for both web-based and classroom exercises. The recommended way to introduce the web-based exercises in DSA courses is by combining these two approaches. There is a set of exercises that are more suitable to be solved and automatically assessed on the web while the rest of the exercises are more suitable for traditional classroom sessions. More detailed information about this repetition study can be found in Laakso et al. (2005b).

The above studies were whole semester studies, in which the focus was on students’ overall performance and drop-out rates. The difference between the treatments were in learning settings: the control groups were in classroom while the treatment groups were on the web. However, the learning objectives were the same for all groups, i.e., the exercises were algorithm simulation exercises. In addition, we studied the students’ attitudes towards web based learning environments.

In contrast to the above studies, Myller et al. (2007) conducted an experimental study focusing on engagement taxonomy in fall 2006 at University of Turku. In the study, the learning outcomes of the students, who learned in collaboration by using visualization on different engagement levels were compared. There were 52 students in the treatment group (EET level: changing) and 53 students in the control group (EET level: controlled viewing), which sums up to 105 participants. The setup was a pre-test, treatment, post-test design. The post-test included the same questions as the pre-test, and additionally more difficult questions in order to see if the differences were apparent in them. The results indicated that the level of engagement had an effect on students’ learning results in favor of the treatment group, although the differences were not statistically significant. Especially students without previous knowledge seemed to learn more from using visualizations on higher engagement level. In this paper, we report on a replication of this study with minor changes in order to repair the flaws in the design of the pre-test and post-test as reported by Myller et al. (2007).

**Experimental Setup**

To summarize the previous sections, the collaborative use of AV tools has been studied only little, yet the need for this kind of research emerges from the increasing use of visualization tools in collaborative learning. We hypothesize that the EET framework can be used to predict performance differences when visualizations are used in collaboration. Previous research supports this view and our hypothesis is based on the previous research on TRAKLA2 and formulated as follows: Students using visualizations collaboratively on EET-level changing (i.e. in pairs) perform better compared to students using only visualization on EET-level controlled viewing (again in pairs).

In order to test our hypothesis, we carried out an experiment in which we compared the learning outcomes of students who were collaboratively using visualizations which were on different EET levels. Participants were (mostly first year) Computer Science major students on a data structures and algorithms course at the Helsinki University of Technology. We utilized TRAKLA2 (Korhonen et al., 2004) in order to provide students with algorithm simulation exercises that act on the EET level changing (treatment group). However, the students did not have the option to reset the exercise to obtain a new similar exercise with new input data, but they had to work with a fixed input data for each exercise during the whole session. The animations that the students used in controlled viewing condition (control group) were similar to those used in model answers provided by the TRAKLA2 system.

Quantitative results were analyzed with one-tailed t-test, ANOVA and $\chi^2$-test depending on the nature of the data. We used the Bonferroni correction when applicable. The justification for using one-tailed t-test is based on the formulation of our hypothesis, which predicts that students using visualizations on EET-level changing perform better than students using visualization on EET-level controlled viewing. The hypothesis is based on the previous research in which it was found that student groups using visualizations on EET-level changing consistently performed better than student groups using visualization on EET-level viewing or controlled viewing although differences were not statistically significant (Myller et al., 2007).
**Method 1: Experimental Study**

The study was a between-subject design with pre-test and post-test (dependent variable). We had one between-subject factor (independent variable): the highest available EET level of the visualizations in the learning materials, namely *controlled viewing or changing*. The unit of analysis was either a student or a pair of students depending on the measure. Each student answered the pre- and post-test individually, but all the observational data collected during the pair learning is not individual but the same for the pair. Moreover, we also report the average performance of the pair in the post-test and use it in the analysis.

![Figure 2: Binary heap insert animation in the tutorial. The student acts on EET level controlled viewing. The user has VCR like buttons (Backward, Forward, Begin, End) to interact with the animation](image)

The learning materials contained textual materials that were the same for both conditions. In the *changing* condition, textual materials were accompanied with TRAKLA2 (Korhonen et al., 2004) algorithm simulation exercises related to the binary heap (see Figure 1). Student pairs in the *controlled viewing* condition were presented with animations about the operations of the binary heap that were similar to TRAKLA2 exercises (see Figure 2). In addition, student pairs in both conditions were given an exercise sheet that asked questions on binary heap that were supposed to be answered during the learning process. In this way, we tried to motivate the learning and make sure that the possible differences are due to controlled variable (level of engagement), and not because pairs in one condition performed cognitively more demanding activities or used more time on the tasks (Grissom et al., 2003; Hundhausen et al., 2002).

**Method 2: Observational Study**

The students’ activities during the controlled experimental study were also recorded utilizing a screen capturing software. The recording accompanied by an audio track contained on-screen activity, i.e., mouse movements, keyboard typings, scrolling of the tutorial page back and forth in the browser window, as well as the conversation between the pair members.
The observed pairs were aware of being observed and we asked a permission to monitor them in advance. In this overt research method, we observed the students in their activities without intervention, i.e., by watching the recordings afterwards (Gall et al., 2006).

A detailed record of the events that occurred during the period of monitoring the students was produced. These events were categorized into the following four engagement levels according to the extended engagement taxonomy: no viewing (e.g., reading phase), viewing (e.g., watching figures), controlled viewing (e.g., watching of animations or model solution step-by-step with user controls) and changing (i.e., solving an algorithm simulation exercise). We separated passive viewing and more active controlled viewing from each other. In passive viewing, there was a still picture on the screen that we assumed the pair was watching. However, some of this time was spent to solve the given exercises on paper, as well. In controlled viewing, however, we knew that students were more actively involved with the animation as we required that they needed to control the animation by pressing VCR-like buttons to execute the animation backwards or forwards, and there were no pauses longer than 20 seconds between each action. The total time-on-task was measured from each four EET levels. Obviously, the students in controlled viewing condition (control group) did not spend time on changing mode. However, not all students in changing condition (treatment group) did either. Based on this analysis, we classified the students to groups based on their behavior.

Participants

Students were mainly first year students, however, some students from other years were also on the course. Students were randomized to the computer lab sessions and sessions were randomly assigned to each condition with the limitation that parallel sessions belonged to different conditions. The total number of participating students was 92. However, not all of them allowed to monitor their performance, nor were they willing to do pair work. In addition, in some of the workstations, the Java applet was not working properly. Moreover, we excluded foreign students from the study as they did not get the same treatment as the others due to the fact that their study materials were in a different language (i.e. English, while the original materials were in Finnish) and did not include animations nor algorithm simulation exercises, but they solved them by paper and pencil. Thus, the total number of analysis units (students) was 75 (n = 75) divided into 7 small groups (3 control groups having viewing condition and 4 treatment groups having changing condition). The original number of lab sessions was 8, but the last one (that would have been control group) was the excluded English speaking group.

All students had been previously using TRAKLA2 during the course to complete three assignment rounds related to basic data structures (e.g., lists and stacks), algorithm analysis, sorting algorithms (i.e., insertion sort, quicksort, and mergesort), and binary tree traversing. Thus, all students should have been able to use TRAKLA2, understand its visualization, and know all its features that were needed to complete the assignments.

Materials

Pre-test consisted of the following questions. In the first question, the student were asked to define concepts array, binary tree, and priority queue. We assumed that the students are able to answer the first two as those concepts were already introduced in the course. The last concept and the rest of the questions were such that we assumed the participants do not have prior knowledge to answer them. However, we wanted to test whether they have some prior knowledge, e.g., due to taking the course already in the previous year (without passing it). The second question was, if a given array is a heap and the third, whether an ordered array is a heap or not. In addition, we asked the students to describe where the smallest value in a minimum binary heap (question 5) and maximum binary heap is located (question 6), respectively. Finally, we asked them to write down a given binary heap’s heap property (question 7). The third question asked the students to draw the binary tree representation of the minimum binary heap, which was given in an array presentation, in the previous question.

The post-test consisted of the following questions. The pre-test and post-test included two questions which were exactly the same. The first question in the pre-test was omitted from the post-test. However, the questions 2, 3, 4, 5, 6 and 7 were the same in both (but the numbering started from 1 in the post-test). In addition, participants needed to do similar exercises that they did in the lab session. One of these was insertion of new items into an initially empty maximum binary heap (question 7 in the post-test). The question 8 asked participants to remove two smallest items
from a minimum binary heap. Finally, we gave a pseudo-code example of a recursive MAX-HEAPIFY procedure and asked several questions, such as for which algorithm one can apply this procedure (question 9). This was a multiple choice question with four alternatives of which the last three were applicable: Heap-Insert, Heap-Extract-Max, (linear-time) BuildHeap, and HeapSort. In addition, we asked them to describe and give an example execution (line-by-line) of what this procedure does and how (question 10). Question 11 requested the participants to provide an example which shows the recursive nature of the algorithm. The code example did not have a complete implementations for how to inquire the left and right child of a node in a complete binary tree implemented as an array. The task was to write this code (e.g., \texttt{LEFT(i)} = 2i and \texttt{RIGHT(i)} = 2i+1) (questions 12). Finally, they needed to analyze the worst case time complexity of MAX-HEAPIFY (question 13).

**Procedure**

Study was performed halfway through the course at the computer lab sessions that lasted for 2 hours. There were a total of 4 + 3 sessions, and they were run on two days in two following weeks. On each day, there were two times two sessions with different conditions running simultaneously. On the second day, there were also 4 sessions, but only 3 of them were included in this study as the last one was the excluded session given in English.

In the beginning of the session, students took the individual pre-test, in which they needed to answer questions related to binary heaps in 15 minutes. After this, they freely formed pairs with their peers and gave their consent to participate in the experiment and to be monitored during the experiment. If there was an odd number of students, one group consisted of 3 students. Each pair was allocated to a single computer.

After the pre-test, students had 45 minutes to go through the learning materials of their condition and complete paper-and-pencil exercises together. The collaboration was monitored by recording their talking and capturing their activities on the computer screens. After the 45 minutes the paper-and-pencil exercises were collected and the session ended with an individual post-test. The students were given 30 minutes to answer the questions in the post-test.

Each question in the pre- and post-tests was analyzed in a scale from 0 and 4. Zero points meant less than 25 percent of the answer was correct in the answer, and each point meant a 25 percent increase in the correctness of the answer.

**Results**

**Randomized Treatment and Control Groups**

In this section, we report the results as they were obtained by using the randomized treatment groups (42 students) and control groups (33 students) \( n = 75 \).

**Previous Knowledge and Motivation**

All the information related to the previous knowledge of the students could be determined only through post-hoc analysis, and thus, we could not make sure before-hand that the randomization did not introduce any bias to the experimental settings. Table 1 represents the students’ previous knowledge in Computer Science and Mathematics for both groups. The first column shows the pre-test scores for the topics studied in the experiment. The column “Prog. Course Results” shows the students’ average grades from a previous programming course. The average number of CS and Math credits units (each credit unit equals to about 30 hours of work) obtained are shown in the next columns, respectively. The difference between groups in the previous programming course grades is approaching statistical significance \( t(73) = -1.94, p = 0.056 \). Other differences are statistically insignificant.

| Table 1: Previous knowledge of the students on Heap data structure, and in CS and Math |
|--------------------------------------------------|-----------------|-----------------|-----------------|
| Pre-test | Prog. Course Grade | CS (Credit Units) | Math (Credit Units) |
| Control (33) | 9.27 (6.87) | 2.61 (1.77) | 10.72 (16.77) | 9.13 (9.33) |
| Treatment (42) | 8.57 (5.04) | 3.36 (1.57) | 10.44 (14.80) | 8.34 (6.87) |
Table 2 shows the results from a motivational questionnaire filled in by the students. The questions were answered in a 7-degree Likert-scale and they were as follows:

Q1. How useful do you regard this course for your working career?
Q2. Do you expect that the on-line learning will help your learning of the course content?
Q3. How well do online exercises fit into this course?
Q4. How useful have the on-line learning tools and materials been in your previous courses?

<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>4.84 (1.25)</td>
<td>4.78 (1.18)</td>
<td>5.38 (1.01)</td>
<td>4.94 (1.39)</td>
</tr>
<tr>
<td>Treatment</td>
<td>5.12 (1.33)</td>
<td>5.24 (1.14)</td>
<td>5.88 (1.05)</td>
<td>5.59 (1.30)</td>
</tr>
</tbody>
</table>

There were no statistically significant differences between the groups in any of the questions in the motivational questionnaire.

Post-test results

In the post-test, we used the same questions as in the pre-test and in addition to this seven more demanding questions. In the questions that were the same as in the pre-test, control and treatment group received on average 16.88 points \((st.dev. \ 4.34)\) and 17.38 points \((st.dev. \ 4.32)\), respectively. When comparing the pre- and post-test scores on the same questions within the group, statistically significant differences were found in both groups’ total scores using pairwise t-test (Control: \(t(33) = -13.48, p < .001\), Treatment: \(t(42) = -25.71, p < .001\)) (see the Table 1 for average pre-test scores and standard deviations). This means that both groups had learned the subject, which seems obvious when they spent 45 minutes to learn the topic.

When the points from all the questions were summed together the control group received on average a total of 30.79 points \((st.dev. \ 6.99)\) and the treatment group 31.55 \((st.dev. \ 6.29)\) points out of 52 points. There were no statistically significant differences found between the post-test scores.

We further calculated pair averages by taking the average of individual post-test scores of the pair. We treat this value as the learning outcome of a pair. The pair averages for control and treatment groups were 30.68 points \((st.dev. \ 4.74)\) and 31.63 points \((st.dev. \ 4.44)\), respectively. There were no statistically significant differences between the final scores or in any individual question scores.

Observational Study

In this section, we report the results as obtained by using a video analysis to match the students activities with the definition of treatment and control group. Based on the analysis, we regrouped students into different groups based on their behavior during the observation. We identified three groups based on their assignment to control and treatment groups and their behavior. Firstly, the students in the control group seemed to behave homogeneously and they watched the animations as expected. We will refer to this group with the name Viewing C (C as in Control).

Secondly, we identified a group of students in the treatment condition, who behaved exactly the same as the control group by only watching the animations and not even once trying to do any algorithm simulation exercises. We will refer to this group with the name Viewing T (T as in Treatment). We will refer to all students who only viewed the animations (i.e. students in groups Viewing C and Viewing T) with the name Viewing A (A as in All). Thirdly, we found the students who behaved as we expected in the treatment group. These students solved algorithm simulation exercises at least one time but most often three to six times. We will refer to this group with the name Changing T. The division of the groups is illustrated in Figure 3.

Based on the video analysis, we classified 33 students to the Viewing C, 17 student to the Viewing T, and 21 students to the Changing T \((n=71)\). We needed to exclude four students from the analysis in this section due to technical problems when matching the students to correct videos. Two of the students would have belonged to the Viewing T and two to the Changing T groups.
In this section, we present two comparisons. Firstly, we analyze the data between three groups, namely Viewing C, Viewing T and Changing T because based on the original randomization and the video analysis these groups are distinct. However, when only the video analysis and groups’ behavior is taken into consideration, we have only two groups, namely Viewing A and Changing T. Therefore, in order to provide a complete account of the results, we provide the analysis of both of these groupings. The validity, justifications and methodological implications of these groupings are further discussed in section Error! Reference source not found.

**Previous Knowledge and Motivation**

The format of Table 3 is similar to the Table 1. None of the differences were statistically significant neither Viewing C vs. Viewing T vs. Viewing T nor Viewing A vs. Changing T. This was different compared to the original experimental design where there was a significant difference in favor of the treatment group in previous programming course grades.

**Table 3:** Previous knowledge of the students on Heap data structure, and in CS and Math

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th>Prog. Course Grade</th>
<th>CS</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Viewing C</strong></td>
<td>9.27 (6.87)</td>
<td>2.61 (1.77)</td>
<td>10.72 (16.77)</td>
<td>9.13 (9.33)</td>
</tr>
<tr>
<td><strong>Viewing T</strong></td>
<td>8.06 (4.49)</td>
<td>3.47 (1.46)</td>
<td>12.56 (21.04)</td>
<td>7.69 (6.63)</td>
</tr>
<tr>
<td><strong>Viewing A</strong></td>
<td>8.86 (6.14)</td>
<td>2.90 (1.71)</td>
<td>11.33 (18.10)</td>
<td>8.64 (8.46)</td>
</tr>
<tr>
<td><strong>Changing T</strong></td>
<td>9.29 (5.72)</td>
<td>3.14 (1.80)</td>
<td>10.43 (9.35)</td>
<td>9.67 (7.21)</td>
</tr>
</tbody>
</table>

Table 4 shows the results from the same motivational questionnaire that was also reported in the Table 2 for the experimental groups (See Section 0 for the description of the questions). None of the differences were statistically significant.

**Table 4:** Motivation of students based on a questionnaire. *Note.* Questions from Q1 to Q4 are discussed in Section 0

<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Viewing C</strong></td>
<td>4.84 (1.25)</td>
<td>4.78 (1.18)</td>
<td>5.38 (1.01)</td>
<td>4.94 (1.39)</td>
</tr>
<tr>
<td><strong>Viewing T</strong></td>
<td>5.00 (1.51)</td>
<td>5.25 (0.93)</td>
<td>5.81 (1.05)</td>
<td>5.44 (1.26)</td>
</tr>
<tr>
<td><strong>Viewing A</strong></td>
<td>4.90 (1.32)</td>
<td>4.94 (1.12)</td>
<td>5.52 (1.03)</td>
<td>5.10 (1.36)</td>
</tr>
<tr>
<td><strong>Changing T</strong></td>
<td>5.19 (1.33)</td>
<td>5.19 (1.36)</td>
<td>5.86 (1.11)</td>
<td>5.67 (1.43)</td>
</tr>
</tbody>
</table>

**Time Allocation between Engagement levels**

Table 5 presents the distribution of the average times spent on each EET level. This was measured by watching the videos and marking times when the EET level changed from one to another, and then summing up the times on each EET level.
Table 5: The distribution of time (45 minutes) between EET levels

<table>
<thead>
<tr>
<th></th>
<th>No viewing</th>
<th>Viewing</th>
<th>Controlled viewing</th>
<th>Changing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viewing C</td>
<td>47.45 % (15.28)</td>
<td>38.26 % (12.24)</td>
<td>14.29 % (6.23)</td>
<td>0.00 % (0.00)</td>
</tr>
<tr>
<td>Viewing T</td>
<td>49.45 % (17.09)</td>
<td>37.82 % (15.01)</td>
<td>12.73 % (5.47)</td>
<td>0.00 % (0.00)</td>
</tr>
<tr>
<td>Viewing A</td>
<td>48.13 % (15.78)</td>
<td>38.11 % (13.10)</td>
<td>13.76 % (5.97)</td>
<td>0.00 % (0.00)</td>
</tr>
<tr>
<td>Changing T</td>
<td>43.22 % (19.20)</td>
<td>38.30 % (15.84)</td>
<td>5.87 % (6.03)</td>
<td>12.61 % (1.98)</td>
</tr>
</tbody>
</table>

Table 6 shows how many times students used materials on each EET level. For example, students in the control group used user-controlled visualizations (controlled viewing) 5 times on average, whereas students in the treatment group used them 2 or 3 times on average.

Table 6: The number of times each EET level was used

<table>
<thead>
<tr>
<th></th>
<th>No viewing</th>
<th>Viewing</th>
<th>Controlled viewing</th>
<th>Changing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viewing C</td>
<td>6.76 (2.11)</td>
<td>7.82 (3.61)</td>
<td>5.15 (2.71)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Viewing T</td>
<td>7.18 (2.19)</td>
<td>7.53 (3.04)</td>
<td>5.29 (2.91)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Viewing A</td>
<td>6.90 (2.12)</td>
<td>7.72 (3.40)</td>
<td>5.20 (2.75)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Changing T</td>
<td>6.24 (1.73)</td>
<td>6.67 (3.20)</td>
<td>2.48 (2.56)</td>
<td>4.10 (1.61)</td>
</tr>
</tbody>
</table>

Post-test results

The results of the post-test are presented in Table 7. When comparing the pre- and post-test scores within the group, statistically significant differences were found in both groups’ total scores between pre- and post-tests when only same questions were compared with pairwise t-test (Viewing C: t(32) = -13.15, p < .001, Viewing T: t(16) = -13.96, p < .001, Viewing A: t(49) = -18.09, p < .001, and Changing T: t(20) = -19.35, p < .001) (see the Table 3 for average pre-test scores and the subtotal in the Table 7 for the comparable average post-test scores and standard deviations).

Table 7: Post-test results. Note. Post-test questions were discussed in Section 0 and composition of the groups in Figure 3

<table>
<thead>
<tr>
<th></th>
<th>Viewing C</th>
<th>Viewing T</th>
<th>Viewing A</th>
<th>Changing T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>2.64 (1.58)</td>
<td>2.12 (1.65)</td>
<td>2.46 (1.61)</td>
<td>2.33 (1.80)</td>
</tr>
<tr>
<td>Question 2</td>
<td>1.76 (1.23)</td>
<td>1.82 (1.29)</td>
<td>1.78 (1.23)</td>
<td>2.19 (1.29)</td>
</tr>
<tr>
<td>Question 3</td>
<td>3.64 (1.08)</td>
<td>4.00 (0.00)</td>
<td>3.76 (0.89)</td>
<td>4.00 (0.00)</td>
</tr>
<tr>
<td>Question 4</td>
<td>2.39 (1.23)</td>
<td>2.18 (1.33)</td>
<td>2.32 (1.42)</td>
<td>2.33 (1.59)</td>
</tr>
<tr>
<td>Question 5</td>
<td>2.61 (1.43)</td>
<td>2.65 (1.58)</td>
<td>2.62 (1.47)</td>
<td>3.38 (0.92)</td>
</tr>
<tr>
<td>Question 6</td>
<td>3.85 (0.71)</td>
<td>3.76 (0.97)</td>
<td>3.82 (0.80)</td>
<td>4.00 (0.00)</td>
</tr>
<tr>
<td>Subtotal</td>
<td>16.88 (4.34)</td>
<td>16.53 (4.90)</td>
<td>16.76 (4.49)</td>
<td>18.24 (3.56)</td>
</tr>
<tr>
<td>Question 7</td>
<td>3.97 (0.17)</td>
<td>3.94 (0.24)</td>
<td>3.96 (0.20)</td>
<td>3.43 (1.29)</td>
</tr>
<tr>
<td>Question 8</td>
<td>3.33 (1.19)</td>
<td>3.65 (1.00)</td>
<td>3.44 (1.13)</td>
<td>3.76 (0.89)</td>
</tr>
<tr>
<td>Question 9</td>
<td>2.48 (0.87)</td>
<td>2.12 (0.78)</td>
<td>2.36 (0.85)</td>
<td>2.67 (0.91)</td>
</tr>
<tr>
<td>Question 10</td>
<td>2.09 (1.44)</td>
<td>2.41 (0.94)</td>
<td>2.20 (1.29)</td>
<td>2.62 (1.40)</td>
</tr>
<tr>
<td>Question 11</td>
<td>0.45 (1.25)</td>
<td>0.71 (1.45)</td>
<td>0.54 (1.31)</td>
<td>1.10 (1.70)</td>
</tr>
<tr>
<td>Question 12</td>
<td>1.30 (1.85)</td>
<td>0.18 (0.73)</td>
<td>0.92 (1.64)</td>
<td>1.24 (1.84)</td>
</tr>
<tr>
<td>Question 13</td>
<td>0.27 (0.45)</td>
<td>0.29 (0.99)</td>
<td>0.28 (0.67)</td>
<td>0.29 (0.46)</td>
</tr>
<tr>
<td>Total</td>
<td>30.79 (6.99)</td>
<td>29.82 (5.71)</td>
<td>30.46 (6.54)</td>
<td>33.33 (6.71)</td>
</tr>
<tr>
<td>Pair Average</td>
<td>30.68 (4.74)</td>
<td>29.88 (4.37)</td>
<td>30.42 (4.55)</td>
<td>33.45 (4.34)</td>
</tr>
</tbody>
</table>

Based on ANOVA, there were no statistically significant differences between Viewing C, Viewing T and Changing T groups in the post-test scores. When comparing the total values from the post-tests between Viewing A and Changing T, statistically significant differences were found in the total and pair average of the post-test scores by using one-tailed t-test (t(69) = -1.73, p < 0.05) and (t(31) = -1.97, p < 0.05), respectively.
Discussion

Interpretation of the Results

We presented an empirical study which analyzed whether the EET framework can be used to predict performance differences when algorithm visualizations are used in collaboration. Two randomized groups of students were involved in this study reading and answering questions related to a hypermedia tutorial presented on a web page. The control group used the algorithm visualizations on controlled viewing level, on which they had the opportunity to watch algorithm animations embedded in the tutorial. The treatment group interacted with the tutorial on changing level, on which they had the option to solve small algorithm simulation exercises and get feedback on their performance. In both groups, the students formed pairs and learned collaboratively about the binary heaps for 45 minutes during the 2-hour closed lab session. The analysis of the video material has showed that students were collaborating and discussing the subject matter during the learning process, therefore we are confident to say that students were truly learning collaboratively in both groups (Myller et al., in press). The null hypothesis of the experiment was that there would be no significant statistical difference between the learning outcomes of the control and treatment group after the session.

Pre- and post-tests were used to analyze the performance. Each student answered these tests individually. There were no significant differences between groups if we analyzed only the pre-test scores. However, post-hoc analysis of some background variables revealed that there was almost a significant bias between the groups. The grades from the previous programming course were better in the treatment group than in the control group. Furthermore, based on the post-test results we could not reject the null hypothesis. This all was (at first) a counter-intuitive result, because a) it was against the theory that we were testing, b) it was against our previous findings and c) even the bias between the groups was in favor of the treatment group.

Fortunately, during the experimental study, we monitored the student pairs in a parallel observational study. After examining the video recordings, we realized that not all of the students in the treatment group were using the tutorial as expected. Some of the pairs did not solve the exercises, but only watched the model solutions instead. Thus, they were interacting with the tutorial only on controlled viewing level, not in changing level as expected. Based on this new evidence, we re-grouped the students. We regarded those students in the treatment group, not behaving on the changing level, belonging to a controlled viewing level. Interestingly, the aforementioned bias in previous programming course grades disappeared, and we found significant differences between the learning outcomes of the groups. Although there were no differences when only three groups were compared, the group working on changing level outperformed all student groups working on controlled viewing level in the total score of post-test. This was true both in the individual performance and the average performance of pairs. Thus, based on this study, we can reject the null hypothesis and confirm our previous findings that the level of engagement on which the students interact with the visualization tool has an influence on the learning. On changing level, they learned better than on controlled viewing level.

Stasko et al. (1993) hypothesize that “algorithm animations will not benefit novice students just learning a new topic as much as the animations will benefit more advanced students”, and moreover, that “the novice students would benefit more by actually constructing an algorithm animation rather than viewing a predefined one.” We can confirm these hypotheses. However, in this first hypothesis, we need to be careful in the definition of a “novice”. In our experiment, all students were exposed to TRAKLA2 before they attended the experiment. They solved similar exercises, but on different topics, a couple of weeks before the experiment took place. Thus, they were not “novices” when it comes to the “graphical notation” used in the experiment. Still, they were novices when it comes to the topic (i.e. they had not studied binary heaps earlier). Therefore, the conclusion is that the first hypothesis holds only if “novice” is defined to be a student who is not familiar with the used notation in the animations. One can still be a novice of the topic but understand the used notation, and benefit as much as more advanced students. Actually, it might even happen that the more advanced students cannot take the full advantage of this kind of learning material, and thus, perform worse, at least in relative scale (Myller et al., 2007). The confirmation of the second hypothesis is a direct outcome of our study in which the treatment group was “constructing an algorithm animation” in terms of changing the visualization, and they outperformed those students in the control group who just were “viewing a predefined” animation.
As discussed in the section on previous research, the learning time has not been a controlled variable in several previous studies, which have used the engagement level as the independent variable (Grissom et al., 2003; Naps et al., 2002; Hundhausen et al., 2002). Furthermore, it has been reported that students using visualizations on higher engagement levels have been motivated to spend more time on learning the topic. This has made it questionable if the time that students spend on learning the topic affects the learning results more than the engagement level, on which the visualization is used, and the engagement level affects only the amount of time students are willing to spend on learning the topic. In this study, we have shown that although we controlled the learning time and monitored students’ activities, the learning results are significantly different between engagement levels. This means that the engagement level has a direct effect on the learning results.

Methodological Considerations

Based on the results, screen capturing and voice recording should be a standard procedure because we cannot always know for sure if the participants really do what we expect them to do. Our study shows that we could not have obtained full understanding of the phenomenon without monitoring the students: not all of them performed on the expected engagement level even though we instructed them to do so. As we can see from our study, the conclusion would have been that we could not find any evidence that the EET level has an impact on learning, which would have been a false negative result. Thus, monitoring should be a standard procedure especially in large scale studies in which the researcher(s) cannot make sure by other means that the conditions remain constant within a group.

However, when using an observational design in the study, we need to pay attention to possible confounds that might affect our results. Due to the fact that in the observational study, we could not control the placement of participants into conditions, but they selected it themselves, this could have caused differences in the final results and there still might be background variables that we have not analyzed or detected affecting the results. However, as stated earlier, we did a post-hoc analysis of several background variables and detected that actually the re-grouping made the groups more similar on one aspect while keeping the other aspects unchanged. Thus, we are fairly confident that the observed differences are due to the claimed causes.

Conclusion and Future Work

Our results confirm that EET framework can predict performance differences also in collaborative use of visualizations. The results substantiate that there is a difference in learning results between viewing and changing modes. The findings of the observational study also explain why the original experimental design failed to reject the null hypothesis. This was due to the fact that students in the treatment group did not perform the learning tasks that we assumed them to do. Thus, they might have outperformed the control group in the experimental design if they only had performed in the changing mode.

From our point of view, the results emphasize the importance of engagement with visualizations, and we should promote systems that support different modes of engagement. The mere viewing of the algorithm animations is not enough, not even when there is a partner with whom to share the understandings and misunderstanding during the viewing of the visualization. Thus, we should, especially, design systems that act on the higher levels of the engagement taxonomy. For example, visual algorithm simulation exercises acting on the changing level produce better results compared to the viewing level. Furthermore, we should encourage the use of the systems on higher engagement levels in classrooms in order to achieve active and more student-centered learning. We hope this paper encourages teachers on different disciplines to try out visualization tools that enable higher engagement between the tool and the students especially in collaborative learning as this seems to increase the learning outcomes.

The future research challenge is to determine the importance and role of collaboration in the EET, i.e., can we repeat this experiment also in the case of individual learning? In this experiment, collaboration was used to encourage discussion in pairs and to collect better evidence of the real behavior in terms of screen capturing. The collaboration, however, has an influence on the performance as well. Thus, one research direction would be toward individual learning, but in a context that can still be monitored in order to prevent inconclusive results due to the fact that the individuals did not behave on the expected EET level.
Acknowledgements

This work was supported by the Academy of Finland under grant numbers 111350 and 210947. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Academy of Finland.

References


Developmental Progression of Referential Resolution in Comprehending Online Texts

Yu-Fen Yang and Ya-Chen Hung
Graduate school of applied foreign languages, National Yunlin University of Science & Technology, Douliu, Yunlin, Taiwan, R. O. C. // Tel:+ 886-5-534-2601 Ext. 3136 // Fax: +886-5-531-2058 // yangy@yuntech.edu.tw

ABSTRACT
The purpose of this study is to examine readers’ comprehension as they develop their mental representation of reference in four sequential online texts. A total of 92 college students from three reading classes were recruited to complete the following steps in each text: (1) identify references, (2) draw the relationships between references, and (3) answer reading comprehension test items. Results of this study showed that the correlation between referential resolution and reading comprehension tests ranged from .68 to .90 in four online texts. This indicated that when readers’ scores in referential resolution increased, their scores in reading comprehension tests were also raised. Among three groups of readers, the more-proficient readers were able to integrate the references in different parts of the text as a coherent network from text 1 to 4. In contrast, average and less-proficient readers initially did not integrate any reference when reading the first text. They eventually clustered different references and referred them to a correct subject in the final text. The keys to the development of college readers’ mental representation of reference lay in whether they were actively engaged in comprehension monitoring and frequently asked for feedback tool as a scaffold.

Keywords
Reading strategy; Feedback tool; Trace result; Comprehension monitoring; Reading comprehension

Introduction
Whether a reader is able to construct a comprehensible and coherent mental representation of textual information in memory is central to comprehension (van den Broek & Kremer, 1998; Walsh & John-Laird, 2004). According to Payne and Reader (2006), the construction of mental representation is a necessary step for comprehension. It aids the reader to encode textual information in a clustered way so that the textual information is more likely to be stored into the reader’s long-term memory (Potelle and Rouet, 2003). Tea and Lee (2004) also state that the reader’s mental map presents his text processing and helps him solve reading difficulties, such as referential resolution.

Referential resolution is the process of searching for events, people, or objects appearing in different parts of a text referring to the same entity (Paterson, Sanford, Moxey, and Dawydiak, 1998). This is essentially difficult for college students who learn English as a Foreign Language (EFL) in Taiwan for they very often fail to recognize the connections among sentences in texts due to the lack of instruction in referential resolution (Bensoussan and Laufer, 1984; Chu, Swaffar, and Charney, 2002). For instance, in a short text “I have a brother. His name is Tom. He is a senior high school student.” The mental map of this short text in referential resolution is shown in Figure 1.

![Figure 1. An example of mental map in referential resolution](image)

Referential resolution in this study is defined as a reading strategy applied by the reader to interpret the references that have the same meanings as other elements in a text, such as “his name”, “Tom”, and “he” refer to “a brother.” While resolving the references, the reader is engaged in comprehension monitoring which he monitors, regulates, and evaluates his own reading process (Hartman, 2001). The management and regulation of one’s own reading process is helpful for meaning construction of text (Paris and Winograd, 1990).
Comprehension monitoring is the awareness that a reader has about the linguistic forms and their relationships with other elements in a text (Potelle and Rouet, 2003). More-proficient readers are found to plan, predict outcomes, and monitor their reading process (Brown, 1987). Particularly, they are able to detect inconsistencies in reading and commonly look back at and recall the text inconsistencies (Zabruck and Ratner, 1989). The results of Yang’s study (2002) also reveal that the more-proficient readers actively engage in monitoring their ongoing reading process as they try to compensate for words that have not been previously decoded. They also employ higher levels of comprehension monitoring in reading such as examining text coherence by internal and external consistency (Baker, 1996).

In contrast, less-proficient readers are indicated to have a lower level of comprehension monitoring (Oakhill & Yuill, 1988). Their lower comprehension monitoring may result from inadequate knowledge and skills in reading comprehension. They commonly have a limited vocabulary and they often fail to comprehend the individual words in a text. As a result, they spend much time and efforts on decoding each word in a sentence rather than comprehending and integrating textual information. Focusing on the decoding process, most less-proficient readers are unaware of connections in linking sentences together (Bensoussan & Laufer, 1984).

Less-proficient readers will not engage in comprehension monitoring unless they are asked to think about their reading process through activities or instruction (Hartley, 2001). The computer assisted learning environment is found to greatly support readers’ engagement of comprehension monitoring as it provides explicit modeling and individualized scaffolding (Potelle and Rouet, 2003). The modeling and scaffolding may help readers build the mental map to show their cognitive structure and the meaningful content of the text. This is fundamentally important for EFL college readers in Taiwan as there are about 45 or more students of varying language proficiency levels involved in one class. The large class size limits the classroom teacher from providing the individualized support and guidance and monitoring each reader’s progress.

In this study, the computer system first informs students of the goal of incorporating the system in instruction and the types of references a student is asked to identify. It then models the procedure in identifying and resolving reference. A trial section is also provided for students to practice. After these three activities, the system requires students to identify the references in reading an online text and then to figure out the relationship between references by drawing their mental maps. That is, the computer system allows the reader to recognize and clarify the meanings of the references in a text. If students encounter difficulties in identifying and resolving references, a feedback tool is provided for their scaffolding. The feedback tool provides three candidate references for each referential device that needs correction back to students. Finally, students are asked to finish an online reading comprehension test in each text.

The purpose of this study is thus to examine readers’ comprehension as they develop their mental representation of reference in four sequential online texts. Three research questions are addressed.

1. How do EFL readers with different English reading proficiency level develop their mental representation of reference in four sequential online texts?
2. How does the incorporation of system in instruction help EFL readers develop their mental representation of reference in comprehending texts?
3. How does the development of readers’ mental representation of reference assist their comprehension of online texts?

Method

Participants

A total of 92 junior and senior college students were recruited from three reading classes in a technological university in central Taiwan. Their language proficiency levels were identified by their reading scores in a simulated online exam Testing of English for International Communication (TOEIC). The maximum achievable score in the reading section of the online exam was 200.
The frequency distribution of all the participants’ scores was used to divide the participants into three groups of readers. The highest frequency falls in the two intervals, 101-110 (8 students) and 151-160 (8 students). These two intervals serve as the benchmarks for dividing the participants into three groups of readers. Participants with reading scores above 151 were identified as the more-proficient readers and those with reading scores below 101 were the less-proficient readers. Participants with reading scores between 101 and 151 were identified as the average readers. Thus, 29 more-proficient readers, 32 average readers, and 31 less-proficient readers were identified in this study. The more-proficient readers showed a mean score of 175.52 with a standard deviation of 14.60, the average readers a mean score of 126.41 with a standard deviation of 15.09, and the less-proficient readers a mean score of 74.03 with a standard deviation of 17.53.

Material

The online referential resolution practice used four texts to examine the participants’ reading comprehension. The four texts were selected from College Reading Workshop (Malarcher, 2005) based on the following four criteria: a number of references for reading practice, similar length, similar readability level, texts written for EFL college students. The four texts were Ideas about Beauty (number of words: 582; number of referring words: 25; Text 1), Fast Food and Teen Workers (number of words: 583; number of referring words: 25; Text 2), Adventure Tours for Charity (number of words: 599; number of referring words: 43; Text 3), and Traditional Markets vs. Modern Markets (number of words: 577; number of referring words: 31; Text 4). The full text of Ideas about Beauty (text 1) is shown in Appendix.

Procedures of Data Collection

Three phases were involved in the procedure of data collection. In the first phase, 92 college students received reading instruction in referential identification and resolution. In this instruction, types of reference and usage of the referring strategy were provided. This allowed students to have opportunities to practice the strategy. Three types of references, personal, demonstrative, and locative references, were investigated in this study as they appeared more frequently in texts (Fortanet, 2004). Personal reference refers to individuals or objects by specifying their functions...
or roles in a context, such as “I,” “me,” and “you.” Demonstrative references substitute nouns when the nouns can be understood from the context. They also indicate whether they are replacing singular or plural words. Examples include “this,” “these,” “that,” “more,” “neither,” etc. Locative references are used to indicate locations. Examples include “here” and “there.”

In the second phase, students were introduced the online system of referential resolution. They were demonstrated how to use the online system and offered chances to practice it. For example, when students encountered difficulties in constructing their initial maps, they could request feedback by clicking an icon in the tool bar. Figure 2 shows the feedback received by a student. The three candidate references were highlighted for the student to make a second attempt at the correct answer. Then, the student connected the chosen reference to construct the mental map (see Figure 3).

In the third phase, participants were asked to complete the online task in class during the period between October 2\textsuperscript{nd}, 2006 and January 3\textsuperscript{rd}, 2007. The online system of referential resolution required the student to follow the steps in reading an online text: (1) identify references, (2) draw relationships between the references, and (3) answer reading comprehension test items. There were ten items in each comprehension test and the full score points of each test were 10. Finally, the system recorded participants’ reading behavior and performance.

![Figure 3](image)

**Figure 3.** The mental map the student drew after feedback

**Procedures of Data Analysis**

The collected data were categorized into reading product and reading process. In reading product, students’ score points in referential identification, referential resolution, the frequency of feedback tool request, and reading comprehension test were analyzed by Statistical Package for Social Science (SPSS) 14.0 version. Mean, standard deviation, and Pearson product-moment correlation coefficient were computed to examine students’ performance on the referential resolution task and to investigate the relationships among referential identification, referential resolution, and reading comprehension test. In reading process, participants’ mental maps of referential resolution, the trace results of their reading process, and the frequency of feedback tool request were examined.
Result

Reading product

In reading product, students’ scores in referential identification and resolution were shown in percentage since the number of referential words is different in each text. For each reading comprehension test, the full score points are 10. The means and standard deviations of referential identification, resolution, and reading comprehension test are shown in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Text</th>
<th>Text 1</th>
<th>Text 2</th>
<th>Text 3</th>
<th>Text 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Referential identification</td>
<td>MR</td>
<td>22.93</td>
<td>1.25</td>
<td>23.72</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>21.16</td>
<td>2.81</td>
<td>21.32</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>LR</td>
<td>19.93</td>
<td>1.75</td>
<td>18.26</td>
<td>1.48</td>
</tr>
<tr>
<td>Referential resolution</td>
<td>MR</td>
<td>21.59</td>
<td>1.86</td>
<td>21.86</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>15.34</td>
<td>3.09</td>
<td>30.59</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>LR</td>
<td>13.13</td>
<td>2.09</td>
<td>14.19</td>
<td>2.27</td>
</tr>
<tr>
<td>Reading comprehension</td>
<td>MR</td>
<td>8.93</td>
<td>.70</td>
<td>8.07</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>6.28</td>
<td>.47</td>
<td>7.78</td>
<td>.42</td>
</tr>
<tr>
<td></td>
<td>LR</td>
<td>5.35</td>
<td>.49</td>
<td>6.45</td>
<td>1.12</td>
</tr>
</tbody>
</table>

Note: N refers to the number of participants; SD refers to standard deviation; MR refers to the more proficient readers; AR refers to average readers; LR refers to less proficient readers; the means of referential identification and resolution are shown in percentage.

As shown in Table 1, all three groups of readers made progress as they read the four sequential online texts. The more-proficient readers outperformed the average and the less-proficient ones as their mean scores in referential identification task ranging from 22.93 to 49.72. In contrast, the mean of the referential identification task for the less-proficient readers only ranges from 21.16 to 38.72. Furthermore, the mean of reading comprehension test for the more-proficient readers ranges from 8.07 to 9.55 whereas the mean for the average readers is from 6.28 to 8.19.

Figure 4. The relationship between referential identification and resolution
Correlation coefficient was also conducted to examine the nature of the relationship between the variables in this study. The correlation coefficient for the referential identification and referential resolution for text 1, 2, 3, and 4 are .81, .84, .85, and .84 which show a positive relationship between the two variables (see Figure 4). That is, as readers’ scores in referential identification increased, their scores in referential resolution also raised. Furthermore, the correlation coefficient between referential resolution and reading comprehension test for four texts are .68, .71, .94, and .90 which also indicate a positive relationship. As readers’ scores in referential resolution increased, their scores in reading comprehension test were also raised.

Reading process

In the following, one student was randomly selected from each reading proficiency group, a total of 3 students, to represent the developmental process of more proficient, average, and less proficient readers in referential resolution and reading comprehension.

Reading process of the more-proficient readers

In this study, the more-proficient readers were found to be able to integrate textual information in four sequential online reading tasks. Figure 5 shows one of the examples.

![Figure 5. The more-proficient reader's mental representation of personal references](image)

In addition to the more-proficient reader’s mental representation of referential resolution, he read and reread the sentences and requested the feedback tool as he was not sure what references referred to. Figure 6 presents one of the more-proficient reader’s reading process in which he requested the feedback tool for assistance.

Figure 6 shows that the more-proficient reader constantly requested the feedback tool 21 times (e.g. line 150). He requested the feedback tool either for overcoming his reading difficulties or confirming his selection. For instance, after he read the sentences, he requested the feedback tool of this (12) (e.g. lines 21–23). Then, he read and reread the sentences and finally made a correct selection. The more-proficient reader also requested the feedback tool as he tried to confirm the answer he chose. For example, after he connected the men (23) to its subject (e.g. lines 150–153), he checked the connection again by requesting the feedback tool for 4 times. Then, he revised his connection by cutting the previous selection.

Table 2 further presents the more-proficient reader’s frequency request of the feedback tool for resolving personal references. It was found that the more-proficient reader increased his request of feedback tool in resolving personal references in four texts.
Table 2. The more-proficient reader’s frequency request of the feedback tool in resolving personal references

<table>
<thead>
<tr>
<th>Text</th>
<th>Reference</th>
<th>Frequency</th>
<th>Reference</th>
<th>Frequency</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>them(5)</td>
<td>0</td>
<td>they(7)</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>her(9)</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(13)</td>
<td>1</td>
<td>their(14)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>their(16)</td>
<td>0</td>
<td>they(17)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(30)</td>
<td>1</td>
<td>their(31)</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>their(11)</td>
<td>1</td>
<td>their(13)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(14)</td>
<td>1</td>
<td>their(15)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(52)</td>
<td>1</td>
<td>their(53)</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>they(4)</td>
<td>1</td>
<td>their(6)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>They(7)</td>
<td>1</td>
<td>their(8)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(10)</td>
<td>1</td>
<td>their(11)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(12)</td>
<td>1</td>
<td>their(20)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(22)</td>
<td>1</td>
<td>their(24)</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

Reading process of the average readers

The average readers were discovered to develop their mental representation of references when reading four online texts. Figure 7 shows one of the average readers’ mental maps of personal references along with the four online texts.

In graph (a) of Figure 7, the average reader initially did not integrate any personal reference when reading the first text. She began to link some personal references in the second text. For example, in graph (b), she referred both their (13) and they (14) to teens. She even could find out the possible subjects that the personal references refer to, such as referring all they (30), they (31), and their (35) to students. In graph (c), she not only referred the personal references to a correct subject but also made these references mutually linked. For instance, the reference they (14) was connected to people and their (15). Eventually, she constantly referred the personal references to a correct subject. In graph (d), she integrated the personal references in different parts of the text referring to a correct subject. For example, she referred all they (4), they (20), they (22), and they (24) to people.
Figure 7. Examples of the average readers' mental maps in personal reference,
From the analysis of the trace results, it was found that the average reader relied much on the feedback tool in their reading process. She requested the feedback tool to either overcome her reading difficulties or confirm her selections of the references. Figure 8 shows one of the average readers’ trace results.

Figure 8. Trace results of the average reader’s reading process

As shown in Figure 8, the average reader requested the feedback tool 96 times (e.g. line 276). He entirely depended on the feedback tool in connecting the references to the correct subject (e.g. lines 22–26). After receiving the feedback, he did not reread the sentences for comprehension. As a result, he tended to have a partial understanding of the textual information and kept requesting the feedback tool for revising his connections. For instance, he requested the feedback tool of the demonstrative reference, this trip (6). He added and erased the text elements that he initially selected. The similar reading process repeatedly occurred as shown in line 272 to 276. He, again, requested the feedback tool of the demonstrative reference, the deposit (19), and made the final connection.

In developing the mental maps of reference, the average reader engaged in comprehension monitoring by the assistance of feedback tool. Table 3 shows the average reader’s frequency in requesting the feedback tool when resolving the personal references.

<table>
<thead>
<tr>
<th>Text</th>
<th>Reference</th>
<th>Frequency</th>
<th>Reference</th>
<th>Frequency</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>them(5)</td>
<td>0</td>
<td>they(7)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>her(9)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(13)</td>
<td>1</td>
<td>their(14)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>their(16)</td>
<td>1</td>
<td>they(17)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>they(30)</td>
<td>1</td>
<td>They(31)</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>their(11)</td>
<td>2</td>
<td>their(13)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(14)</td>
<td>3</td>
<td>their(15)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(52)</td>
<td>1</td>
<td>they(53)</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>they(4)</td>
<td>1</td>
<td>their(6)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>They(7)</td>
<td>1</td>
<td>their(8)</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>their(10)</td>
<td>1</td>
<td>their(11)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>their(12)</td>
<td>1</td>
<td>they(20)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>they(22)</td>
<td>1</td>
<td>they(24)</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

291
Table 3 shows that the average reader asked help from the feedback tool more frequently from text 1 to 4. He was able to monitor his reading process and frequently asked for assistance as he encountered difficulty in resolving the personal references. His performance in the reading comprehension test, 8 out of 10 points (80% correct), further illustrated his better textual understanding in text 2.

After the average reader had drawn the relationships among references, he still needed to take a multiple-choice comprehension test in each text. A test item in reading comprehension test of the text Fast Food and Teen Worker is shown as follows.

---

5. Which of the following is NOT true?
(A) Fast food workers quit their jobs easily.
(B) Most Americans have worked in fast food restaurants.
(C) Teenagers can hardly find jobs in fast food restaurants.
(D) McDonald’s is the largest fast food chain in the United States.

In answering the item #5, the reader first had to identify the main character in the text and select the correct description of the character. Since he referred and integrated the personal references in the text to the correct subject, most teenagers (see graph (b) of Figure 7), he was able to choose one of the four options that is mainly related to the teenagers, namely, option C.

**Reading process of the less-proficient readers**

The less-proficient reader was also found to develop his mental map of reference as he tried to finish the four sequential online tasks. Figure 9 presents the less-proficient readers’ mental maps in resolving personal references.
In graph (a) and graph (b) of Figure 9, the less-proficient reader did not integrate any of the demonstrative reference. Instead, he referred each demonstrative reference to its subject separately. In graph (c) and graph (d), he started to connect some of the demonstrative references like *the cow* (29) and *its* (30) to the correct subject, *his* (26) cow, in a successive way. *The store* (45) and *its* (42) were linked together with an *international chain store*.

In Figure 10, the trace results revealed that the less-proficient reader rarely requested the *feedback tool* in text 1. Most of the time, he repeatedly selected and erased the incorrect references and eventually referred the references to an incorrect subject when reading the online text.

Figure 10 shows that the less-proficient reader never requested the *feedback tool* for assistance even though he was not sure about what the references referred to in the text. Instead, he only read and reread some sentences (e.g. line 78–81). He then randomly connected a reference to a subject.

In developing his mental map of reference, the less-proficient reader was more aware of his reading process and actively asked for help. Table 4 shows the less-proficient reader’s frequency in requesting the *feedback tool*.
Table 4. The less-proficient reader’s frequency request of the feedback tool

<table>
<thead>
<tr>
<th>Text</th>
<th>Reference</th>
<th>Frequency</th>
<th>Reference</th>
<th>Frequency</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>this(20)</td>
<td>0</td>
<td>that(22)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>that(26)</td>
<td>0</td>
<td>It(36)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>this(2)</td>
<td>0</td>
<td>these(3)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>these(4)</td>
<td>0</td>
<td>This(5)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>its(6)</td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>his(23)</td>
<td>1</td>
<td>the field(24)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>the farmer(25)</td>
<td>1</td>
<td>his(26)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>the field(27)</td>
<td>1</td>
<td>the cow(29)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>its(30)</td>
<td>1</td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>the seller(26)</td>
<td>5</td>
<td>this(28)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>the shopper(29)</td>
<td>1</td>
<td>the shopper(31)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>it(42)</td>
<td>1</td>
<td></td>
<td></td>
<td>9</td>
</tr>
</tbody>
</table>

Table 4 revealed that the less-proficient reader requested the feedback tool more frequently when reading the final online text. By the assistance of the feedback tool, he not only solved his reading problems but also enhanced his reading comprehension. He initially got 3 score points (30% correct) and 5 points (50%) in text 1 and 2. He further advanced to 6 points (60% correct) and 7 points (70% correct) in text 3 and 4. An exemplary test item in Adventure Tours for Charities is shown below.
4. What is NOT mentioned as the way for people to raise money?
(A) Some people ask for help from friends.
(B) Some people ask the party guests pay for attending the parties.
(C) One worker sells his car.
(D) One farmer sells his one-meter square of field.

Since the less-proficient reader successfully figured out the demonstrative references referring to the farmer (see graph (c) of Figure 9). He had better understanding of the textual information related to the farmer. He could select the correct answer, option D, and further abandoned the incorrect answer, option C. Another exemplary test item in Traditional Markets and Modern Markets is shown as follows.

6. What would you suggest people who want to open a new store in Brazil bear in mind?
(A) They should remember the names of the shoppers.
(B) They should make friends with the sellers at the ferias.
(C) They should remember the names of the streets in Brazil.
(D) They should know that shoppers might still prefer the ferias.

The less-proficient reader correctly linked the demonstrative references like the shopper (29) and the store (45) (see graph (d) of Figure 9) to correct subjects. He eventually integrated the textual information and enhanced his comprehension. As a result, he selected the correct option, namely, option D.

Conclusion

Some conclusions can be gleaned from the result of this study. First, the incorporation of computer system in instruction supported the average and the less-proficient readers to develop their mental representation of reference when reading four sequential texts. It was found that the average reader initially did not integrate any reference when reading the first text. She did begin to link some references in a successive way in the second text. This also resulted in her performance in the reading comprehension test, 8 out of 10 points (80 % correct) in text 2. She even could find out the possible subjects that the references referred to. In the third text, she not only referred the references to a correct subject but also made these references mutually linked. In the final online text, she clustered the references in different parts of the text referring to a correct subject.

The developmental process of the average reader’s mental map in referential resolution is also true for the less-proficient reader. The less-proficient reader did not integrate any reference when he read the first online text. Instead, he referred each reference to its subject separately. In the third and fourth text, he started to connect some references to the correct subject in a successive way. In the process of developing his mental map of reference in interpreting textual information, the less-proficient reader was more aware of his reading process and actively asked for help than before. As a result, he was able to select a correct answer in the multiple-choice reading comprehension test. In the reading comprehension tests, the less-proficient reader initially got 3 score points (30% correct) and 5 points (50%) in texts 1 and 2. He further advanced to 6 points (60% correct) and 7 points (70% correct) in text 3 and 4. That is, the more the reader was engaged in comprehension monitoring in drawing their mental maps of referential resolution, the higher score he obtained in the reading comprehension tests.

Second, it was found that the keys to the development of readers’ mental representation of reference lay in whether the reader was actively engaged in comprehension monitoring and frequently asked for help. In the engagement of comprehension monitoring, the reader read and reread the sentence many times to make sure if his comprehension was coherent or not, such as adding or erasing a text element. This helped the reader monitor, regulate, and evaluate his own reading process and construct the meanings of the textual information.

Third, the request of feedback tool is essential to develop the reader’s mental representation of the reference. It assists the reader to grasp the main idea of the text as the reader tries to connect sentences together by reference. In this study, the more-proficient readers requested the feedback tool as they were not sure about the correct referential resolution. They depended on the feedback tool to confirm the choices they made. The average readers relied on the feedback tool when they encountered difficulty. They also requested the feedback tool to make sure if their selection was correct. In contrast, the less-proficient readers seldom requested the feedback tool. They very often failed to
monitor their reading process and hardly sought help from the feedback tool. Instead, they gave up reading the texts easily. It was found that the modeling and practicing instruction for using the feedback tool in the computer system was particularly necessary to the less-proficient readers.

Finally, the importance of the tracing and recording the reader’s reading behavior and process should be emphasized. The trace result provided by the computer system makes the intangible reading process visible to the teacher. Every reading action that the reader takes is recorded in the system. This enables the teacher to observe the difficulties that readers encounter and the performance among the students with various English reading proficiencies. Based on these information, the teacher could modify his follow-up instruction to help the reader overcome their difficulties and better develop their integrative skills in reading.

References


Appendix

Online reading text: Ideas about Beauty

Most people would agree that “beauty is in the eye of the beholder.” That is, everybody has a different measure of what (or who) is beautiful and what is not. And in fact, researchers in different fields have collected convincing evidence that even people from different cultures tend to rate beauty in much the same way. However, researchers do not agree on whether the factors which influence how most people judge beauty come from genetics (nature) or culture (society).

Devendra Singh, a psychologist at the University of Texas at Austin, conducted an experiment in 1993 to find out if different men found different female body shapes attractive. Dr. Singh gave drawings of different female body shapes to a variety of men and asked them to choose the most attractive body shape. Even though the men came from a wide range of cultural backgrounds, they all tended to rate the “hourglass” body shape as the most attractive. In fact, Dr. Singh found that any woman whose waist is 70% as wide as her hips is judged as attractive by most men no matter how big the woman is overall. Body shape, not weight, seemed to be viewed as the critical factor for attractiveness by men in this survey.

Dr. Singh explained this result from the perspective of evolution. Women who develop an hourglass shape have a relatively high level of estrogen, the female hormone. Because estrogen levels also influence fertility, men may subconsciously view these women as good candidates for producing children. Therefore, according to Dr. Singh, the men choose this type of women who have the potential for having more children. Over time, evolution would favor men who have inherited genes from their fathers which influence the selection of this type of fertile woman.

Douglas Yu, a biologist at Imperial College in London, has a different theory about men’s ideas of beauty. Dr. Yu thinks that culture, especially culture developed through exposure to the media, has had the largest influence on how men judge beauty. In order to test this, Dr. Yu traveled to southeast Peru to interview men in an isolated community far from the reach of modern television, movies, and magazines. Through his own survey, Dr. Yu found that the men in this isolated community preferred heavier women with wider waists, and not particularly women with “hourglass” shapes. Because this small community has lived apart from western mass communication, their own culture has not been influenced by outside standards of beauty. Dr. Yu points out that this group has experienced the same genetic evolution as all humans do, but a different standard.

In order to check the reliability of his study, Dr. Yu surveyed two other groups of men from this same community. However, the second and third groups surveyed by Dr. Yu had had more exposure to Western media. The results of these later surveys showed that as men from this isolated community came into contact with Western media, their standards of beauty began to change more toward the Western standard of beauty. Dr. Yu concluded from these findings that even if evolution played a part in men’s selection of mates, cultural influences were more powerful in the end and worked faster in changing men’s standards.

With both satellite communication and the Internet broadcasting images and information, globalization has become almost impossible to avoid. It is becoming harder and harder to find isolated communities like the one surveyed by Dr. Yu. The genetics vs. culture debate may soon become irresolvable simply because there will be no uninfluenced groups left to ask.
Attitudes and Satisfaction with a Hybrid Model of Counseling Supervision

Steven R. Conn, Richard L. Roberts and Barbara M. Powell
Department of Counseling and Student Development, Eastern Illinois University, 600 Lincoln Ave, Charleston, IL 61920, USA // srconn@eiu.edu // rlorberts@eiu.edu // bmpowell@eiu.edu

ABSTRACT
The authors investigated the relationship between type of group supervision (hybrid model vs. face-to-face) and attitudes toward technology, toward use of technology in professional practice, and toward quality of supervision among a sample of school counseling interns. Participants (N = 76) experienced one of two types of internship supervision: a hybrid model (N = 41) or face-to-face (N = 35). Data analyses indicated that the hybrid model of group supervision was positively related to attitudes toward technology in counselor education, future professional practice, and the overall supervisory experience. Further, differences between the approaches in delivery of supervision showed no effect on perceptions of quality of supervision. Implications for extending the use of technology-mediated supervision to practicing professionals are presented.

Keywords
Hybrid model, Technology, Counseling supervision, School counselors

The study of distance learning is a major focus in higher education (Anakwe, 1999; Benigno & Trentin, 2000; Davies & Mendenhall, 1998; DeBourgh, 1999) and an emerging concern among counseling related programs (Alterkruse & Brew, 2000; Bobby & Capone, 2000; Kjosness, 2002). The body of research examining the general satisfaction with distance learning, although mixed (Salas, Kosarzycbi, Burke, Fiore, & Stone, 2002; Smith, 1999), does reflect a consensus about characteristics associated with student satisfaction (DeBourgh, 1999). These include clear course expectations, prompt response to student questions, encouragement of student participation, use of varied instructional techniques, access to the instructor, and timely feedback to students about their work. However, due to limited research, there is less agreement about what characteristics are associated with satisfaction among counselors-in-training concerning technology-mediated supervision (Janoff & Schoenholtz-Read, 1999).

In spite of the limited research, there is growing support for the use of technology in training and supervision (Alpert, 1986; Casey, Bloom, & Moan, 1994; Christie, 1998; Lambert, Hedlund, & Vieweg, 1990; Myrick & Sabella, 1995). Olson, Russell, and White (2001) suggested using technology in supervision to meet the need for outreach to rural areas, for faculty who have limited time to supervise face-to-face, for increasing students’ access to qualified supervisors, and to manage the cost of supervision. Other advantages include removal of time and space restrictions, more time to reflect on information, and a permanent record for later reflection (Hara, Bonk, & Angeli, 2000).

Although there is research on various aspects of technology and supervision (Gamon, Sorlie, Bergvik, & Hoifodt, 1998), much of the investigation has focused on the efficacy of email. In an early study of its use, Myrick and Sabella (1995) suggested that email has a place in supervision by providing students with multiple opportunities for feedback and enhanced reflection. Other researchers found that email is a useful supplement to traditional modes of supervision (Olson, Russell, & White, 2001), increases personal reflection (Clingerman & Bernard, 2004), and “offers a way to know students’ thought processing and development at a level not before practically feasible” (Graf & Stebnicki, 2002, p. 48).

Researchers have also identified several limitations to using technology in supervision. These include concerns over variations in levels of computer skills among users, loss of non-verbal information, limited bonding between supervisor and student, slow response time, lack of confidentiality, and slow band-width speed, to name a few (Hara, Bonk, & Angeli, 2000; Janoff & Schoenholtz-Read, 1999; Myrick & Sabella, 1995; Olson, Russell, & White, 2001). However, Gamon, et al. (1998) found that, paradoxically, the forced limitations of technology-mediated supervision increased the development of insights and communication and enhanced the quality of supervision.

The nature of technology mediated communication is evolving from Web 1.0 tools such as basic email, chat room, threaded discussion, instant messaging and interactive video to Web 2.0 tools such as Skype, blogs, social networking, Wikis, podcasts and Folksonomy (collaborative tagging). However, faculty who are adopting technology into their instruction prefer the hybrid model, defined as a combination of face-to-face meetings and...
technology delivered instruction (Young, 2002). In a review of the research on computer-mediated communication, Janoff and Schoenholtz-Read (1999) suggested that a hybrid model of supervision offers benefits such as 1) access to peer and expert supervision when not meeting face-to-face, 2) access to information supplied by the supervisor when not meeting face-to-face, 3) opportunity for equal and evolving participation among all members, and 4) opportunity for other members to observe interactions between supervisor and supervisee.

Although there are various models of supervision that guide the process, most of the research has focused on practice in a clinical setting (Bernard & Goodyear, 2004). Nelson and Johnson suggested that supervision of school counselors should focus primarily on skill building. Others have suggested that supervision should also include consultation (Kahn, 1999) and be organized around several primary functions: clinical, developmental, administrative, and peer supervision (Barret & Schmidt, 1986; Henderson, 1994). We have chosen the combination of these recommendations to guide our research.

Like technology-mediated supervision, there is limited research concerning the use of supervision among school counselors (Kahn, 1999) due to 1) the multiple roles school counselors play compared to counselors in clinical settings, 2) the lack of formal training among school supervisors (Nelson & Johnson, 1999), and 3) uncertainty in terms of focus. Therefore, more investigation is needed to guide the growing support of supervision with Web 1.0 and 2.0 tools that include the use of a hybrid model. Our investigation compared two groups of school counseling interns. One group experienced a hybrid model of supervision which included face-to-face, email, and live chat room organized around the ideas proffered by Henderson (1994), Barret and Schmidt (1986), and Kahn (1999); the other group experienced only face-to-face supervision. We investigated three research questions. The first question focused on the degree to which use of the hybrid model of supervision was positively associated with attitudes toward the use of technology in counselor education. Students who classify themselves as “high computer users” have been shown to have more positive attitudes toward technology (Hayes & Robinson, 2000). Thus, our first research hypothesis was: the group experiencing the hybrid model of supervision will report more positive attitudes toward use of technology in counselor education than the face-to-face group.

The second research question focused on the degree to which use of the hybrid model of supervision was positively associated with attitudes about the use of technology in future professional practice. Prior research has suggested that technology is not a neutral influence on the user and may have an impact on future practice (Barnard, 1997). Thus, our second research hypothesis was: interns who participate in the hybrid model of supervision will report more positive attitudes toward use of technology in future professional practice than the face-to-face group.

The third research question focused on the degree to which the two groups differed in their perceptions of the quality of supervision. Although the positive working relationship important to quality supervision is typically established face-to-face (Ladany, Ellis, & Friedlander, 1999), much of the research indicates students are equally satisfied with distance learning as they are with face-to-face learning (DeBourgh, 1999). Thus, our third research hypothesis was: the group experiencing the hybrid model of supervision will report no difference in satisfaction with quality of supervision than the face-to-face group.

Method

Participants

Participants were 76 graduate students enrolled in the School Counseling program at a mid-sized Midwestern University accredited by the Council for Accreditation of Counseling and Related Educational Programs (CACREP). The study was conducted during their first semester of a two-semester internship. All participants completed the survey at the end of the first semester of their school counseling internship. During the 3 years that data were collected, participants were unaware of the different models used in the research and were assigned to one of two groups based on their enrollment in either the on-campus section of internship which received face-to-face supervision or the off-campus section which received a hybrid model combining technology-mediated supervision with face-to-face meetings. Although a convenience sample was used, because of the method of assignment to experimental groups, there is no reason to believe there were prior differences in attitude toward technology among group members.
The technology-mediated groups, with 41 participants, received group supervision by means of a combination of online supervision using WebCT (10 meetings) and face-to-face meetings (5 meetings). The other groups, made up of 35 participants, received all of their supervision face-to-face. Surveys were given to participants at the end of their first semester of internship. In order to insure complete anonymity of participants, the surveys contained no identifying information.

Demographic data, such as gender, ethnic background, and professional status were collected from participants. The majority of the participants were female (N=63), 12 were male, and one participant did not respond. Seventy-one participants were white (not of Hispanic origin), two were black (not of Hispanic origin); and three did not respond. Due to the homogenous sample, the demographics were not factored in the statistical analyses.

**Instruments**

The Supervisory Working Alliance Inventory: Trainee Form was developed to measure the relationship between supervisor and trainee on two subscales, Rapport and Client Focus (Efstatio n, Patton, & Kardash, 1990), and supported in research on supervision (Chen and Bernstein, 2000; Patton and Kivlighan, 1997). There were 19 items on this inventory using a 7-point Likert scale ranging from 1 = “almost never” to 7 = “almost always.” The Rapport subscale consists of 12 items and the Client Focus subscale is made up of 7 items. Trainees were asked to indicate the frequency of specific characteristics (for example “I feel comfortable working with my supervisor”) within the supervisory relationship and to rate accordingly. Confirmatory factor analytic techniques supported the three-factor and two-factor structures for the supervisor and trainee version of the SWAI (Gold, 1993). Further, convergent and discriminant validity for the supervisee form of the SWAI was supported in statistically significant correlations with scales of the Supervisory Styles Inventory (Friedlander & Ward, 1984) and the Personal Reactions Scale – Revised (Holloway & Wampold, 1983). Cronbach’s alpha indicating internal consistency was calculated for the Trainee scales with alpha coefficients of .90 for Rapport and .77 for Client Focus (Efstation, Patton, & Kardash, 1990).

The Supervision Questionnaire was developed to measure the trainee’s satisfaction with their supervisory experience (Ladany, Hill, & Nutt, 2004; Larsen, Attkisson, Hargreaves, & Nyguyen, 1979). The questionnaire consisted of eight items using a 4-point Likert scale ranging from 1 = “excellent” to 4 = “poor.” This survey includes items such as “How would you rate the quality of the supervision you have received?”, and “Did you get the kind of supervision you wanted?” Originally derived from the Client Satisfaction Questionnaire (Larsen et al., 1979), this questionnaire was found to be related to supervisee nondisclosure with previous research reporting the SSQ to be alpha = .96 (Ladany et al., 2004). In a study by Ladany, Lehrman-Waterman, Molinaro, and Wolgast (1999), internal consistency was reported as alpha = .97.

The Web-Based Distance Group Satisfaction Survey (Roberts, Powell, & Fraker, 2002) used a 5-point Likert scale ranging from 1 = “strongly agree” to 5 = “strongly disagree.” The survey is made up of 3 subscales: 1) Perceptions of the Usefulness of Technology in Professional Practice (8 items), 2) Perceptions of the Usefulness of Technology in Counselor Education (6 items), and 3) Perceptions of the Effectiveness of Online Supervision (16 items). Participants were asked to indicate the answer that most likely fit their current thoughts or feelings on the topic. This unpublished 33 question survey (Roberts et al., 2002) was designed using a pilot group (N= 43) to measure students’ perceptions of the effectiveness of online supervision (Cronbach’s alpha = .91), students’ perception of the usefulness of technology in professional practice (Cronbach’s alpha = .63), and student’s perceptions of the usefulness of technology in counselor education (Cronbach’s alpha =.82).

**Procedure**

All participants were enrolled in a semester long internship (15 meetings) in school counseling in which they received group supervision. Data were collected over a three year period. Two of the researchers using identical syllabi taught both the technology-mediated and face-to-face sections over the course of data collection. Students were required to email “cases” to the faculty member and classmates one week prior to discussion to allow for advanced reflection. Students in the technology-mediated group used either WebCT “live chat” or met face-to-face to discuss the cases. In the hybrid model section, the class began with a face-to-face meeting followed by
approximately 2 technology-mediated classes for every face-to-face class. In total, there were 5 face-to-face meetings and 10 technology-mediated meetings.

Data Analyses

To test the three working hypotheses, scores were compared between the online and face-to-face supervision groups. ANOVAs were performed to analyze the relationship between the hybrid model of supervision and attitudes toward current and future practice using technology. In addition, ANOVAs were also performed to analyze the relationship between the hybrid model of supervision and attitudes toward the quality of supervision. Following the ANOVAs, power analyses were conducted.

Results

Table 1 contains the descriptive statistics for the present study. All instruments administered in the present study were found to have moderate-to-high reliability estimates, ranging from .66 to .93 for alpha coefficient estimates and ranging from .71 to .90 for split-half reliability estimates. Substantially higher correlations among the respective subscales of the Working Alliance measure and the Satisfaction Survey over the correlations with other measures, including the Supervision Questionnaire, suggested that the three constructs under investigation (i.e., satisfaction with the quality of supervision, working alliance, and satisfaction with the use of technology) were relatively independent of one another.

Table 1. Descriptive Statistics (N = 76)

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>sd</th>
<th>α</th>
<th>Split Half†</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>1. Supervision Quest.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working Alliance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Rapport</td>
<td>6.4</td>
<td>0.64</td>
<td>.93</td>
<td>.88</td>
<td>.26</td>
</tr>
<tr>
<td>3. Client Focus</td>
<td>7.2</td>
<td>0.93</td>
<td>.89</td>
<td>.90</td>
<td>.39*</td>
</tr>
<tr>
<td>Satisfaction Survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Subscale 1</td>
<td>54.7</td>
<td>11.73</td>
<td>.90</td>
<td>.90</td>
<td>-.26</td>
</tr>
<tr>
<td>5. Subscale 2</td>
<td>10.9</td>
<td>3.60</td>
<td>.77</td>
<td>.82</td>
<td>-.20</td>
</tr>
<tr>
<td>6. Subscale 3</td>
<td>11.0</td>
<td>3.14</td>
<td>.66</td>
<td>.71</td>
<td>-.16</td>
</tr>
</tbody>
</table>

† = split half reliability estimates were corrected for length using the Spearman-Brown formula.

*p < .05 (2-tailed); ** p < .01 (2-tailed).

Analysis of Variance results are presented in Table 2 addressing each of the three research questions under investigation. Our first research question focused on the degree to which use of the hybrid model of supervision was positively associated with attitudes toward use of technology in counselor education. Subscale 3 of The Web-Based, Distance Group Satisfaction Survey (Roberts, et al., 2002) was used to determine perceptions of usefulness of technology in counselor education. A one-way ANOVA was computed comparing the means of the hybrid model of supervision group (M = 9.46, SD = 2.76) and the face-to-face supervision group (M = 12.71, SD = 2.90). The analysis indicated that there was a significant difference between the means of the two groups (F(1,74) = 25.06, p < .001, η = .50). Thus, our hypothesis was supported, indicating that use of the hybrid model of supervision was in fact positively related to attitudes toward use of technology in counselor education.

Our second research question focused on the degree to which use of the hybrid model of supervision was positively associated with attitudes toward use of technology in future professional practice. Subscale 2 of The Web-Based, Distance Group Satisfaction Survey (Roberts, et al., 2002) was used to determine perceptions of usefulness of technology in future professional practice. The one-way ANOVA computed indicated a significant difference between the hybrid model of supervision group (M = 9.56, SD = 2.72) and face-to-face supervision group (M = 12.71, SD = 3.16) (F(1,74) = 21.87, p < .001, η = .48). Thus, our hypothesis was supported, indicating that use of the
hybrid model of supervision was positively associated with attitudes toward use of technology in future professional practice.

Table 2. Analysis of variance results, including Cohen’s effect size (ES) F-values

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
<th>η (η²)</th>
<th>p</th>
<th>ES(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web-Based, Distance Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction Survey: Subscale 1</td>
<td>1/74</td>
<td>28.94</td>
<td>.53 (.28)</td>
<td>.001</td>
<td>.61</td>
</tr>
<tr>
<td>Satisfaction Survey: Subscale 2</td>
<td>1/74</td>
<td>21.87</td>
<td>.48 (.23)</td>
<td>.001</td>
<td>.55</td>
</tr>
<tr>
<td>Satisfaction Survey: Subscale 3</td>
<td>1/74</td>
<td>25.06</td>
<td>.50 (.25)</td>
<td>.001</td>
<td>.58</td>
</tr>
<tr>
<td>Supervisory Working Alliance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subscale 1: Rapport</td>
<td>1/65</td>
<td>1.08</td>
<td>.14 (.02)</td>
<td>.30</td>
<td>.15</td>
</tr>
<tr>
<td>Subscale 2: Client Focus</td>
<td>1/65</td>
<td>2.48</td>
<td>.20 (.04)</td>
<td>.12</td>
<td>.20</td>
</tr>
<tr>
<td>Supervision Questionnaire</td>
<td>1/74</td>
<td>.08</td>
<td>.03 (&lt;.01)</td>
<td>.78</td>
<td>.03</td>
</tr>
</tbody>
</table>

Our third research question focused on the degree to which the two groups differed in their perceptions of the quality of supervision. We predicted that use of the hybrid model of supervision would result in no significant difference between groups on perceptions of quality of supervision. Subscales 1 and 2 of the Supervisory Working Alliance Inventory: Trainee Form (Effstaton, Patton, & Kardash, 1990), the Supervision Questionnaire (Ladany, et al., 2004), and Subscale 1 of The Web-Based Distance Group Satisfaction Survey (Roberts, et al., 2002) were used to determine attitudes toward quality of supervision.

The one-way ANOVA computed on the Rapport subscale of the Supervisory Working Alliance Inventory: Trainee Form indicated no significant difference between the hybrid model of supervision group ($M = 6.53, SD = 0.58$) and face-to-face group ($M = 6.37, SD = 0.69$) groups ($F(1,65) = 1.08, ns, η = .14$). Thus, our hypothesis was supported, indicating that participants in the hybrid model of supervision did not significantly differ from the face-to-face group in their perceptions of supervisory rapport.

The one-way ANOVA computed on the Client Focus subscale of the Supervisory Working Alliance Inventory: Trainee Form indicated no significant difference between the hybrid model of supervision group ($M = 6.39, SD = 0.66$) and face-to-face group ($M = 6.09, SD = 0.90$) groups ($F(1,65) = 2.48, ns, η = .20$). Thus, our hypothesis was supported, indicating that participants in the hybrid model of supervision group did not significantly differ from the face-to-face group in their perceptions of supervisory client focus.

The one-way ANOVA computed on the data collected from the Supervision Questionnaire indicated no significant difference between the hybrid model of supervision group ($M = 29.44, SD = 3.52$) and face-to-face group ($M = 29.20, SD = 3.87$) groups ($F(1,74) = .08, ns, η = .03$). Thus, our hypothesis was supported, indicating that students experiencing the hybrid model of supervision did not significantly differ from students experiencing face-to-face supervision in terms of satisfaction with supervisory experience.

The one-way ANOVA computed on the data collected from the Web-based, Distance Group Supervision survey indicated a significant difference between the hybrid model of supervision group ($M = 54.39, SD = 11.55$) and face-to-face group ($M = 68.94, SD = 11.99$) ($F(1,74) = 28.94, p < .001, η = .53$). Thus, our hypothesis was rejected using this subscale since results indicated that use of the hybrid model of supervision was in fact positively associated with attitudes toward the quality of supervision.

Effect Size and Power Analyses

Traditionally, statistical analyses of empirical findings test the proposition that the phenomenon under investigation is either present or not in the population (Cohen, 1988). This is accomplished by testing the null hypothesis positing that the phenomenon does not exist in the population, after which it is rejected or not. However, such results say little about the actual results found in the study. For instance, if the null hypothesis is rejected, are the “significant” findings small, medium, or large? Moreover, what is the likelihood that the results found in one sample would subsequently be found in other samples from the same population? These questions can be addressed by the investigation of a nonzero effect size and power analyses.
In the present study, Cohen’s (1988, pp. 274-288) $f$-value was used as the effect size index; $f$-values were determined using Cohen’s tables. An effect size was evaluated as follows: $f < .10$ as no effect, $10 \leq f < .25$ as a small effect, $0.25 \leq f < .40$ as a medium effect, and $f \geq .40$ as a large effect size. Each of the three Satisfaction Survey subscales yielded large effect size values (i.e., .61, .55, and .58, respectively). Both Supervisory Working Alliance subscales (i.e., Rapport and Client Focus) produced small effect size values (i.e., .15 and .20, respectively), and the Supervision Questionnaire yielded no effect size (i.e., .03).

Power analyses were conducted only for large effect-size findings (i.e., Satisfaction Survey subscales 1, 2, and 3). To minimize the capitalization on chance due to the multiple statistical tests, power analyses were conducted at the Type I error rate of $\alpha = .01$. Results of the power analyses showed a very high power for all three subscales (i.e., power exceeding .995). Thus, for each Satisfaction Survey subscale, one would expect to reject the null hypothesis in 99 out of 100 random samples from the present population studied.

Discussion

Results of the study indicated that the technology-mediated group had more positive attitudes about the use of technology in counselor education than the face-to-face group (hypothesis 1). These results support earlier research (Hayes & Robinson, 2000) which found that “high computer users” were shown to have more positive attitudes toward technology than those not classified that way. As mentioned in the procedure section, the hybrid model consisted of a 2:1 ratio between technology-mediated and face-to-face supervision giving students consistent experience over time with technology. The fact that there was consistency may have alleviated some of the limitations of the use of technology in supervision identified by Olson, Russell and White (2001). For instance, slowness in responding to supervisees’ consultation questions becomes a moot point when groups meet weekly. Moreover, another limitation, that computer-mediated communication can overwhelm the student and instructor with endless opportunities to interact (Hara, Bonk, & Angeli, 2000), was not experienced by the technology-mediated group. This may have been a result of specific time frames used for both face-to-face and technology-mediated chat room discussions.

Similarly, hypothesis 2 was supported; that is, the hybrid model of supervision was positively associated with attitudes toward use of technology in future professional practice. The benefits of the hybrid model, which includes opportunity for equal and evolving participation among all members (Janoff & Schoenholtz-Read, 1999) may have influenced the technology-mediated group to envision themselves as users of technology in the future. Although the face-to-face group also experienced equal and evolving participation, they did not use technology to do so; thus any attitude toward technology in future professional practice would not be substantiated by actual experience. Also, the technology-mediated group had the opportunity to experience the development of insights and communication as a result of the hybrid model of supervision (Gamon, Sorlie, Bergvik, & Hoifodt, 1998), and most likely assumed that this type of communication would be possible in the future.

Hypothesis 3 was also supported. With the exception of one subscale, there was no significant difference between groups on perceptions of quality of supervision. This finding supports DeBourgh’s (1999) research which indicated that the satisfaction of students experiencing distance learning was not significantly different from students receiving face-to-face learning. Moreover, it appears that the results found in the Roberts et al. (2004) subscale substantiates earlier research (Gamon, Sorlie, Bergvik, & Hoifodt, 1998) which found that the hybrid model of supervision actually may enhance the quality of supervision.

Given these findings, the hybrid model of supervision may be the answer to the need for outreach to rural areas and for increasing students’ access to qualified supervisors (Olson, Russell, & White, 2001). Likewise, adopting a hybrid model of supervision appears to address one the limitations, namely, that technology-mediated relationships can take longer to form than face-to-face relationships (Hara, Bonk, & Angeli, 2000; Myrick & Sabella 1995). In acknowledging that a positive working relationship is critical (Ladany, Ellis, & Friedlander, 1999), we suggest that the initial meeting be held face-to-face, which will enable the students to meet and interact with the supervisor. At the same time, students will be able to begin their own bonding process with each other. The initial meeting can also address other identified limitations such as confidentiality and ethics, particularly as they relate to supervision via the internet. Moreover, these topics can be discussed not only during face-to-face meetings, but also in the weekly chat–room discussions.
Our research indicates that the hybrid model of supervision is a positive experience for students. But it will likely become a negative experience if students are not trained adequately or if the supervisors lack access to computer support (Janoff & Schoenholtz-Read, 1999). We found that after students receive information about the chat-room process (during the first meeting), it is important to hold a “practice” session, where students and supervisor meet informally. This allows both supervisor and students to confirm that the process is working and allows them to address any technical difficulties before supervision occurs. Also, it is necessary for the supervisor to have computer support in case any problems arise during the live chat-room time. All of these necessary arrangements must be made before the first face-to-face meeting so that the technology aspect of the chat runs as smoothly as possible and the focus can remain on supervision.

As with any study, ours has limitations. The Midwestern location and homogeneity of participants may limit the applicability of the research to other areas. Two of the surveys are unpublished and although pilot studies were initially done on the Web-Based Distance Group Supervision Survey (Roberts, et al., 2002) more work is needed to ensure its validity and reliability. Future studies of the hybrid model of supervision might include measures of satisfaction among practicing school counselors. It may be that school counselors in rural areas will feel less alienated from colleagues if they participate in a weekly chat.

Conclusion

One of the important findings of this study is that school counseling interns who experienced technology-mediated supervision were more satisfied with their experience than were the interns who met face-to-face with their supervisor. Additionally, interns in the technology-mediated group were more likely than the face-to-face group to have a positive attitude toward the use of technology in professional practice. Counselor Educators can use the hybrid model of supervision to show students how they can seek supervision or consultation about the varied kinds of issues they will be facing when they are working as school counselors. For example, our hybrid model of supervision includes discussion of cases and situations similar to the primary functions suggested in previous research: clinical, developmental, and administrative tasks (Barret & Schmidt, 1986; Henderson, 1994) as well as case conceptualization skills (Butler & Constantine, 2006).

These findings have implications for school counseling interns as well as practicing professionals. This not only includes school counselors, but also other fields such as health (e.g., nursing, medical and psychology training) and education (e.g., student teaching). The importance for supervision has been recognized by school counselors and other professionals. It is particularly important that these professionals utilize emerging technologies including Web 2.0 tools for supervision, especially for those who live in rural areas and have little access to colleagues. However, communication in a cyber environment is not without problems. The hybrid model used in this study addresses the concerns of some researchers about the difficulty of establishing the interpersonal connectedness required for effective communication absent face-to-face communication (Wilczenski & Coomey, 2006). Thus, if practicing professionals who utilize supervision make use of technology for establishing connections, they might be better served if they also included occasional face-to-face meetings.

References


Christie, B. S. (1998). Distance clinical supervision in cyberspace: A qualitative study, Oregon State University, USA.


Modeling Academic Education Processes by Dynamic Storyboarding

Yoshitaka Sakurai¹, Shinichi Dohi¹, Setsuo Tsuruta¹ and Rainer Knauf²

¹School of Information Environment, Tokyo Denki University, Chiba, 270–1382, Japan // Tel: +81 476 46 8491 // Fax: +81 476 45 8449 // ysakurai@sie.dendai.ac.jp // dohi@sie.dendai.ac.jp // tsuruta@sie.dendai.ac.jp
²Faculty of Computer Science and Automation, University of Ilmenau, 98684 Ilmenau, Germany // Tel: +49 3677 69 1445 // Fax: +49 3677 69 1665 // rainer.knauf@tu-ilmenau.de

ABSTRACT
In high-level education such as university studies, there is a flexible but complicated system of subject offerings and registration rules such as prerequisite subjects. Those offerings, connected with registration rules, should be matched to the students’ learning needs and desires, which change dynamically. Students need assistance in such a maze of dynamically changing opportunities and limitations. To cope with this problem, a new storyboard concept for academic education, called “dynamic storyboarding” is proposed to assist university students. Dynamic storyboarding is based on the idea of semi-formally representing, processing, evaluating, and refining didactic knowledge. This storyboarding is more appropriate in managing high-level education than is general artificial intelligence knowledge representations such as frames. This is because the structure of dynamic storyboarding is driven by the semi-formal and multilayered nature of didactic knowledge in university education. A feasibility study showed that storyboarding can be used to supplement an academic educational system, such as the dynamic learning need reflection system (DLNRS) of Tokyo Denki University (TDU) in Japan. Concretely speaking, didactic knowledge in the university curricula was proven to be easily and clearly represented by dynamic storyboarding. This contributed to the students’ dynamic learning activities by supporting features that help students review and adapt their own individual curricula.

Keywords
Process modeling, Learning processes, Storyboarding, Dynamic learning need reflection system, Knowledge engineering

Introduction

University studies are characterized by a high degree of flexibility with respect to the subjects to be included in a student’s curriculum. On the other hand, there are complex and dynamic rules of subject registration that can constrain the student’s selections. Some of these rules are as follows:

- when subject registration can be made
- how to register (in-person, online, via post)
- department permission for particular subjects or courses
- limitations due to class capacity and time conflicts among classes
- particular subject or course combinations necessary for graduation
- prerequisites

Each of those subjects and their registration rules should be matched with students’ learning needs and desires, which change dynamically depending on factors such as the learning status attained in the previous semester. Under these circumstances, it is not easy to finish a study plan or curriculum in time for registration.

The flexibility and complexity of regulations increase students’ burden in planning their university studies. The authors’ experiences with Japanese, German, and US universities indicate that this may be a world-wide problem.

Generally, university students suffer from a lack of clarity due to the above-mentioned complex situation. A significantly high number of students do not succeed in their desired subjects or courses because they cannot comply with some of these regulations or do not even know about them. Therefore, students sometimes cannot finish their studies in the designated time. Avoiding the resulting frustration is one objective of introducing the new approach of storyboarding as a means to keep track of the big picture in the maze of dynamically changing opportunities and restrictions mentioned above. University students need assistance in this maze.

Flexibility has both advantages and disadvantages. Being flexible helps students in their learning and, as long as the flexibility doesn’t threaten their success, increases students’ motivation. Increasing student motivation is the other
goal of using storyboards. But, this advantage can turn out to be a problem when students are unable to meet the requirements because they violated the complex and dynamic rules/constraints mentioned above. Helping students meet the requirements without such violation is the other objective of the proposed storyboard application.

In contrast to basic-level education, such as in primary and secondary schools, higher education is characterized by

- a large variety of opportunities for creating the academic curricula and timelines, and
- teachers (professors and tutors) who are experts in their subjects but do not necessarily have the didactic skills to teach.

Particularly due to the first point, at the School of Information Environment (SIE) of Tokyo Denki University (TDU), students can be more flexible than students of other schools or universities in designing their studies according to their needs, wishes, interests, and talents. Thus, students at SIE of TDU need more qualified assistance or guidance in the maze of dynamically changing opportunities.

To address the above-mentioned problems, an education system, which we currently call dynamic learning need reflection system (DLNR system or DLNRS), was developed and introduced at SIE of TDU (Dohi & Nakamura, 2003; Dohi, Nakamura, Sakurai, Tsuruta, & Knauf, 2006a). Its objective is to maintain or increase the students’ motivation by clarifying and dynamically reflecting students’ individual learning needs. The system is characterized by

- the abolition of the traditional rigid academic year
- the introduction of prerequisite conditions instead of a fixed, pre-determined subject sequence
- the replacement of fixed yearly tuition by a subject-based payment system
- a grade point average (GPA) system that evaluates the learning results and helps create or modify the upcoming class schedule, namely, the next semester’s curriculum, in deriving appropriate results or consequences.

As shown in Dohi & Nakamura (2003), the introduction was revolutionary and a remarkable success. In a questionnaire that was filled out by 203 students, almost 90 percent said that the DLNRS is very understandable and useful for both creating each semester’s schedule and making a long-term graduation plan. However, the level of understanding of the prerequisite conditions was about 60 percent. Thus, the method of displaying the prerequisite conditions needs to be improved, according to Dohi and Nakamura (2003).

Qualified guidance for factors such as the prerequisite conditions mentioned above needs adaptation with respect to varying learning needs, context conditions (the students’ performance in the previous stages of the study), and the students’ educational history. Such adaptation, however, presumes an anticipation of various structurally complicated alternatives and their explicit representation. Didactic variants have to be subject to discussion and, more systematically, subject to quality assurance. For this purpose, an appropriate didactic design methodology and tool needs to be established.

The current state of affairs in learning in general and in e-learning in particular shows some obvious reluctance to such didactic design. Scientists discuss and learners complain about the insufficient learning adaptability offered to the learners’ needs (Schulmeister, 2003).

The employment of dramaturgy and storyboarding seems promising in moving toward qualified guidance for the above didactic design. This is because storyboarding is considered to express the process flow through structured diagrams and, therefore, seems to simplify the representation of structurally complicated alternatives.

However, the current employment of such representations is characterized by misunderstandings. The so-called storyboard concepts in use (Meissner, 2003) are mostly substitutes for technological documents of high-level design, but are not specific to the instructional design process (Gagne, Briggs, & Wager, 1992; Rothwell & Kazanas, 2004). Didactic concepts (Flechsig, 1996; Jank & Meyer, 2002) are not made explicit and, therefore, modeling, evaluating, and refining such didactic issues are not sufficiently enforced. Even quite recent approaches (Schewe & Thalheim, 2004) could not go beyond IT systems. That is, they could not sufficiently consider handling educational contents such as didactics or educational knowledge.

There are contrasting approaches (Forsha, 1994) that are conceptually very useful, but syntactically too different from a workflow directed to technology-enhanced learning implementations. Forsha’s storyboards stick to the classic
booklet form of storyboards. They are a linear sequence of pages (called “panels” in Forsha, 1994). Although each page is quite flexible in terms of its content, there is no way to clearly represent and anticipate alternative paths and to use a hierarchical structuring in Forsha’s storyboards to support the user’s overview of the general structure of the plan of study and alternative choices in it. Purely software-driven concepts avoid such drawbacks. However, they do not provide an opportunity to represent and discuss details of more informal human learning (Bransford, Brown, & Cocking, 2000; Damasio, 1999). Learning is much more than logically formalized memorizing. “Learning imposes new patterns of organization on the brain, and this phenomenon has been confirmed by electro-physiological recordings of the activity of nerve cells” (Bransford et al., 2000, p. 121).

Learning is reasonably understood as an interactive knowledge construction process. Illustrative case studies are discussed in much detail in Davis, Sumara, and Luce-Kapler (2000). Chapter 3B, “Organizing Shapes,” for instance, nicely reports a lengthy process of conversation and co-operation between a teacher and his/her students in which a variety of media types, forms of interaction, and learners’ activities are dovetailed.

Didactic design implies the anticipation of those teacher-learner communication processes (Flechsig, 1996), and storyboards may provide the expressive power suitable to the didactic design and implementation in the learning processes. This, however, needs to go beyond the limits of software systems specification — the crucial question for innovations in didactic design.

The new storyboarding approach adopted here (Jantke & Knauf, 2005; Knauf, Sakurai, & Tsuruta, 2007) avoids the drawbacks and enjoys the advantages of both kinds of approaches (Forsha-like [1994] and technologically driven approaches). Such storyboarding, as introduced in Jantke & Knauf (2005) and Knauf et al. (2007), is a very general concept. In the context of DLNRS, this storyboarding is expected to complement the currently existing system such as the dynamic syllabus (DS) system or tool.

To illustrate the dependence between the subjects (courses) as well as to propose useful and efficient timetables with respect to the students’ needs, this storyboard concept is appropriate.

There is a maze due to dynamically changing opportunities, restrictions, and constraints. These opportunities, restrictions, and constraints are difficult to meet and their changes are difficult to predict. At some point of complexity, humans are not able to understand the rules as characterized above, which are driven by various factors, such as
- topical prerequisites, i.e., subjects to complete before taking new ones
- mental prerequisites, such as the student’s maintaining concentration throughout the class or at certain times of the day
- performance requirements, such as a high GPA,
- personal aspects (needs, wishes, talents, aims)
- organizational aspects (schedules without conflicts in time or location, and without undesired time gaps)
- resource management, including money, time, necessary equipment, and textbooks.

However, many explanations of such factors (on a web page, in a textbook, in the schedule) are quite informal and do not really help students to manage a program of study in a way that meets all regulations and the students’ learning needs.

In the authors’ universities (in Japan and Germany), students often study very ineffectively (too many semesters, few related subjects, and so on) or sometimes even fail in their desired academic career because of not keeping an accurate plan of study and/or not respecting regulations. To make choices in a legal and optimal way is difficult without a method such as this newly proposed storyboarding.

Although this storyboarding application was originally created to introduce e-learning systems, its reach goes far beyond the limits of current e-learning systems. The concept is simple as well as general and allows for modeling the didactics of any learning activities. New storyboards model the didactics behind learning in general and provide appropriate paths in a system of nested graphs, depending on the students’ needs. Dynamic storyboarding presented in this paper is expected to contribute to the learning in a wider context than to just a particular subject. This is because it can help students to keep track of both their program of study and the constantly changing opportunities and limitations in university education.
In general, learning activities need to be composed and designed at different levels of detail. For example, a lower level is the design of each lesson or, even lower, the discussion about a particular problem. On the other hand, the storyboard introduced here is of a higher level of detail and covers a complete/whole study in university education.

The larger the scope of learning, the more involved human activities are in the field of management. In other words, by enlarging the scope, the subject modeled by storyboards is extended from just didactics to didactics and management. For example, shifting from a difficult theory to a small example for its explanation is a tactic called didactic decision that can be easily made by the teacher without worrying about resources or about the cooperation of other individuals outside the current situation. On the other hand, inserting an additional subject in a curriculum can have a greater impact on resources such as time and cost. The latter adds strategic or management issues to the tactical or purely didactic (low-level or detail-level) ones in composing or modifying a detailed storyboard.

By adopting the new storyboard concept for a complete university study, even the management of the study becomes accessible for evaluation and refinement, that is, quality assurance.

As a deeper benefit of this work, some data mining can be performed on the paths of particular students after they have completed their studies at Tokyo Denki University. This will answer questions such as: What do the successful students’ paths have in common? How do the paths of successful students differ from those of less successful students? This will also allow future students to create curricula with an optimal chance of success.

Before showing how to adapt the new storyboarding concept for the intended application, we provide a short introduction to the DLNR system.

The paper is organized as follows: The next section describes the DLNRS as successfully introduced at SIE of TDU. This is followed by a section on the new storyboard concept as developed so far. After that, we introduce the enhancement of the storyboard concept to support the development of curricula and long-term graduation timetables for effective and target-oriented university studies at TDU. Finally, the suggested approach is summarized, conclusions are derived, and benefits are illustrated.

**Dynamic learning need reflection system**

Dynamic learning need reflection system (DLNRS) (Dohi & Nakamura, 2003) primarily aims at encouraging the students’ motivation by allowing students to create or modify their own class schedule each semester as well as their own graduation timelines. This develops a spirit of independence and keeps up with globalization. Key features of DLNRS are as follows:

- **Elimination of the traditional rigid academic year.** There is no academic year with fixed courses (here, a course is an ordered combination of subjects) and a fixed fee. Instead, there is a semester-based system where students can take their preferable subjects or change their course each semester, without restriction by rigid academic year, and pay the fee for each subject. The only restrictions to taking a particular subject in a particular semester are the prerequisites. Thus, the students are able to study at their own adaptive pace.

- **Elimination of compulsory subjects.** Compulsory subjects are replaced by the concept of prerequisite conditions. These conditions are expressed in two levels of recommendation: subjects that must be taken beforehand (prerequisites) and subjects recommended to be taken beforehand. The grades given to students in the prerequisite subjects are formally checked.

- **Replacement of a fixed charge per year by a subject-based payment system.** Students pay a subject-based fee in proportion to the number of units in each subject. Therefore, they carefully check their learning needs to pick the right subjects to achieve their academic goal at a minimum cost. Furthermore, a subject-based fee motivates them to make the maximum effort in order to pass the subject so that they don’t waste money.

- **Class length.** The usual length of a class is reduced from 90 minutes to 50 or 75 minutes. Typically, a subject is taught in three units, either as $3 \times 50$ minutes or as $2 \times 75$ minutes a week. The intended effect is that students will be able to concentrate the entire length of a class. This results in greater learning benefit from the subjects and, thus, more learning for the amount of money spent on a subject.

- **Grade Point Average (GPA).** This is a system to rate the learning results, especially to measure the learning quality that students attained. The GPA is calculated by equation 1, with $g_i$ being the points earned for a
particular subject, \( u \) being the number of units of the subject, and \( n \) being the number of subjects in the semester.

\[
(1) \quad GPA = \frac{\sum_{i=1}^{n} u_i g_i}{\sum_{i=1}^{n} u_i}
\]

According to equation 2, the number of grade points (GP) per subject ranges from 4 (> 80%) to 0 (< 40%).

\[
(2) \quad GP = \begin{cases} 
4, & \text{if performance = [80\%, 100\%]} \\
3, & \text{if performance = [70\%, 80\%]} \\
2, & \text{if performance = [60\%, 70\%]} \\
1, & \text{if performance = [40\%, 60\%]} \\
0, & \text{if performance = [0\%, 40\%]} 
\end{cases}
\]

The intention of this measure is that the maximum number \( (n_{\text{max}}) \) of units for which a student can register is controlled by the GPA of the previous semester as shown in equation (3).

\[
(3) \quad n_{\text{max}} = \begin{cases} 
25, & \text{if GPA} \geq 3.0 \\
21, & \text{if } 1.0 < \text{GPA} < 3.0 \\
12, & \text{if GPA} \leq 1.0 
\end{cases}
\]

The regulation demonstrated by equation 3 is a result of the experience with students who are not able to self-estimate their capacity. In the trade-off between a high learning quality, indicated by a high GPA, and a high learning quantity, indicated by a high number of units, some students tend to promote the latter at the expense of the former.

Obviously, this is not only a typically Japanese phenomenon, but is also apparent in German universities, judging from the experience of the German author, who reports of students who tend to just pass examinations but do not care about the learning quality.

This extrinsic motivation for examination results usually reveals itself when students feel that they will not necessarily need the subject’s topical content in their future career. Therefore, the DLNR contributes to avoiding this phenomenon by either making students open to recognizing the potential future need of the subject’s content or encouraging them to refrain from choosing this subject.

The introduction of the DLNRS at the SIE, is supported by a curriculum planning class, which aims to teach the student to develop an individual curriculum that meets his/her needs and desires by him/herself and a workshop that aims to develop an ambience of mutual trust between a professor and his/her students.

### Table 1. Unit composition for graduation

<table>
<thead>
<tr>
<th>Field of the subjects</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to SIE, computer literacy</td>
<td>2</td>
</tr>
<tr>
<td>Liberal arts subjects</td>
<td>40</td>
</tr>
<tr>
<td>Major subjects</td>
<td>60</td>
</tr>
<tr>
<td>Any subjects other than free subjects</td>
<td>22</td>
</tr>
<tr>
<td><strong>( \Sigma )</strong></td>
<td><strong>124</strong></td>
</tr>
</tbody>
</table>

The relationship among prerequisite conditions GPA and GP, the quantitative unit composition requirement for graduation as shown in Table 1, and other aspects are difficult to obtain and to keep track of. Instead of compulsory subjects, there are prerequisite subjects whose GP should not be E level (the lowest rating) for students to register in the related subject. Instead of the traditional rigid academic year, there is a unit composition regulation for graduation as shown in Table 1, where the total number of units that students can register in each semester is regulated based on their GPA of the previous semester, whether the number of units that students can get differs in each subject, and whether or not some subjects have prerequisites.
Furthermore, free subjects, such as those for qualifications to teach in high schools, don’t count for the 22 units in the last row of Table 1 because they don’t directly contribute to the career itself established by SIE but contribute to other qualifications instead.

To sum up, the development of class schedules and long-term graduation timetables is a quite challenging task.

Indeed, the GPA is a means to control the quantity (the total number of units) of subjects, but it doesn’t help the students to select appropriate subjects to achieve their individual goals under the terms of the complicated rules mentioned in the introduction.

On the one hand, there is a combinatorial explosion of opportunities for selecting subjects (e.g., 10 out of 100 possible ones for each class period or each semester). Yet there is a system of constraints that is driven by both topical and quantitative reasons.

Depending on the student’s individual academic history (subjects previously taken), the opportunities for a student’s academic future (the subjects in the upcoming semester) are narrowed. Here, as shown in Figure 2, the DS can indicate only a direct part of prerequisite subjects for which the prerequisite conditions are not directly met (constraint violation). Meanwhile, the storyboard application (presented later as “dynamic storyboarding”) helps keep track of the student’s program of study by marking all prerequisite subjects hierarchically, including consecutive prerequisite relations in which the prerequisite conditions are met in multiple stages. For example, prerequisite subjects for each of the subjects of a certain area of study are shown in a multistage hierarchy.

However, practice shows that there still remains a number of possible combinations even after masking the non-options, due to the prerequisite conditions. Fortunately, some combinations require a certain number of units beyond the number allowed by the GPA restriction. Here, again, the storyboard application helps by also masking the non-options resulting from the GPA regulations.

It is really difficult to keep track of the remaining choices in a huge catalogue of subjects after considering the individual combination of topical and quantitative restrictions. Therefore, a dynamic syllabus system has been developed, which supports the students in this complex task by way of the following four-step process:

1. acquisition of student data (model course, field after graduation, career goal, individual preferences)
2. selection of subjects
3. simulation of the schedule
4. registration for the subjects

Steps 2 and 3 form a repeated process until a satisfactory solution, one that meets all requirements and regulations, is found.

![Figure 1. Information on a subject](image)

For each subject, the dynamic syllabus system provides the number of units (the number in brackets, see Figure 1), the particular syllabus of this subject (in a separate window, accessed by clicking on the book icon), and information about the prerequisites, accessed by clicking on the star icon (see Figure 2).

The creation or modification of a class schedule is a complex process of repeated prototyping and simulation with the dynamic syllabus system. Using the system, students can interactively check the syllabus of each semester or, more correctly, just one semester, to verify if prerequisites are met and if there is a time or location conflict or an undesired time gap between classes.

However, the system does not provide an overview of the complete interdependencies nor a long-term schedule that meets the individual students’ desires and needs regarding the unit composition rules for graduation shown in Table 1. In fact, the storyboard concept is a way to add this feature to DLNRS. We propose to illustrate the interdependencies among subjects in a dynamic storyboard and thus help the students see and understand these interdependencies. This way, students will have an overview of the prerequisites for each subject. With a view to
future trends, this approach might become consistent with the representation of each student’s personal curriculum, including the didactics of its creation. Thus, a storyboard presentation here forms nested storyboards for a complete overview of the academic process toward graduation.

The effects of the DLNRS were investigated through a questionnaire at the end of the curriculum planning course, which was administered to 203 students. The rate of understanding the prerequisite conditions was about 60 percent, which means the method of displaying the prerequisite conditions needs to be improved. The rate of understanding class schedules for a current semester and until graduation was greater than 80 percent. Thus, using the DS tool to compose curricula seems to be effective and will become even more effective if the method for displaying prerequisite conditions is improved. The significantly decreased time frame in which students create their timetable supports this thesis about the DS. More than 50 percent of the students were able to create their class schedules in two hours or less. About 50 percent of the students were not able to complete their class schedule in the curriculum planning class. Those students were given a week to complete it as homework. A total of 162 out of 173 students (96.3%) were finally able to submit their class schedules.

Two classes under the same conditions, except for GPA and the credit system, were compared: an undergraduate class taught in the new system (DLNRS) and a postgraduate class taught in the old system. The class taught by the new system had no drop-outs, but in the other class, 50 percent of all students had dropped out by mid-semester. Students in both classes had much homework. In the conventional system, students did not need to pay for the class and they were not considered formally registered in the class until they had taken the class examination. Their GPA did not change even if they did not take the examination, that is, if they had dropped the subject. As a consequence, we introduced the GPA and credit mechanisms, which provide a harsh yet not unfair assessment of the students’ performance. Students should clarify their learning needs and be moderately controlled so that they can overcome difficulties such as too much homework. Thirty percent of the remaining students in the old system could not pass. Meanwhile, fewer than five percent in the new system failed. This suggests that a shorter class period (50 minutes vs. 150 minutes), a higher frequency (three times per week vs. once a week), and a significant amount of homework for review is effective for learning. However, because of the increased number of shorter classes, students’ class schedules became very complicated.

Thus, we suggest supplementing the DLNRS concept with the concept of storyboarding, which we describe in greater detail in the following section. After that, a separate section tells how to adapt this concept to supplement DLNRS (Dohi et al., 2006a; Dohi, Sakurai, Tsuruta, & Knauf, 2006b).
Storyboarding

Roots and related work

Storyboarding as a means to model information and learning processes was introduced around 1998 (Feyer, Schewe, & Thalheim, 1998; Feyer & Thalheim, 1999). One of the first storyboarding languages called SiteLang (Thalheim & Dusterhoft, 2001) was introduced in 2001. However, this language, created to model web information systems (WIS) had related but limited expressivity and suffered from non-usability by non-IT experts.

The latter, however, was not a drawback for the application described by Thalheim and Dusterhoft (2001). Chang, Lin, Shih, and Yang (2005) attempted to apply SiteLang for learning processes. However, this application was limited to e-learning only. An application to use that language for the reverse purpose, that is, not to describe WIS but to represent stories or data contents in WIS, was presented by Kaschek, Matthews, Schewe, and Wallace (2003a).

In fact, this work inspired the authors of the present paper to investigate successful didactic patterns with the more general storyboarding approach presented here. However, in the limited application field WIS and e-learning, there was a need to formally distinguish media types (Feyer, Kao, Schewe, & Thalheim, 2000) and to integrate context (Kaschek, Schewe, Thalheim & Zhang, 2003b). A first approach in distinguishing various abstraction layers in storyboards was introduced by Kaschek et al. in 2003a. In Kaschek’s publication, a limited number of different layers were modeled by different means. Again, all the mentioned approaches suffer from a lack of generality and from an intended application (e.g., WIS are different from e-learning, which is limited to electronic material or the data in internet sites of WIS).

The new storyboard concept

The storyboard concept newly developed and adopted here (Jantke & Knauf, 2005; Dohi et al., 2006a; Knauf & Jantke, 2006) is built upon standard concepts which enjoy the following:

- clarity by providing a high-level modeling approach
- simplicity, which enables everybody to become a storyboard author
- visual elements such as graphs

With respect to a better formal composition, processing, verification, validation, and refinement, the concept as introduced so far (in Dohi et al., 2006a; Jantke & Knauf, 2005; Jantke, Knauf, & Gonzalez, 2006; Knauf & Jantke, 2006; Knauf et al., 2007) has been refined by Sauerstein (2006), who introduced some modifications aimed at improving the formal verification of knowledge represented by storyboards and the machine support of its usage in practice. In particular, Sauerstein introduced a distinct start node and end node for each graph and bi-colored edges to express, in a graphic way, conditional interdependencies between incoming and outgoing edges of a node.

Here, we adopt above-mentioned modifications. This new storyboard concept is as follows:

- A storyboard is a nested hierarchy of directed graphs with annotated nodes and annotated edges.
- Nodes are scenes or episodes.
- Scenes denote no further structured elements (elements with no sub-graph behind it) of the nesting hierarchy and represent a basic learning activity, which cannot be broken down and can be implemented in any way. The storyboard can be the presentation of a media document, the opening of any other software tool that supports learning, such as a URL and/or an e-learning system or an informal description of the activity. There is no formal representation at or below the scene level.
- Episodes denote a sub-graph.
- Graphs are interpreted by the paths on which the graphs can be traversed.
- There is a start node and an end node for each graph. The start node of a graph defines the starting point of a legal graph traversing. The end node of a graph defines the final point of a legal graph traversing.
- Edges denote transitions between nodes. The outgoing edge from a node must have the same color as the incoming edge by which the node was reached; and if there is a condition specified as the edge’s key attribute, this condition has to be met for leaving the node by this edge.
- Nodes and edges can carry key attributes and/or free attributes (annotations). Key attributes of nodes specify application-driven information that is necessary for all nodes of the same type, such as actors and locations. Key
attributes of edges specify conditions that must be true for continuing traversing on this edge. Free attributes may specify whatever informal information the storyboard author wants the user to know: didactic intentions, useful methods, necessary equipment, and so on.

The interpretations of these terms are described after presenting a small example. Figure 3 shows a top-level storyboard representing the anticipation of the diverse ways to study this paper according to the reader’s individual goals. The sections of the paper that are currently under the reader’s (your) consideration appear as the storyboard’s episodes if the reader has a substructure, and as its scenes if the reader doesn’t.

Further structured sections are episodes (with subsections). Episodes need to be implemented by constructing a related sub-graph. Episodes are represented by a rectangular with double vertical lines. Each episode is followed by a pentagonal reference node, which is the episode’s re-entry point into the graph after the episode has reached the end node of the sub-graph. Sections with no further structure are scenes without subsections and are represented as rectangles. If a scene does not introduce new topical content (such as the reference list), it is represented by an ellipsis.

The representation as a graph (as opposed to a linear sequence of sections) reflects the fact that different readers trace the paper in different ways, according to their particular interests, prerequisites, a current situation (e.g., being under time pressure), and other situations.

The alternative paths may be driven by the reader’s role as follows:
- Members of the Ilmenau research group may skip the introduction, summary, and outlook sections, as well as the section on the storyboarding concept, because they are familiar with it.
- Members of the Tokyo research group may also skip the introduction, summary, and outlook sections as well as the section on the dynamic learning need reflection system, because they are familiar with it.
Referees, on the other hand, may want to read all sections. After reading the summary and outlook sections, they can read the acknowledgements and references independently from each other. Although they don’t have to read the acknowledgements, they should read the references at least.

A storyboard can be traversed in different ways, based on users’ interests, objectives, and desires; didactic preferences (such as the need for examples or illustrations for better understanding); the sequence of nodes and other storyboards visited before, the availability of resources such as time, money, equipment to present material; and other application-driven situations.

Free attributes were introduced after practical experimental usage showed that much of the information necessary by a user to execute a storyboard cannot be formalized well. However, such informal information is essential for the proper implementation of the storyboard.

The new storyboard is a semi-formal knowledge representation for the didactics of a teaching subject. Informal knowledge represented as free attributes and scenes can be integrated in a formal structure. Because of the formal graph hierarchy above scene level, the storyboard is effective as a firm base for processing, evaluating, and refining this knowledge. The vision of the further effect of these ideas is to gain didactic knowledge by analyzing storyboard paths by means of data mining methods.

**Implementation details of the new storyboard concept**

The following are supplements to the implementation of the new concept. Some of the features might not be applicable to particular implementation of the storyboards. Many of them are implicit in the general concept. Those details are described only to show implementation feasibility in order for readers to become a little more familiar with our ideas, aims, and intuition.

- The nodes called episodes may be expanded into sub-graphs on double click, since storyboards are hierarchically structured graphs by their nature. A double click on the end node of a graph will result in a return to the related reference-node in the related super-graph.
- Comments in nodes and edges are intended to carry information about didactics. Comments express goals and variants.
- Educational meta-data, such as a degree of difficulty (e.g., basic or advanced) or a style of presentation (e.g., theory-based or illustrated) may be added as key attributes.
- Edges may carry conditions as key attributes.
- Edges may carry other information about conditions or recommendations.
- Certain scenes may represent documents of different media types, such as pictures, videos, PDF files, Power Point slides, Excel tables, and so on.

Clearly, the sophistication of our storyboards can go further. The concept allows for deeply nested structures, comprising different learning knowledge (forms of learning), getting many factors involved (actors, locations, materials) and permitting a large variety of alternatives.

This is intentionally different from all the modeling approaches so far, which are driven by software technology.

In Microsoft Visio (Walker & Eaton, 2004), so-called hyperlinks can be defined on any graphic object either to open a local file of any media type with the appropriate tool or to open the standard browser with a specified URL or mail tool, if it is an e-mail address. These hyperlinks appear as a symbol when the mouse passes over a related node. With a left mouse-key click, the list of hyperlinks shows up. We make use of this opportunity for the scenes and episodes. In this way, the off-line and on-line teaching materials and tools (such as e-learning systems) are collected and structured according to individual paths of users through the storyboard.

**Nodes and edges**

The nodes types, their visual appearance, their behavior when double-clicked, and their behavior when following a hyperlink, are as specified in Tables 2 through 6.
### Table 2. Scene

<table>
<thead>
<tr>
<th>Symbol</th>
<th>OPEN BOX</th>
</tr>
</thead>
</table>
| Behavior when double-clicked | • opening a document (*.doc, *.pdf, *.wav, *.vsd, *.ppt, *.xls, etc.)  
• nothing, if just a verbally described scene |
| Behavior when following a hyperlink | • opening a document (*.doc, *.pdf, *.wav, *.vsd, *.ppt, *.xls, etc.)  
• visiting a website with the standard browser, if it is a URL  
• opening the standard mail tool, if it is an e-mail address |

### Table 3. Episode

<table>
<thead>
<tr>
<th>Symbol</th>
<th>OPEN BOX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior when double-clicked</td>
<td>OPEN BOX</td>
</tr>
</tbody>
</table>
| Behavior when following a hyperlink | • opening a document (*.doc, *.pdf, *.wav, *.vsd, *.ppt, *.xls, etc.)  
• visiting a website with the standard browser, if it is a URL  
• opening the standard mail tool, if it is an e-mail address |

### Table 4. Start node

<table>
<thead>
<tr>
<th>Symbol</th>
<th>start</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior when double-clicked</td>
<td>jumping to the start node of the related graph</td>
</tr>
<tr>
<td>Behavior when following a hyperlink</td>
<td>not applicable</td>
</tr>
</tbody>
</table>

### Table 5. End node

<table>
<thead>
<tr>
<th>Symbol</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior when double-clicked</td>
<td>jumping to the reference node that succeeds the associated episode node in the related graph</td>
</tr>
<tr>
<td>Behavior when following a hyperlink</td>
<td>not applicable</td>
</tr>
</tbody>
</table>

### Table 6. Reference node

<table>
<thead>
<tr>
<th>Symbol</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior when double-clicked</td>
<td>jumping to the end node of the sub-graph that is associated with the preceding episode node</td>
</tr>
<tr>
<td>Behavior when following a hyperlink</td>
<td>not applicable</td>
</tr>
</tbody>
</table>

For edges, it is not meaningful to define double-click actions or hyperlinks. The edges are not intended to carry subject content, but are intended to carry didactics of a mandatory, conditioned, or recommended switch between the nodes of the graph. However, the way that the sequence of nodes is defined by edges needs to be defined, particularly in the case of alternatives and forks. As introduced above, there is a rule that an outgoing edge of a node should have the same starting color as the incoming edge of the node. This serves to express conditions for leaving a node that refers to the incoming edge. The sequence of node traversing is defined as specified in Table 7 through Table 10.
Table 7. Simple edge

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Simple edge symbol" /></td>
<td>Defines a unique successor node</td>
</tr>
</tbody>
</table>

Table 8. Fork

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Fork symbol" /></td>
<td>Defines several successor nodes, which have to be traversed independently from each other, i.e., in any sequence or parallel</td>
</tr>
</tbody>
</table>

Table 9. Fork with conditions

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Fork with conditions symbol" /></td>
<td>Defines several successor nodes, which have to be traversed independently from each other, i.e., in any sequence or parallel according to the specified condition, e.g., ( \text{take n out of m specified paths} )</td>
</tr>
</tbody>
</table>

Table 10. Alternatives

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Alternatives symbol" /></td>
<td>Defines alternative successor nodes, i.e., one of them has to be traversed</td>
</tr>
</tbody>
</table>

Characteristics of the new storyboard concept

Generally, the basic but important approach behind the new storyboarding is that teaching consists of activity sequences, which can be described by a hierarchy of nested directed graphs.

On the first view, the purpose of our storyboards is similar to the purpose of traditional storyboards that are produced for plays or movies. The materials and tools of the storyboard learning activities such as textbooks, scripts, slides, hardware and software models, e-learning systems, and others are comparable to the props of a play or movie.

However, there are basic differences between the storyboards proposed here and the traditional storyboards used to depict a play or movie. The differences are as follows:
- primary purpose (learning vs. entertainment)
- degree of formalization
- the constraint of everything above the level of scenes, which is not (and should not be) applied to traditional storyboards in performing arts, since the director has some freedom of individual interpretation (which is good for arts, but bad for our purpose)
- opportunities to formally represent, process, evaluate, and refine new storyboards, which do not apply at all to storyboards in performing arts

In fact, the last two differences are due to the degree of formalization.

Also, the newly proposed storyboard concept has features in common with the following:
classic artificial intelligence (AI) knowledge representations such as semantic networks and frames
process modeling languages
state diagrams such as Petri nets (see Chang et al., 2005, for examples of their use in learning processes
workflow diagrams (see Lin, Ho, Orlowska & Sadiq, 2002, for examples of their use in learning processes
flow charts (see Sykes & Franek, 2003, for examples of their use in learning processes
the educational modeling language (EML) (Koper, Hermans, Vogten, Brouns et al., 2000)

However, items that make the storyboard concept more expressive for didactic knowledge than do the conventional representations mentioned above are as follows:
the potentially unlimited nesting of graphs
the opportunity to express “conditioned” edges by using the colors or respective key annotations to edges
the opportunity to use two kinds of fork-edges
the potential of nodes to carry many different teaching materials and tools as hyperlinks,
the fact that a scene can be implemented in any way, i.e., a scene is not restricted to something electronically available or even formally structured, such as any knowledge representations in AI, and any material that may be included in process models
free attributes are not restricted to formal information

A good example of the usefulness of hyperlinks is a storyboard by Knauf (Knauf & Jantke, 2006) for an AI course at a US university, which included hyperlinks to material from his own AI course in Germany. Now, this storyboard serves both Knauf’s university and the US university for which the storyboard was originally developed, and is also a common platform for internationally sharing teaching materials.

Furthermore, at a first glance, storyboards are quite similar to modeling approaches that are usual in software engineering, such as Petri nets and other process modeling languages. However, these languages are completely formal. There is no room to represent informal parts of didactics knowledge.

Clearly, the scenes of our storyboards define this level. A scene can be anything (media content, an informal activity description, etc.). There is no formalism at or below this level.

In the application that this paper focuses on, there is still a chance to refine the granularity through storyboarding the topics of the study. At some point, we can’t go deeper. There is no way to formalize a movie or a song, for example. However, such media content needs to be involved in the learning processes. Consequently, a hybrid knowledge/media representation through storyboarding needs to be used for (semi-)formalizing the didactics of learning processes.

**Dynamic storyboarding for modeling academic education**

**High-level dynamic usage of new storyboards for modeling academic education**

The above-mentioned new storyboards can be used to represent the didactics or tutorial knowledge of academic education, especially of DLNR. Through this, learning plans have several advantages. Examples are students’ curriculum plans for their academic career. Besides being visual in nature, these curriculum plans are formally created and can also be dynamically verified and modified each semester.

In order to optimize the new storyboard concept for an academic graduation path, we propose a new method of storyboard use: a high-level use that we call dynamic storyboarding. The results of a feasibility study are shown here.

In dynamic storyboarding, subjects to be taught are represented as scenes. In the original storyboarding, subjects are represented as top-level episodes.

In learning plans, namely curriculum plans, some high-level structure (i.e., episodes) lies above the subject level.
Figure 4. Top-level storyboard for graduation ("Sem." means "semester")

Figure 5. Alternative major subjects

Figure 6. Network computing course
As shown in Figures 4 through 8, dynamic storyboarding provides a chance to make explicit the high-level complex and dynamic education structure of university study. This makes a complex and dynamic education structure more understandable to students.

Yet, if necessary, this storyboard can also be refined later through exchanging a scene representing a subject for an episode and through storyboarding (expanding the episode hierarchically into sub-episodes, sub-sub-episodes, etc. and/or scenes representing content or syllabi). Here, each episode representing a subject is defined as sub-graphs by the professor or tutor in charge of the subject.

This paper is expected to lead us towards multilayered, high-level, and dynamic storyboarding, which is useful for university study. As shown in Figure 4 at the top-level and in Figures 5 through 7 at the lower level, it is easier to get an overview by nesting the storyboard graphs using episodes throughout all levels except for the lowest. This overview cannot be provided by any printed document commonly used.

In some dynamic storyboards, as shown in Figures 4 through 7, there is a semester scale at the left-hand side of each graph. This shows a time sequence of semesters. The group of subjects at the right side of each semester in the scale should be taken as the curriculum of a semester.

At SIE of TDU, the education based on the DLNR concept is very adaptable to students’ needs, but the cost of this feature is a complex system of conditions to meet. The regulations on composing a curriculum for studying subjects at SIE are as follows:

- The subjects to study are grouped into four subject blocks (as explained in the section describing DLNR and summarized in Table 1): major subjects (specialized subjects most related to one’s graduation path); general cultural subjects (liberal arts related subjects); optional subjects (less related but specialized subjects, cultural subjects, or free subjects); and orientation subjects (which introduce students to the new study system DLNRS and to computer operation).
- This is storyboarded quite simply, as shown in Figure 4. Of course, at this level all nodes but graduation ceremony are episodes that are further storyboarded or structurally expanded.
- As illustrated in the storyboard graph shown in Figure 5, the study of major subjects can be performed in one of three subject groups called courses, namely network computing, advanced system design, and media human environment design.
- Each course is composed of several cores. For example, the network computing course consists of the following three cores (see Figure 6), namely the network core, the computer core, and the programming core.
- With regard to nesting, each core contains about 60 units of recommended major or specialized subjects. Some of the subjects belong to more than one core.
- For the other two general cores, (a) at least two units of orientation subjects and (b) 40 units of general cultural subjects (liberal arts and related subjects) are indispensable for graduation.
- Twenty-two other units of any subjects less related to the selected course or career are recommended for graduation. Optional subjects can be taken but not counted as any unit or credit.

This situation calls for a representation in nested structures. To show the complexity for requesting the nested structure, Figure 7 illustrates the subjects and their relationship within the network core. It shows only the precondition prerequisite edges that are relevant for a particular student. The introduction of the dummy node “summing-up” is an idea for a mechanism to process the information of all nodes in this sub-graph and to avoid too many edges from every node in the figure. The processing is, for example, to add units of all succeeded subjects in the sub-graph and to check if the graduation conditions for these subjects are met. As described later, such results are passed on to the super-graph (i.e., network core, network computing course) of the sub-graph. This introduction also helps students see the big picture.

Meanwhile, Figure 8 shows the general dynamic storyboard on learning English with fewer semester restrictions or with a loosely fixed sequence of subjects before the storyboard is simplified according to the particular student’s situations. Therefore, unlike in the master dynamic storyboards in Figure 4 through Figure 7, there is no semester scale at the left-hand side of the Figure 8 graph, even though this is a sub-sub-graph expanded from the episode of general cultural subjects (see also Figure 4). Figure 8 illustrates that students do not strictly depend on a fixed sequence of subjects. In accordance with their individual needs (for example, an upcoming internship abroad) and their situations (for example, a time gap in the current curriculum that should be filled for effective study), students
In addition, dynamic storyboards need to be more individual and dynamic than general or master-level storyboards because the composition of a learning plan for academic education depends not only on general regulations (such as in master dynamic storyboards) but also on individual dynamic facts, as follows:

- Individual goals, such as a company to serve, a position to reach, an individual talent to support, an amount of money to make
- Pre-conditions, such as prerequisites, a required level of success in the preceding semester or in the school-leaving examinations before entering university, equipment such as a notebook or car, proficiency in a particular language, or resources such as money or time to attend the class
- Talent, such as creativity, analytical skills, leadership skills, or athletic ability

For example at SIE of TDU, there is a system of grade-point average that limits the number of units an individual student can register in an upcoming semester (see section 2) and, as a result, narrows the individual student’s storyboard for an upcoming semester.

Thus, in order to take individuality and dynamics into account when composing a high-level dynamic storyboard for academic education, factors such as career goals, pre-conditions, current or previous semester’s GPA, etc. are necessary. These factors, which are key attributes or annotations in dynamic storyboards, should be formalized and associated with both the related episodes and the students.
The dynamic storyboarding system can analyze these annotations or attributes such as the unit number or grade points attached to subjects as well as the annotations or attributes attached to students, such as the achieved GPA. After each semester, high-level dynamic storyboards can be individualized and dynamically updated according to their new status.

Through temporarily omitting or hiding those items already traversed or not focused on by the student (e.g., Figure 7 shows only prerequisite edges along with their subject nodes and hides others) and items unnecessary or impossible to traverse (because they are outside of the student’s area of interest or not useful for his/her career, or because of the student’s low GPA in the previous semester), the storyboard is dynamically updated each semester and becomes easier for students to keep track of their program.

Thus, dynamic storyboarding is helpful for supporting the individual planning of the study. As well, it ensures that the plan satisfies students’ goals and all related regulations.

**Benefits through multilayered nature of dynamic storyboarding for universities**

Based on the multilayered nature of dynamic storyboarding, namely, through nesting the dynamic storyboard graph, the students obtain an overview of the structure of their study and its variety of options and enjoy the graphical representation of the interdependencies of the various subjects. One graph contains a number of nodes that are easy to track. In case the graph becomes too complicated, a new nesting level is introduced. Thus, this representation is much better than any textual or tabular representations such as DS (Dohi & Nakamura, 2003) of subject interdependencies.

Resulting from these general effects of dynamic storyboarding, the “space of opportunities or selections” becomes a “space” and is no longer represented as merely a flat list or a table. Since such a structural and graphical representation is much easier to interpret by humans, students get to know this space much more thoroughly. Thus, they can come up with new ideas to compose their subjects in a way that meets both the university’s regulations and the student’s individual settings, such as their learning needs, learning preferences, career goals, and so on.

The fundamental benefit of dynamic storyboarding for students is that they can enjoy a much better overview of options (or alternatives) and possible ways to obtain individual dynamic plans of more flexible but more mazy university study than before. This is because the storyboard graphs can hide sub-graph structures of episodes, revealing them only on request (double-click), and the student is not bombarded with undesired information. Before, students used textual descriptions about the content of the subjects, pre-conditions for each subject, constraints regarding individual GPAs, and so on. The graphical representation and automatic reduction by dynamically masking the individually impossible options according to each student’s individual situation is a remarkable advantage for maintaining an overview of the maze of dynamically changing opportunities and constraints mentioned previously.

Due to the remarkable advantage of the above-stated dynamic overview capability, the dynamic storyboard is useful for long-term planning (over several semesters). Meanwhile, the dynamic syllabus (DS) tool introduced earlier as a part of the DLNR system is a planning tool for a particular semester. These benefits of dynamic storyboarding are concretely explained below.

Let’s consider a student, who
- developed a storyboard such as Figure 7 (but a slightly different sub-storyboard of it that doesn’t contain all its scenes) during his/her curriculum planning class,
- finished the third semester with a GPA that allows him/her to follow the long-term plan without restrictions as to the number of units achieved, and
- did not change this plan by learning need reflection (if his/her GPA is low, this might happen by dynamically changing his/her career goals).

For the student’s fourth semester, he/she is strongly advised to take the recommended subjects (denoted as scenes with a white background in Figure 7), as follows:
- system programming /OS (operating systems)
As well, the student is advised to take the less recommended subjects (denoted by scenes with a grey background), as follows:

- Unix and C programming
- digital measurement
- database systems

Finally, if these recommendations are accepted by the student, the aforementioned six subjects are sent to the DS to register a semester schedule for the fourth semester. However, each of subjects, which is a starting point of a prerequisite edge for the fifth or a later semester, remains in the dynamic storyboards as shown in Figure 7.

Further, the dynamic storyboard system is able to check or suggest an individually optimal curriculum by manually providing individual interests such as desired courses (e.g., network computing and advanced system design), interesting cores e.g., network, computer programming, or a certain preferred application domain), and individual career goals such as vocation names (e.g., a network engineer, interior designer, a preferred position level, a preferred enterprise, or a preferred location of a future entrepreneur) as well as by automatically including individual educational data from the electronic data processing (EDP) system such as the GPA history of previous semesters and the results of entrance examinations and/or preparatory education examinations.

In fact, the above-mentioned individual learning needs may dynamically change during the course of study. Thus, the curriculum will also change over the upcoming semesters. The change from one semester to the next is usually not serious. However, over many semesters, the course of study may come to look quite different from the original plan submitted in the first semester.

Difficult subjects such as mathematics and English often tend to decrease a student’s GPA. These subjects are included automatically (but minimally) by the EDP system, based on students’ previously achieved GPA and level.

Suggestions for successful learning patterns for students who want to become network engineers are illustrated in Figure 4 through Figure 7. Figure 7 is a particularly good example.

Manually collected information comes into the system through a related interactive webpage and is incorporated into the curriculum creation support system using a Prolog-based (inference) program. Automatically collected data is provided by the university administration’s own EDP system and is also incorporated into the curriculum creation support system using a Prolog-based (inference) program.

The processing or inference mechanism in Prolog is based on the storyboard represented by facts for the following predicates:

- includes(graph , [<elements>] )
- edge( [<begin>] , [<end>] , [<color> , <color>] )

Here, the predicate includes the elements (scenes, episodes, start nodes, end nodes, and reference nodes) of each particular graph of the nested graph hierarchy. There is one fact for each graph in the knowledge base along with a list of its elements.

Also, each edge is specified by a related fact. Since edges may have several start nodes and/or end nodes (such as the forked edge in Table 8, and its counterpart or a joining edge that re-unite the forked paths), the starting node and ending node of each edge are represented as a list of nodes. In case of simple edges, there may be just one element. Since edges may be bi-colored (see Table 7), the edge’s colors are specified by a list of two elements that are identical in case of single-colored edges.

Based on such a storyboard representation, an individual curriculum can be automatically checked, edited, and displayed by a kind of artificial intelligence program or inference program. The individual curriculum suggestions are sub-storyboards, that is, dynamic storyboards that are derived from the more complete storyboard by omitting
nodes and edges that don’t match the student’s individual situation or request (such as learning needs, learning preferences, career goals, the previous semester’s GPA, and so on).

In the case of automatic curriculum suggestions based on this storyboard representation, the top predicate of the inference looks as follows:

```prolog
compose curriculum (Student)        :-                       /* suggests an individual sub-storyboard for a student */
    collect information (Student, database, StudentsInfoList),  /* collects information via a web interface */
    collect information (Student, EDP, EDPInfoList ),              /* collects information from an EDP system */
    process (StudentsInfoList, EDPInfoList ),                          /* derives an individual sub-storyboard */
    present result (Student ).                                     /* presents the result of the composition */
```

“Student” refers to a student’s unique identification, “database” refers to information that a student has manually input via a related web interface.

The predicate “collect_information” gathers the information needed to compose the sub-storyboard. Depending on the type of information this predicate is called twice in the above Prolog-rule. Once to receive the manually input data from a database and once to receive the data automatically input by the EDP system.

The predicate “process” produces facts similar to those that represent the complete storyboard, but contain the student’s identification:

- includes (Student , graph , [<elements>] )
- edge (Student , [<begin>], [<end>] , [<color> , <color>])

These facts represent the student’s individual sub-storyboard.

The predicate “present result” calls a program, which graphically presents the derived sub-storyboard.

For example, in the case of the dynamic storyboard of the studies offered by SIE of TDU, our experiments indicate that these sub-storyboards are basically unique sequences with only a few remaining options. Namely, there are only a few different options left after all nodes and edges that were inappropriate, not possible, or not useful for an actual student were automatically omitted. Having few options left is key to maintaining an overview of the maze of opportunities. By narrowing down the options, students can manually make their individual choices in the maze of dynamically changing opportunities and constraints. Dynamic storyboarding has such benefits.

The results so far are quite promising. Several (at least five) students of SIE in TDU joined the experiment to prove the benefits mentioned above. The students confirmed that this approach is helpful in creating individually tailored and optimized curricula. Indeed, due to the multilayered overview nature of dynamic storyboarding, students could easily create their individual storyboard or curriculum. Meanwhile, teachers created, displayed, checked, and modified general or master dynamic storyboards using Microsoft Visio (Walker & Eaton, 2004) for the storyboard (usually for representing the syllabus of subjects) mentioned previously in the subsection “The new storyboard concept” of the “Storyboarding” section.

Using Microsoft Visio, teachers created, checked, and modified master storyboards. Such master storyboards are initially prepared so that students can use them for easy selection and combination (called individualization) of their needed, permitted, and preferable subjects in their specified program of study. Here, the model is represented at the master level of dynamic storyboards. The master dynamic storyboards are pre-generated by teachers using Microsoft Visio as the master knowledge base, displayed by the system, and checked/modified by students. The modification input by students make the master dynamic storyboards into individual dynamic storyboards with regulations such as prerequisite conditions and recommendations satisfied by the support of the system’s checks, warnings, and suggestions.

The system checks the regulations using the data, including regulations and students’ achievements such as GPA and subjects passed. The data are sent from the EDP system using the Prolog-based program as mentioned above.
The concrete procedure and interaction among the system, students, and teachers derive useful information for high-level, complex education support to create individual curriculum. In this experiment, the system is called the individual curriculum creation support system (ICCSS), which uses the Prolog-based program.

Using ICCSS, teachers generate master dynamic storyboards that represent curriculum creating didactics. The master dynamic storyboards are utilized to create students’ individual curriculum easily. Teachers draw their subjects’ master storyboards by using Microsoft Visio. More concretely, such master dynamic storyboards are derived from general dynamic storyboards, such as those shown in Figure 8, which were initially made by teachers using ICCSS and Visio. ICCSS verifies and displays the master dynamic storyboards as nested hierarchical diagrams such as those in Figure 4 through Figure 7. This allows teachers and students to easily view and check the entire storyboard. Once verified by teachers, the master dynamic storyboards are stored in the ICCSS and used by the students for curriculum planning as follows:

1. **ICCSS loads students’ information.** When students enter or log in to ICCSS, it loads and displays the master dynamic storyboards or the previously created and/or partly executed individual dynamic storyboards according to the students’ login information. ICCSS also loads students’ personal data, such as their affiliations or achievements from the administration’s own EDP system.

2. **Students select their related course/path from the master storyboards using the multilayered overview.** Students and/or ICCSS select the related path or thread from the master or previously created dynamic storyboards (Figures 4–7), which have been loaded and displayed hierarchically to allow students to obtain an overview of complicated education didactics, which includes subject groups (from Figure 4), courses (from Figure 5), core courses (from Figure 6), and subjects (from Figure 7) that they prefer or need to fulfill their academic goals or vocations after graduation. This selection is done semi-automatically. Using the information loaded in step 1, ICCSS automatically suggests the subject group courses, core courses, and subjects for students to confirm or change. Later, in step 4 and step 5, the recommendations by ICCSS and the selections by students are described in further detail.

3. **ICCSS indicates students’ dynamic status diagrammatically.** ICCSS indicates the current semester, subjects already passed, credits obtained, credits necessary to graduate, etc. on the dynamic storyboard.

4. **ICCSS highlights recommendations on the dynamic storyboards.** ICCSS further contributes to students’ curriculum creation by automatically suggesting the recommended combination of subjects and highlighting them on the dynamic storyboards, as shown in Figure 7. The system does this dynamic recommendation by automatically checking didactics such as prerequisite/recommended subjects in the storyboard loaded in step 1, and dynamic factors such as students’ GPA from the previous semester, the number of credits required for graduation, etc. that have been loaded from the administration’s EDP system.

5. **Students select unnecessary subjects, referred to ICCSS’s dynamic indication.** Students manually select unnecessary or un-preferable subjects to make their own curriculum. In such instances, ICCSS stops highlighting selected subjects as well as those whose prerequisite subjects are deemed unnecessary, and makes them light-colored and changes the regulation/recommendation information as well. Of course, students can override the results or can select mandatory or preferable subjects from the unnecessary subjects, which are marked by a light color. Then, ICCSS highlights the selected subjects and light-colored subjects whose prerequisite subjects were deemed necessary.

6. **ICCSS verifies and displays the resulting individual storyboard.** When students finish their selections (individualization) and store the results for the semester or all semesters until graduation, ICCSS verifies the results. ICCSS also displays the resulting dynamic storyboards as nested hierarchical diagrams so that teachers and students can view the results and check them for accuracy. The verified storyboards are stored by ICCSS as knowledge or didactics for subsequent curriculum recommendations. Thus, after totally checking, displaying, and saving the dynamic storyboard representing student’s own individual curriculum, ICCSS sends the information necessary for administration, such as students’ ID numbers registered subjects, to the EDP system.

Before using the ICCSS system, students could never adequately check if these combinations (the students’ individually created curricula) followed the recommendations based upon their achievements in previous semesters. However, due to the automatic suggestion and verification described in the processes step 2 through step 6, especially step 4 through step 6, these student-created curricula match their individual situations, including abilities, desires, and restrictions much better than traditional curricula do.

Meanwhile, teachers also benefit from this approach when they create general or master dynamic storyboards as well as receive the curricula that have been composed by the students based on their individual settings. Teachers can
much more easily estimate the usefulness of the submitted plans because it is easy to obtain an overview of the plans. Moreover, the overview makes it easier for the teacher and school administration to logically decide whether the specified subject compositions are really what the students intended to specify. In other words, storyboards are used to logically evaluate specifications involved with various restrictions, especially for long-term plans. In particular, through the dynamic storyboarding system, the curricula can be checked unconsciously or even against the intentions of the system users.

While these benefits are based on better knowledge processing by students, teachers, and school administration staff, the knowledge formally represented in dynamic storyboards can be used for more advanced knowledge processing by machines. This is outlined in the next section as benefits of the formal nature of dynamic storyboarding.

As described above, the dynamic storyboarding was found to be successful in supplementing DLNRS. That is, didactic knowledge in the flexible but complicated DLNRS proved to be easily and clearly represented by high-level dynamic storyboarding. Storyboards are represented as hierarchical graphs without a particular implementation tool to support their use by machine. Even at this early stage of introducing the storyboard or dynamic storyboarding technology, students already enjoy its benefits. Students can study more effectively by individually determining the subjects and their sequence based on their dynamically changing needs and circumstances. Namely, the dynamic storyboarding adaptively assists or guides students through the multilayered overview of an individual curriculum in the maze of opportunities and limitations, and beyond by improving quality management of complicated academic programs. The latter is described more in detail as follows.

Benefits of the formal nature of dynamic storyboarding for universities

The proposed dynamic storyboarding also has some promising benefits because of increased use of AI methods that are implemented as Prolog programs based on the abovementioned formal representation of dynamic storyboards.

The first benefit is that dynamic storyboarding increases logical reliability, such as consistency, completeness, and so on. To ensure consistency and completeness of our storyboards, several verification procedures were developed and implemented (Sauerstein, 2006):

1. A test for the correctness of the graph hierarchy focuses on questions as follows:
   - Does every episode have exactly one related graph?
   - Does every sub-graph have exactly one related episode node in exactly one related super-graph?

2. Reachability issues for each node are formally checked:
   - Does every traversing path terminate? (In other words: Is the end node reachable on every possible path in each sub-graph?)
   - Is every node reachable from the start node in every sub-graph?

3. Completeness and non-contradictoriness of alternative outgoing edges with the same beginning color are checked by logically analyzing conditions expressed as annotations of each node’s outgoing edges with the same start color.

4. Edge colors, which express incoming/outgoing conditions, are verifying through checking the following issues:
   - Is there a unique start color? The start node’s outgoing edges should have a unique color. A start node has no incoming edge, so if the outgoing edges have different colors, it is difficult to know which edge is to go out from the start node.
   - Is there at least one outgoing edge with the same beginning color for each incoming edge’s finishing colors?

These verification procedures are useful for composing the general dynamic storyboard. This one represents all possible combinations of recommended subjects and their sequences, which makes it somewhat complicated. Since the individual plans are derived from the general or complete dynamic storyboard, it is important that the general dynamic storyboard is free from errors such as logical anomalies (non-reachable nodes, for example).

A second benefit is the inheritance concept that was implemented within the graph hierarchy, which distinguishes several inheritance types such as sum, maximum, or set union (Xu, 2006). This is a deductive inference over the knowledge represented as dynamic storyboards. Because of this, some features of an episode are inherited from the nodes in the related sub-graph. For example, already achieved units (e.g., in major subjects) are totalled and checked.
(e.g., for graduation) using such inheritance features. Thus, students can know which of the three cores (network, computer, or programming) is most difficult, since the feature “difficulty” can be a key attribute of each node, as shown in Figure 6 by the nodes in the sub-graph. There are different ways through each core, with different levels of difficulty. The maximum value indicating the worst or most difficult case can also be found at an even higher level. Further, the following information can also be passed from each graph to its super-graph by selecting the maximum value of the graph’s nodes the: Which is the less expensive one? the less time-consuming one (in terms of how many units are needed to finish a certain core, for example)? and so on.

The third benefit is a set of reliable operations, defined below, that incorporate features such as checking for logic. Their exhaustive use automatically leads to a legal dynamic storyboard. This ensures that the students’ own curricula will be more reliable and free from logical anomalies so that students can enjoy a flexible but mazy university education. These operations are:

- adding paths
- adding nodes
- changing scenes to episodes by introducing a related sub-graph
- adding a concurrent path
- merging equivalent nodes by introducing related bi-colored edges to make sure that the linkage with the remaining graph isn’t changed.

The fourth benefit of the dynamic storyboarding is the fact that it is easy to learn because it formally represents of knowledge for university study. Since students at SIE of TDU create or modify their own timetables as mentioned before, dynamic storyboarding is expected to be effective.

A basic approach is to use AI technologies such as data mining and case-based reasoning on these formal process models for learning successful storyboard patterns and recommendable paths. This learning is based on an analysis of the paths on which former students went through the storyboard and is also based on their success that associated with these particular paths (Boeck, 2007).

The introduced storyboarding approach is very useful in assisting students suffering from the maze of opportunities and constraints in university education. More concretely, a simple prototype was recently developed to evaluate curricula created or modified by the students in advance of their study (Boeck, 2007). The basic idea is as follows:

- The construction and successive refinement of a decision tree based on paths that have been taken by former students, that is, paths with the students’ known level of success.
- The application of the decision tree to estimate the possibility of success of a planned path, where current and future students want to go. Moreover, Boeck (2007) introduces a technology to supplement a given curriculum that leads to an optimum chance of success. By simulation, we found that at least half of the planned schedules have the potential to be refined towards a plan with a higher chance of success.

Since the last of these operations might not be easy to understand and imagine, it is illustrated in Figure 9. Here, V1 and V4 are equivalent, and V2 and V3 are equivalent. Since different users visit V1, V2, V3 and V4 in different sequences, they are represented as different nodes on the left-hand side. By introducing a new color to express these different sequences, the equivalent nodes can be merged together.

The fourth benefit of the dynamic storyboarding is the fact that it is easy to learn because it formally represents of knowledge for university study. Since students at SIE of TDU create or modify their own timetables as mentioned before, dynamic storyboarding is expected to be effective.

A basic approach is to use AI technologies such as data mining and case-based reasoning on these formal process models for learning successful storyboard patterns and recommendable paths. This learning is based on an analysis of the paths on which former students went through the storyboard and is also based on their success that associated with these particular paths (Boeck, 2007).

The introduced storyboarding approach is very useful in assisting students suffering from the maze of opportunities and constraints in university education. More concretely, a simple prototype was recently developed to evaluate curricula created or modified by the students in advance of their study (Boeck, 2007). The basic idea is as follows:

- The construction and successive refinement of a decision tree based on paths that have been taken by former students, that is, paths with the students’ known level of success.
- The application of the decision tree to estimate the possibility of success of a planned path, where current and future students want to go. Moreover, Boeck (2007) introduces a technology to supplement a given curriculum that leads to an optimum chance of success. By simulation, we found that at least half of the planned schedules have the potential to be refined towards a plan with a higher chance of success.
In greater detail, the construction of the decision tree is based on former students’ paths represented as dynamic storyboards that model the opportunities or selections. Namely, from the space, the students take a particular path represented as a dynamic storyboard. Each of these paths can be associated with the degree of success that has been achieved by these students. In case a set of students choose the same path, the degree of success can be estimated by the average degree of success of all students that traverse this path.

In particular, this path begins at the start node of the top-level storyboard and terminates at its end node. Each episode on this path is replaced by its sub-graph. This replacement continues throughout the entire hierarchy of nested graphs. As a result, such a path contains atomic scenes only (Boeck, 2007).

Each scene of this dynamic storyboard for the education of SIE of DLNR represents subjects for which students can register. Figuratively speaking, the decision tree is constructed on the basis of a flattened storyboard. “Flattened,” in this context, means the graph is hierarchically collapsed to just one level with no sub-graph.

The decision tree is based on the concept of bundling common starting sequences (Boeck, 2007) of the various paths to a node of the tree. In Boeck, these starting sequences are called the least common denominator. Of course, all paths traversed by the students begin with the start node that forms the root of the decision tree. Different subsequent nodes of the paths will result in different sub-trees below the actual root on the last node of the common starting sequence. This continues for each lower level sub-tree accordingly. If there are different paths with a common starting sequence from the root to the actual node different in the subsequent nodes, related sub-trees will be established.

Each node in the tree that represents a final node of a path is followed by a label node. Label nodes contain a list of marks that students receive after going through this path. Each mark appears with the number of occurrences (the number of students getting the mark). Additionally, the weighted arithmetic average value of these marks (GAM—“Gewichtetes Arithmetisches Mittel” in German) (Boeck, 2007) is also attached to this label. The value of GAM serves as an estimate of the chance of success for future students who plan to follow the same path.

The application of the decision tree, which contains knowledge gained by data mining of storyboard paths, is as follows: If a student submits a curriculum plan that is already represented in the decision tree (as a path from its root to a node that is succeeded by a label node), the prediction or estimation is very easily done by presenting the content of this label.

On the other hand, if a student submits a curriculum plan that is not represented in the decision tree, the most similar sub-path in the decision tree will be identified. In this context, similarity refers to the number of same subjects in the common part of paths that are represented in the tree. In other words, those paths in the decision tree that have the longest leading (starting and subsequent) parts in common with the path representing the submitted curriculum plan will be identified. The last node of this common starting sequence of the common leading part forms the root of several sub-trees that represent remaining paths. They are different from the submitted remaining path. The GAM value is used to estimate the chances of success.

Of course, in such a case, the student may be interested in suggestions to modify the submitted path in a way that the chances for success reach an optimum or become highest in value. For example, it is suggested that students substitute the submitted remaining path by the most successful alternative remaining path with the best GAM value among the remaining paths in the decision tree whose common leading part is the longest.

Based on this modification suggestion for the submitted path, the student can decide whether to keep the submitted curriculum or modify it in accordance with the optimization result by considering the chance of success.

Due to the benefits of the formal nature of dynamic storyboarding, which improves quality management of complicated academic programs, students are assisted and guided in the maze of opportunities. Namely, the quality of complicated academic programs or students’ curriculum plans can be improved incrementally through validation (looking at the students’ success chance) and refinement (restructuring the storyboards in a way that paths with bad results are removed).
Thus, academic education modeling by dynamic storyboarding opens the door for applying AI technologies such as data mining to estimate chances of success and classic inductive reasoning to derive successful didactic patterns. Didactics of teaching university subjects has not been modeled so far. As a result, dynamic storyboarding is considered effective as a modeling approach for academic education, because of the following:

- **formality** (causing the applicability of reasoning methods whose effects include abovementioned estimation of success chances and derivation of successful didactic patterns as well as logical reliability of complicated academic programs or students’ curriculum plans)
- **clarity** (being simple enough to be applied without particular software tools by every university teacher and student—not only by computer scientists)
- **visibility** (hierarchical graph representations [instead of tables, for example] that make it easier for teachers and students to obtain an overview of the flexible but complicated university education or programs).

**Summary and outlook**

In contrast to basic education, such as that of primary and secondary schools, academic education at universities is characterized by a large variety of dynamically changing opportunities to compose academic timelines or class schedules and teachers (professors and tutors) who are usually experts in their subject but do not necessarily have the didactic skills to teach. In particular, at the School of Information Environment (SIE) of Tokyo Denki University (TDU), students are required to be more flexible in designing their study according to their dynamically changing needs, wishes, interests, and talents.

However, in university education, there are usually requirements and rules to guarantee a certain level of academic quality. These rules are often complex, possibly dynamic in some aspects, and difficult to keep track of. A remarkable number of students, at least those of some universities in Germany and Japan, suffer from violating such regulations. They often make mistakes because they do not know or they forget these regulations. Students need qualified assistance in this maze of dynamically changing opportunities and limitations.

A basic property of a qualified guidance is adaptability, that is, a certain dynamic with respect to varying learning needs, context conditions, and the students’ educational history. At first blush, the basic benefit of the proposed storyboards compared to that of any formal knowledge representation is the fact that the storyboards provide a semi-formal but easy-to-use overview of the relevant class schedules by nesting the graphs and reducing them to individual and dynamically possible choices.

A more important benefit for our future work is that dynamic storyboarding is a step towards making academic education processes a subject of reasoning with AI technologies such as data mining and identifying successful didactic patterns for individual students. This is possible due to the fact that the proposed storyboards have a semi-formal but sufficient degree of formal knowledge representation that is controlled by a set of construction operations that ensure formal correctness when designing a storyboard (Sauerstein, 2006), and is verified for further formal features by automatic structure tests after the storyboard has been designed (Duesel, 2007).

This opens the door to designing learning plans that ensure a certain degree of learning quality, that is, a strong indication that learning ends up with a high level of success in academic education as a result of the storyboard being incrementally refined based on an automatic analysis such as data mining of its usage. In other words, dynamic storyboarding supports quality management in academic education.

Currently, this system for SIE of TDU and a similar system for supporting the computer science study at the University of Ilmenau are under construction. In the experimental use, the suggested approach was found to be very general and can easily be adopted for any other university education.

Our upcoming work focuses on the following issues:

- A definition and representation of criteria that allows the specification of individual goal-driven storyboards. In fact, this is very different depending on cultures, countries, and universities. Therefore, we plan to do that prototypically for SIE of TDU.
- Storyboards have a high performance with respect to didactical issues of planning education processes. However, there is still no way to manage these processes according to their resources (e.g., to concretely plan...
weekly timetables based on requests and available capacities such as rooms, teachers, equipment, and so on). Therefore, a desirable synergy effect is expected when incorporating the planning capabilities of the dynamic syllabus tool of the DLNRS into the storyboards.

- Including meta-knowledge is another focus to infer learning needs, learning desires, preferences, and talents. This meta-knowledge is useful for maintaining the university’s educational resources by predicting, to a certain extent, upcoming students’ learning needs. This is desirable not only for need-oriented and effective planning at universities, but also for suggesting content, such as class schedules represented by storyboards according to the students’ desires.

- As well, individual learning plans should not be based solely upon individual quantitative capability (such as GPA) or upon the success of former students who went similar ways. But also individual properties, talents, and preferences should be considered. For example, some students demonstrate talent for analytical challenges, some are more successful in creative and compositional tasks, and others may have an extraordinary talent for memorizing a lot of factual knowledge. Consequently, at some point we need to include some sort of user profile to avoid overwhelming the students with suggestions that don’t match their individual preferences and talents.

Acknowledgements

The authors gratefully acknowledge the contributions of Professor Klaus P. Jantke at the University of Ilmenau in Germany for his storyboarding approach.

References


Meissner, H. (2003). Ein multimediales Lehr- und Lernsystem über die Erstellung von Drehbüchern für multimediale Lehr- und Lernsysteme (German) [English: A multimedia teaching and learning over the building by scenarios per multimedia teaching and learning], In Hambach, S., Urban, B. (Eds.), Proceedings of 4. IuK-Tage Mecklenburg-Vorpommern (pp.18–20), Stuttgart: IRB.


Xu, G. (2006). Ein Vererbungskonzept für Storyboards (German) [English: An inheritance concept for storyboards], Project Work, Germany: Ilmenau University of Technology, Faculty of Computer Science and Automation.
Technology-mediated Learning Environments for Young English Learners
(Book Review)

Reviewer:
Griff Richards
Associate Professor and Director
Technology Enhanced Learning Research Institute (TELRI)
Athabasca University
griff@athabascau.ca

Textbook Details:
Technology-mediated Learning Environments for Young English Learners
Connections in and Out of School
L. Leanne Parker (Ed.)
ISBN: 978-0-8058-6233-1
Publisher: Routledge
Pages: 336

In 1980 I published a paper in the NALLD Journal on how computer-based learning activities could be integrated into the modern language lab (Kenner and Richards, 1980). My co-author and I thought we were pretty avant-garde having collected several computer programs and installed a couple of terminals in the university’s language lab. One of the terminals was actually an Apple II microcomputer and we had figured out a way to store its programs on the mainframe and download them from a menu as local hard drives were not yet available. The computer corner quickly became far more popular than the rows of listen and repeat cassette stations and we congratulated ourselves in the constant chatter of foreign students who clustered there. One day we discovered most were Computer Science students who had figured out a way to bypass our English language tutorials and access the mainframe to do their homework instead of queuing up in the science lab. Parker’s book of collected chapters helps put that early experience into perspective. Technology is important not just because of intended teachings, but because it can become a focal point for student interaction, a gateway to new worlds of information and a fulcrum for socially constructed learning.

Over the past 30 years I never ceased to be amazed by the clever offerings of technology, and how they have so little impact on intentional school-based learning yet so much impact on everything else we do. Years of research into teaching programs have generated good teaching algorithms, particularly for the acquisition of vocabulary and practicing grammar, but seldom has this penetrated the primary school classroom. At first we thought, “if only we had graphics”, and then “we need an installed base” of machines to create a market. Well, now schools have the computers (with graphics) and they are connected to a vast internet of resources, but other than typing assignments, doing web searches and making audio-visual presentations we don’t often reap rewards from the technology seeds sown. And for learners of second languages (or even for learners of first languages) we seem to be light years away from harnessing the potential benefits. Parker and her co-authors seem to share this perspective. Language learning has not enjoyed a return on technology investment. Their book examines the many social, and technical issues that need be addressed, and supplies a refreshing number of examples of projects where technology appears to have been used with positive results for young second language learners. A feature of the book that I enjoyed was that each main perspective was followed by a critical reflection by another contributor.

The context for this book is an emerging and pressing social need in the United States. The melting pot where minorities quickly shed their cultural identities and blended into the social fabric within one or two generations is being rapidly transformed into a new multicultural society. By 2030 an estimated 40% of the K-12 school population will be second language learners. The largest and most rapidly expanding language group is Hispanic, but other immigrant populations are also establishing linguistic beachheads in the USA. On the one hand, multiculturalism in an era of international trade should enrich the society, but on the other, a public school system dominated by a unilingual white Anglo-Saxon culture is ill equipped to cope with the demands of a swelling population of second-language learners. As second-language children struggle to simultaneously learn academic English and course
content, many fall farther and farther behind the achievement norms established for the dominant linguistic group. Allington and Cunningham (2002) noted the same pattern for poor readers, those who are behind at the start of school fall further and further behind each year. The growing gap spirals into social and economic disparity. The same fear may also apply to the new technology literacy – the ability to seek, sort and re-use information from the internet. The digital linguistic divide is exacerbated by the current predominance of resources in English.

Just as technology is evolving, so is our understanding of how children learn language. Basically, there is only so much one can do with low-level vocabulary and grammar drills. Indeed, in one chapter Carla Meskill points out that young children do not benefit from meta-linguistic approaches that might appeal to adults with well-developed language schemas. Language is socially constructed and is learned best from those with whom the child interacts in what Vygotsky would have called their “zone of proximal development”. Children have a natural ability to learn languages, and language has meaning when it is used to further the child’s own purposes. School English is different from the English of causal daily intercourse – it is a rich vocabulary and a set of cultural formalities of composition, methods for critical thinking and algorithms for test taking. For child whose “proximal zone” at home and on the street contains no English, there may be few opportunities for preparation for the language used in school.

It is impractical to summarize all the perspectives brought together in this volume, so let me highlight a few key chapters:

Jim Cummins provides a stimulating chapter on technology, literacy and young learners. He describes three nested pedagogical orientations: transmission, social constructivist and transformative. He points out that one cannot expect technology to stimulate constructive dialogue in schools if such practices are not already present in the classroom. He goes on to provide a set of design criteria extracted from analysis of case studies where technology seems to make a difference. Olga Vsquez follows with a wonderful rebuttal, closely examining Cummins criteria (she calls them his “Rules of Engagement”) and notes they too lack empirical validation. For her, the real problem is that schools are “intractably resistant to change” and thus a solution to the lack of progress in the schools is to foster change in the community. She goes on to describe La Classe Magica - a community based technology program that has 17 years of effective results.

Jill Castek leads a strong contribution “Developing new literacies among multilingual learners” and cites Coiro et al that “The Internet is today’s defining technology for literacy and learning.” No matter how you define “new literacies”:
1. They are here and are central to participation in today’s globalized community,
2. They are “deictic” and the rapid change in technology brings rapidly changing perspectives, and
3. They are multifaceted and require analysis from several points of view.

She notes that new literacies have also inverted the power relationship in education – when it comes to technology they young people are often “more literate than their teachers”. In short, if teachers become guides on the side, if schools would see multilingualism as a strength rather than a problem and if technology was welcomed as an opportunity to create change, then there is hope at the bottom of the digital box.

Several of the chapters provide overviews and examples of the ways in which Information and Communication Technologies (ICT) are used with young English Language Learners. These are interesting, not just in terms of providing a catalogue of useful ideas, but also in reminding us of the lack of coordination of the effort in this area.

The key impressions I brought away from my reading are:
1. ICT can help young language learners accelerate their learning of a second language, but to be really effective it has to go beyond drills for phonetic awareness, vocabulary acquisition and grammar to become a fulcrum for social constructive learning. It is not learning from the computer, but learning through the computer that is likely to improve English comprehension.
2. ICT approaches can and should reinforce and build on first language skills. By helping to reinforce each child’s cultural identity and the value of multiculturalism, ICT can be transformative not just for young children but for society as a whole.
3. One of the strongest methods for technology to assist children in learning their second language is to create collaborative activities in which they interact with native English speakers in a meaningful way. In essence, this
expands the child’s proximal environment to include language models and experiences that may be missing in their home or community.

4. ICT fails to have widespread language learning impact because the many interventions are scattered geographically, isolated technologically and discontinuous in scope and range of curriculum. The time has come for large-scale longitudinal approaches involving TV, internet, games, print and social networking.

5. ICTs cannot be ignored because they are already highly integrated into our society and constitute the \textit{de facto} new literacy. Schools that will not adapt deserve to be by-passed. Adaptation requires more than an infusion of technology – it requires extensive and on-going professional development in technological skills, in learning new socio-constructive pedagogical approaches and a sensitivity to help young learners build on their first language skills. It also requires a change in political will to recognize and value a multilingual approach that will reflect the future make-up of American society.

Of course, much of what is said about ICT and language might also be applied to other subject matter areas in school. While I don’t disagree that solid language skills are fundamental to school success, the same lack of comprehensive planning to lever ICT to accelerate learning and remediation is also found in subjects such as math, social studies and science. The school system is ultra conservative and the overselling of technology in the past has left a bitter legacy of highly expensive computers and little impact in terms of academic achievement.

Technology-mediated Learning Environments for Young English Learners is not a recipe book for success. It is a collection of perspectives that may help re-define the start line for initiatives in the use of technology by young language learners. This language challenge is not only an American issue. It exists wherever education is offered in a language other than one with which the child is comfortable. While some children thrive and become natural polyglots, others struggle to keep up with the dual demands of learning the content and learning the language at the same time. They fall further and further behind and often drop out. In my Canadian experience with minority language francophone children, those who are uncomfortable in the language of instruction can fall into the safety net of the neighbourhood English school. But seldom is there such a safety net in the USA. Perhaps, innovators and policy makers who read Parker’s collection, might weave one with technology.

References


Tips for Teaching with CALL: Practical Approaches to Computer-assisted Language Learning
(Book Review)

Reviewer:
Ferit Kilickaya
Middle East Technical University
Faculty of Education
Department of Foreign Language Education
06531 Ankara/TURKEY
kilickay@metu.edu.tr
ferit.kilickaya@gmail.com
http://www.metu.edu.tr/~kilickay

Textbook Details:
Tips for Teaching with CALL: Practical Approaches to Computer-assisted Language Learning
Carol Chapelle and Joan Jamieson
Publisher: Pearson Education
ISBN: 978-0-13-240428-0
256 Pages

Most teachers are interested in introducing technology to their students to assist them with opportunities to learn English. Yet, some of them may be easily intimidated by technology and/or need practical hands-on techniques for using CALL (Computer-assisted language learning) with their students. Moreover, they need to know how as well as why while benefiting from CALL.

The book “Tips for Teaching with CALL: Practical Approaches to Computer-assisted Language” introduces CALL to English language teachers and students by providing practical techniques and tips as well as the rationale for using any specific CALL activity with their students.

While discussing the websites, software and the practical techniques, the authors raise awareness the rationale behind such techniques by summarizing the current research related to the use of CALL materials.

What impressed me first about the book is how it is organized. Chapters are devoted to specific skills such as Vocabulary and Grammar, which I believe is a very convenient way to present activities. Moreover, the introduction entitled “What is CALL?” presents a precise and awareness-raising explanation of what CALL is and what roles teachers and computers play while guiding students.

Chapter 1 introduces CALL materials for Vocabulary teaching such as WordSmart, Vocaboly and Gerry’s Vocabulary Teacher.

Chapter 2 focuses on Grammar and presents the websites such as ESLgold and Chemnitz Internet Grammar, and the software such as Understanding and Using English Grammar and Newspaper Editor.

Chapter 3 deals with Reading with the help of CALL materials on the websites such as ESL Reading and Adult Learning Activities and the software such as Longman English Interactive 3 and Issues in English.

Chapter 4 addresses another skill, namely Writing, by presenting materials from Easy Writer, Linguistic Funland and other websites and software.

Chapter 5 deals with Listening by introducing the readers into popular websites such as Randall’s ESL Cyber Listening Lab and ELLLO: English Language Listening Lab Online, and the software such as Rosetta Stone and Longman English Interactive (LEI) 3.
Chapter 6 exemplifies the activities that can be used for Speaking and discusses the software including *Speech Works 4* and *text-to-speech* applications.

Chapter 7 addresses the issue of communication tools that will help students learn English. Popular tools such as *Microsoft® Windows Messenger™*, *Nicenet* and *forums* such as *UsingEnglish.com* are presented in this chapter.

Chapter 8 discusses how technology and communication tools can be applied to content-based instruction.

It is important to note here again that each chapter not only introduces websites/software in order to help students learn English but also equip teachers with what research says about the use of CALL materials. Better still, each chapter is ended with a special section presenting guidelines while selecting CALL for the activities they are planning to do.

The book, however, seems to have ignored some of the very useful tools/websites that are available on the Internet. I would have liked to read about, in Chapter 4, the use of Blogs and Wikis, especially PBwiki (Kilickaya, 2008), which I think lend themselves to collaborative writing; see the inclusion of the popular website “WordChamp” in Chapter 1 (Kilickaya, 2007), which is very useful for vocabulary learning and teaching. Moreover, another chapter on *Online Teaching and Learning* which discusses the use of tools such as WiZiQ (Kilickaya, in press), Dokeos and Moodle would be very beneficial. A small discussion or a set of links about virtual learning (Second Life), Mobile Learning and Simulation Gaming would be highly appreciated.

Tips for Teaching with CALL: Practical Approaches to Computer-assisted Language Learning really does contribute to English language teaching and especially ELT methodology and materials design, in that it helps language teachers be aware of the materials and activities that are offered by CALL. I suggest any teachers interested in CALL have this extremely valuable book together with the ones by Dudaney and Hockly (2007) and Erben and Sarieva (2008) as they complement each other in terms of discussion of specific tools, research findings and step-by-step instruction.

**References**


