Journal of Educational Technology & Society

Published by
International Forum of Educational Technology & Society

January 2010
Volume 13 Number 1

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 The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to *Educational Technology & Society* and three months thereafter.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

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.

ISSN 1436-4522 (online) and 1176-3647 (print).

Abstracting and Indexing

Educational Technology & Society is abstracted/indexed in Social Science Citation Index, Current Contents/Social & Behavioral Sciences, ISI Alerting Services, Social Scisearch, ACM Guide to Computing Literature, Australian DEST Register of Refereed Journals, Computing Reviews, DBLP, Educational Administration Abstracts, Educational Research Abstracts, Educational Technology Abstracts, Elsevier Bibliographic Databases, ERIC, Inspec, Technical Education & Training Abstracts, and VOCED.

Guidelines for authors

Submissions are invited in the following categories:

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

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

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

Submission procedure

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

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

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

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

The submissions should be uploaded at http://www.ifets.info/ets_journal/upload.php. In case of difficulties, 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.

Journal of Educational Technology & Society

Volume 13 Number 1 2010

Table of contents

Book review(s)

Intelligent Tutoring Systems (Guest Editorial)

Roger Nkambou
Computer Science Department, University of Quebec at Montreal, QC, Canada // nkambou.roger@uqam.ca

Intelligent Tutoring Systems (ITS) are meant to provide useful tutoring services for assisting the student. These services include coaching, assisting, guiding, helping, and tracking the student during problem-solving situations. To offer high-quality tutoring services, an ITS must be able to establish the correct student profile, then understand and diagnose the student cognitive as well as its affective state. This special issue of Educational Technology & Society presents recent works dealing with those matters.

Extracting Procedural Models Using Educational Data Mining

The main goal of an intelligent tutoring system is to actively provide guidance to the student in problem-solving situations. Relevant feedback should be founded on a thorough understanding and diagnosis of student responses. Building such understanding and diagnosis model is a difficult issue that is also a time-intensive process involving human experts. This issue becomes even more difficult in ill-defined domains where an explicit representation of the training task is hard, if not impossible, to set up. Educational data-mining (EDM) brings some promising solutions to this issue.

You will find in this special issue two EDM-based solutions proposed for coping with this problem. Each of these solutions consists of a model that can constantly learn from new learner or user data and thus, guaranties that the tutor provides an up-to-date feedback.

In one hand, Barnes and Stamper propose a novel application of Markov decision processes (MDPs) to automatically generate hints for an intelligent tutor that learns. This approach eases the process of building the understanding and diagnosis model of student actions. The authors extracted MDPs from four semesters of student solutions created in a logic proof tutor, and calculated the probability of being able to generate hints for students at any point in a given problem. The results indicate that extracted MDPs and their proposed hint-generating functions are able to provide hints over 80% of the time. The results also indicate that they can provide valuable tradeoffs between hint specificity and the amount of data used to create an MDP.

In the other hand, Fournier-Viger *et al*. present a novel framework for adapting the behavior of intelligent agents based on human experts' data. The framework consists of an extended sequential pattern-mining algorithm that, in combination with association rule discovery techniques, is used to extract temporal patterns and relationships from the behavior of human learners of multiple profiles, executing a procedural task. The proposed framework has been integrated within CanadarmTutor, an intelligent tutoring system aimed at helping students solve procedural problems that involve moving a robotic arm in a complex virtual environment. CanadarmTutor acts in an ill-defined domain where the problem space associated with a given task consists of an infinite number of paths. The framework was used to improve the behavior of a cognitive agent that adapts its decision by learning from data gathered during past cognitive cycles. The results of the experimentation demonstrate the benefits of the framework for tutoring systems acting in ill-defined domains.

Filling the Gap Between Student Profiles Through Metacognitive Problem-Solving Strategy

One benefit of tutoring is of narrowing, even eliminating the gap between High and Low learners. Low learners are those who are more sensitive to variations in learning environments. Effective ITS should narrow the gap as much as possible without pulling the High learners down. In their paper, Chi and VanLehn present a study that investigates this issue. The study involved two groups of college students who studied probability first and then physics. The experimental group studied probability with Pyrenees, an ITS that explicitly taught and required them to employ a general problem-solving strategy; the control group studied probability with Andes, an ITS that does not teach or require any particular strategy. During subsequent physics instruction, both groups used Andes.

Results showed that an Intelligent Tutoring System teaching a domain-independent problem-solving strategy indeed closed the gap between High and Low learners, not only in the domain where it was taught (probability) but also in a second domain where the strategy had not been taught (physics). The strategy includes two main components: one is solving problems via Backward-Chaining (BC) from goals to givens, named the BC-strategy, and the other is drawing students' attention on the characteristics of each individual domain, named the principle-emphasis skill. Evidence suggests that the Low experimental group transferred the principle-emphasis skill to physics while the High experimental apparently already possessed it and thus mainly transferred the BC-strategy.

Coping with Affective Issues in Tutoring Systems

Considering learners' affective responses during learning episodes is a key issue for more effective tutoring dialogue. Hence, recent work has begun to investigate the emotions experienced during learning in a variety of environments. McQuiggan *et al*. contribute to this effort by investigating the likelihood of affective transitions that occur throughout narrative-centered learning experiences. The study was conducted with the Crystal Island, a learning environment in which narrative is used as a mechanism to contextualize learning.

The results suggest two directions for future work. First, they call for investigation of what type of feedback pedagogical agents should consider when empathy does not promote desirable affective states for learning. For instance, reactive empathy was likely to encourage transitions to either flow or frustration. Second, analysis of individual differences is necessary to determine the affective transitions common across a variety of demographics such as gender, but also across learning attributes such as efficacy, goal orientation, interest, and abilities to self-regulate both learning and affect.

Barnes, T., & Stamper, J. (2010). Automatic Hint Generation for Logic Proof Tutoring Using Historical Data. *Educational Technology & Society, 13* (1), 3–12.

Automatic Hint Generation for Logic Proof Tutoring Using Historical Data

Tiffany Barnes and John Stamper
University of North Carolina at Charlotte, Computer Science Department, Charlotte, NC, USA // Tiffany.Barnes@gmail.com // John@stamper.org

ABSTRACT

In building intelligent tutoring systems, it is critical to be able to understand and diagnose student responses in interactive problem solving. However, building this understanding into a computer-based intelligent tutor is a time-intensive process usually conducted by subject experts. Much of this time is spent in building production rules that model all the ways a student might solve a problem. In our prior work, we proposed a novel application of Markov decision processes (MDPs) to automatically generate hints for an intelligent tutor that learns. We demonstrate the feasibility of this approach by extracting MDPs from four semesters of student solutions in a logic proof tutor, and calculating the probability that we will be able to generate hints for students at any point in a given problem. Our past results indicated that extracted MDPs and our proposed hint-generating functions will be able to provide hints over 80% of the time. Our results also indicated that we can provide valuable tradeoffs between hint specificity and the amount of data used to create an MDP.

Keywords

Educational data mining, Hint generation, Intelligent tutoring, Propositional logic proofs

Introduction

According to the Joint Task Force on Computing Curricula (2005), discrete mathematics is a core course in computer science, and an important topic in this course is solving formal logic proofs. However, this topic is of particular difficulty for students, who are unfamiliar with logic rules and manipulating symbols. To allow students extra practice and help in writing logic proofs, we are building an intelligent tutoring system on top of our existing proof-verifying program. Results from student surveys and our experience in teaching discrete math indicate that students particularly need hints when they get stuck.

The problem of offering individualized help is not unique to logic proofs. Through adaptation to individual learners, intelligent tutoring systems (ITS) can have significant effects on learning (Anderson & Gluck, 2001). However, building one hour of adaptive instruction takes between 100 and 1000 hours of work for subject experts, instructional designers, and programmers (Murray, 1999), and a large part of this time is used in developing production rules that model student behavior and progress. A variety of approaches have been used to reduce the development time for ITSs, including ITS authoring tools and constraint-based student models. ASSERT is an ITS authoring system that uses theory refinement to learn student models from an existing knowledge base and student data (Baffes & Mooney, 1996). Constraint-based tutors, which look for violations of problem constraints, require less time to construct and have been favorably compared to cognitive tutors, particularly for problems that may not be heavily procedural (Mitrovic, Keodinger, & Martin, 2003). However, constraint-based tutors can only provide condition-violation feedback, not goal-oriented feedback that has been shown to be more effective (Van Lehn, 2006).

Some systems use teacher-authored or demonstrated examples to develop ITS production rules. RIDES is a “Tutor in a Box” system used to build training systems for military equipment usage, while DIAG was built as an expert diagnostic system that generates context-specific feedback for students (Murray, 1999). These systems cannot be easily generalized, however, to learn from student data. In example-based authoring tools, teachers work problems in what they predict to be common correct and incorrect approaches, and then annotate the learned rules with appropriate hints and feedback. The Cognitive Tutors Authoring Tool (CTAT) has been used to develop example-based tutors for genetics, Java, and truth tables (Koedinger, Aleven, Heffernan, McLaren, & Hockenberry, 2004). This system has also been used with data to build initial models for an ITS, in an approach called Bootstrapping Novice Data (BND) (McLaren, Koedinger, Schneider, Harrer, & Bollen, 2004). However, in both of these approaches, considerable time must still be spent in identifying student approaches and creating appropriate hints.

Machine learning has also been used to improve tutoring systems. In the ADVISOR tutor, machine learning was used to build student models that could predict the time students took to solve arithmetic problems, and to adapt instruction to minimize this time while meeting teacher-set instructional goals (Beck, et al., 2000). In the Logic-ITA tutor, student data was mined to create hints that warned students when they were likely to make mistakes using their current approach (Merceron & Yacef, 2005). Another logic tutor called the Carnegie Proof Lab uses an automated proof generator to provide contextual hints (Sieg, 2007).

Similar to the goal of BND, we seek to use student data to directly create student models for an ITS. However, instead of feeding student behavior data into CTAT to build a production rule system, our method generates Markov decision processes (MDPs) that represent all student approaches to a particular problem, and uses these MDPs directly to generate hints. This method of automatic hint generation using previous student data reduces the expert knowledge needed to generate intelligent, context-dependent hints and feedback. The system is capable of continued refinement as new data is provided. In this work, we demonstrate the feasibility of our hint generation approach through simulation experiments on existing student data.

Background and Proofs Tutorial context

The Proofs Tutorial is a computer-aided learning tool on NovaNET (http://www.pearsondigital.com/novanet/). This program has been used for practice and feedback in writing proofs in university discrete mathematics courses taught by the first author and others at North Carolina University since 2002 and at UNC Charlotte since 2006. In the Proofs Tutorial, students are assigned a set of 10 problems that range from simpler logical equivalence applications to more complex inference proofs. In the tutorial, students type in consecutive lines of a proof, which consist of four parts: the statement, reference lines, the axiom used, and the substitutions that allow the axiom to be applied. After the student enters these four parts, the statement, reference lines, axiom, and substitutions are verified. If a mistake is made, a warning message is shown, and the line is deleted (but saved for later analysis). In this work, we examine student solutions to Proof 1. Table 1 lists an example of a student solution that includes three errors.

In Barnes (2006), the first author has applied educational data mining to analyze completed formal proof solutions for automatic feedback generation. However, this work did not take into account student errors, and could only provide general indications of student approaches, as opposed to feedback tailored to a student's current progress. In Stamper (2006), the second author performed a pilot study to extract MDPs for a simple proof from three semesters of student data, and verified that the extracted rules conformed to expert-derived rules and generated buggy rules that surprised experts. In Barnes & Stamper (2007), we used visualization tools to explore how to generate hints based on MDPs extracted from student data. In Croy, Barnes, & Stamper (2007), we applied the technique to visualize student proof approaches to allow teachers to identify problem areas for students. This was just one method used to identify students at risk of falling behind in the course.

Table 1. Sample Proof 1 solution

	Statement	Line	Reason
1.	a → b		Given
2.	c → d		Given
3.	¬ (a → d)		Given
	¬ a v d	**3**	**rule IM (error)**
4.	a ^ ¬ d	3	rule IM implication
5.	a	4	rule S simplification
	b	**4**	**rule MP (error)**
	b	**1**	**rule MP (error)**
6.	b	1,5	rule MP modus ponens
7.	¬ d	4	rule S simplification
8.	¬c	2,7	rule MT modus tollens
9.	b ^ ¬c	6,8	rule CJ conjunction

Markov decision processes

A Markov decision process (MDP) is defined by its state set S, action set A, transition probabilities P, and a reward function R (Sutton & Barto, 1998). On executing action a in state s, the probability of transitioning to state s' is denoted $P(s' \mid s, a)$ and the expected reward associated with that transition is denoted $R(s' \mid s, a)$. For a particular point in a student's proof, our method takes the current statements and conclusion as the state, and the student's input as the action. Therefore, each proof attempt can be seen as a graph with a sequence of states (each describing the solution up to the current point), connected by actions. Specifically, a state is represented by the list of statements generated in the student attempt, and actions are the axioms used at each step.

We combine all student solution graphs into a single graph by taking the union of all states and actions and mapping identical states to one another. Once this graph is constructed, it represents all of the paths students have taken in working a proof. Typically at this step, value iteration is used to find an optimal solution to the MDP. For the experiments in this study, we set a large reward for the goal state (100), penalties for incorrect states (10), and a cost for taking each action (1). Setting a non-zero cost on actions causes the MDP to penalize longer solutions (but we set this at 1/10 the cost of taking an incorrect step). These values may need to be adjusted for different sizes of MDPs. We apply the value iteration technique using a Bellman backup to assign reward values to all states in the MDP (Sutton & Barto, 1998). The equation for calculating values $V(s)$ for each state s, where $R(s)$ is the reward for the state, γ is the discount factor (set to 1), and $P_a(s,s')$ is the probability that action a will take state s to state s':

$$V(s) = R(s) + \gamma \cdot \max_a \left(\sum_{s'} P_a(s,s') \cdot V(s') \right)$$

For value iteration, V is calculated for each state until there is little change in the value function over the entire state space. Once this is complete, the optimal solution in the MDP corresponds to taking a greedy traversal approach in the MDP (Barnes & Stamper, 2007). The reward values for each state then indicate how close to the goal a state is, while probabilities of each transition reveal the frequency of taking a certain action in a certain state.

Generating individualized help

We propose to provide real-time, individualized hints to support on-going student proof construction efforts. As described in (Barnes & Stamper, 2007), we generate an MDP for each problem and use it to generate hints for new students who are solving proofs. Since our tutorials have been used as a computer-aided instructional tool for a number of years, we have many semesters of data from which to create large MDPs for each proof problem. We first use these MDPs to add intelligent hints to every problem. As a new student works a problem, we match each of their states to those in the MDP. If their state is present in the problem's MDP, we enable a hint button to give contextual help.

In Barnes & Stamper (2007), we have proposed several reward functions that could be used in hint generation, including expert, typical, and least error-prone. The reward function we have described herein reflects an expert reward function, where the value for a state reflects the shortest path to the goal state. Given the current state, when the hint button is pressed, we select a reward function for the current student based on his or her student profile. If we have identified the student as an at-risk student, we may select the "least error-prone" reward function for generating hints. On the other hand, high-performing students would likely benefit from expert hints, while students between these two extremes may benefit from hints reflecting typical student behavior (Barnes & Stamper, 2007). After we've selected a reward function, we select the next state with the highest reward value. We create four levels of hints from this state, as follows:
(1) Indicate a goal statement to derive (goal-setting hint).
(2) Tell the student what rule to apply next (rule hint).
(3) Indicate the statements where the rule can be used (pointing hint).
(4) Tell the student both the rule and the statements to combine (bottom-out hint).

The hint sequence is constructed to provide goal-setting, pointing, and bottom-out hints. Hints that help students set intermediate goals have been shown to be effective in (McKendree, 1990). Pointing hints help focus user attention, while bottom-out hints essentially tell students the answer (Van Lehn, 2006).

We plan to limit the number of hints a student can use and still receive credit for working the problem. We believe that four hints is a fair number, to be used on a single state in sequence as above or on separate states in the same problem. This results in giving the student one full step of the proof, or allowing rule hints up to four times.

If a student’s state is not found in the MDP, the hint button will be disabled. The student can then get the tutor’s built-in feedback that indicates the correctness of each step, but will not get strategic help. However, we can add the student’s action and its correctness to our database and periodically run value iteration to update the reward function values. Before an update is applied, we will test the update to be sure that the instructors agree with the generated hints.

Method

This experiment uses data from the four fall semesters of 2003–2006, during which an average of 220 students took the discrete math course at NC State University. Students in this course were typically engineering and computer science students in their second or third year of college, but most had not been exposed to a course in logic. Students attended several lectures on propositional logic and completed online homework in which students completed truth tables and filled in the blanks in partially completed proofs. Students then used the Proofs Tutorial to solve 10 proofs, directly or indirectly. Sixty percent of students used direct proof when solving proof 1. We extracted 537 of students’ first attempts at direct solutions to proof 1 from the Proofs Tutorial.

The data were validated by hand by extracting all statements generated by students and removing those that 1) were false or unjustifiable, or 2) were of improper format. We also removed all student steps using the axioms Conjunction, Double Negation, and Commutative, since students were allowed to skip these steps in the tutorial. After cleaning the data, there were 523 attempts at proof 1. Of these, 381 (73%) were complete and 142 (27%) were partial proofs, indicating that most students completed the proof. The average lengths, including errors, were 13 and 10 steps, respectively, for completed and partial proofs. When excluding errors and removed steps, the average number of lines in each student proof was 6.3 steps.

We performed several experiments to explore the capability of our method to generate automated hints. In each experiment, we isolated the data into training and test sets, where the training set was used to generate the Markov Decision Process (MDP) as described above, and the test set was used to explore hint availability. The process for comparing the test set to the MDP consisted of several steps. Because of the structure of the tutorial, we first removed all error states from the MDP and from student attempts before comparison, since the tutorial provides error messages and deletes the corresponding error from the student proof. Then, each attempt in the test set was mapped onto a sequence of states. For each test state, there are two requirements for a hint to be available: 1) there must be a “matching” state in the MDP, and 2) the “matching” state must not be a leaf in the MDP. The closer the match between a test state and the corresponding MDP state, the more context-specific the hint based on that match is.

Leaves

Because we included partial solutions in our training datasets, there are leaves in the MDPs or, in other words, statements with no subsequent actions taken. Therefore, we potentially had a higher percentage of matches than the number of states where hints could be generated. To investigate this, we examined all leaf nodes, and found that the overwhelming majority of leaves occurred in only one semester, for one student. This means that these states are never matched, and therefore do not count toward being able to provide hints. There were a total of three leaf states that occurred in multiple semesters. For two of these states, they occurred twice in one semester, and once in another. For the remaining leaf, it occurred once in one semester and once in another. Therefore, we have over-counted the number of times we can give hints to students by at most two matches in any semester. Since the minimum number of state visits in any semester is 304, this represents an error of at most 2/304 (0.06%) in any reported statistic.

Matching functions

In our experiments, we considered four matching functions that would allow us to select a source statement for hint generation: 1) ordered, 2) unordered, 3) ordered minus the latest statement, and 4) unordered minus the latest statement. An ordered, or exact, state match means that another student has taken the same sequence of steps in solving the proof. An unordered state match means that there is a state with exactly the same statements, but the states were not necessarily reached in the same order. An “ordered-1” match looks for an exact match between the student's previous state and an MDP state. An "unordered-1" match looks for an unordered match between the student's previous state and an MDP state. Once a match is made, we generate a hint based on knowing the next optimal (highest reward value) step from the matching state. The more specific the match, the more contextualized the hint.

To determine hint availability, we calculated two numbers. The first was the "move matches," the number of test set states or "moves," including duplicates, that had matches in the MDP, divided by the total number of test set states. The second was the "unique matches," where we determined all unique test states and calculated the percentage of these that have matches in the MDP. Move matches gave us a measure of the percentage of the time a particular student was able to attain a hint while working a proof. Unique matches gave us a measure of the percentage of overlap in the states in the test set and the MDP.

We conducted two experiments to test the feasibility of automated hint generation. The first was something like a cross-validation study, comparing the hints we could generate using various semesters of data for MDP creation. The second was a simulation of creating MDPs incrementally as students worked proofs and calculating the probability of being able to generate hints as new attempts were added to the MDP.

Experiment 1: Comparing classes

In this experiment, we explored the ability of our system to provide hints after one, two, three, or four semesters of data were used to build MDPs. Table 2 shows that each semester was used as a test set (denoted by f and the semester), while all the remaining semesters were used as training sets for MDPs. For example, when fall 2003 was used as test set f3, it was compared with MDPs created from one semester of data each (e.g., M4 = fall 2004), two semesters of data each (e.g., M45 = fall 2004 and 2005), and three semesters of data (e.g., M456 = fall 2004 to 2006).

Table 2. Experimental design for comparing classes

Test	1-sem. MDP	2-sem. MDP	3-sem. MDP
f3	M4, M5, M6	M45, M46, M56	M456
f4	M3, M5, M6	M35, M36, M56	M356
f5	M3, M4, M6	M34, M36, M46	M346
f6	M3, M4, M5	M34, M35, M45	M345

This experiment provides us insight into the number of semesters of data we might need to provide hints a reasonable percentage of the time while students are solving proofs. Table 3 presents the data for each semester. We note that semester fall 2005 was unusual: there was a small number of states, but a large number of moves, suggesting that students solved these proofs in very similar ways.

Table 3. Semester data, including attempts, moves, and MDP states

Semester	# Attempts	MDP states	# Moves
f3	172	206	711
f4	154	210	622
f5	123	94	500
f6	74	133	304

We hypothesized that we could provide hints a majority of the time, using just one semester as our MDP training data. Table 4 shows the percentage of ordered matches between each semester and the remaining combinations of training sets. We were very encouraged by these data, which suggest that our system would provide highly contextualized hints over 66% of the time, in the worst case, after just one semester of training. In all cases, adding more data increased the probability of providing hints, though we do see diminishing returns when comparing the marginal increase between 1–2 (6.8%) and 2–3 (2.8%) semesters of data.

Table 4. Average % move matches using the ordered function

Test set	1-sem. MDPs	2-sem. MDPs	3-sem. MDPs
f3	68.73%	75.67%	78.62%
f4	69.77%	77.71%	81.03%
f5	86.33%	90.80%	92.00%
f6	66.34%	74.12%	77.63%
Average	72.79%	79.57%	82.32%

Tables 5–7 show the results of this experiment using the unordered, ordered-1, and unordered-1 matching techniques. These results show consistent increases within each table, going from 1-semester MDPs up to 2-semester MDPs, as expected. However, the increases between 2- and 3-semester MDPs are decreasing, suggesting consistent diminishing returns for adding more data to the MDPs.

Table 5. Average % move matches using the unordered function

Test set	1-sem. MDPs	2-sem. MDPs	3-sem. MDPs
f3	76.62%	82.16%	84.37%
f4	75.35%	81.99%	84.41%
f5	91.93%	94.40%	95.40%
f6	74.56%	82.35%	84.87%
Average	79.62%	85.22%	87.26%

Table 6. Average % move matches using the ordered-1 function

Test set	1-sem. MDPs	2-sem. MDPs	3-sem. MDPs
f3	76.92%	85.14%	89.00%
f4	76.26%	85.69%	90.35%
f5	90.78%	96.19%	97.80%
f6	75.55%	84.32%	89.14%
Average	79.88%	87.84%	91.57%

Table 7. Average % matches using the unordered-1 function

Test set	1-sem. MDPs	2-sem. MDPs	3-sem. MDPs
f3	82.63%	89.19%	91.99%
f4	81.73%	90.14%	93.41%
f5	94.60%	97.00%	98.00%
f6	81.03%	89.69%	92.43%
Average	85.00%	91.50%	93.96%

Table 8 lists the average percentage of matches for each of our experiments using the four match functions. This table gives an indication of the tradeoffs between using multiple semesters of data versus multiple techniques for matching. Here, we see that, on average, for 72% of moves, we can provide highly contextualized (ordered) hints using just one semester of data. With two semesters of data, we can provide these hints almost 80% of the time, but this only increases to 82% for three semesters of data. If we wished to provide hints after collecting just one semester of data, we could provide less contextualized hints for those who didn't have ordered matches in the MDP. There is a nearly identical benefit to providing hints using unordered versus ordered-1 searches, increasing the match rate to almost 80%. We did not calculate the marginal benefit of providing one of these over the other. However, we can provide hints to an additional 5% of students if we add the unordered-1 match function.

When analyzing these data, we observed a skew in all statistics because of the unusual distribution of states and moves in f5. We therefore repeated all experiments excluding f5, and the results are given in Table 9. The differences caused by skew in f5 had a smaller effect as you move from top left to bottom right, suggesting that more data or less sensitive matching can mitigate the effect of unusual training data.

Table 8. Comparison of % move matches and matching techniques

Matching	1-sem. MDPs	2-sem. MDPs	3-sem. MDPs
Ordered	72.79%	79.57%	82.32%
Unordered	79.62%	85.22%	87.26%
Ordered-1	79.88%	87.84%	91.57%
Unordered-1	85.00%	91.50%	93.96%

Table 9. Comparison of % move matches, excluding f5

Test set	1-sem. MDPs	2-sem. MDPs
Ordered	70.97%	78.05%
Unordered	78.69%	83.59%
Ordered-1	79.02%	87.99%
Unordered-1	85.77%	91.86%

Table 10 shows the marginal increase of each matching technique, for each MDP size, to illustrate the tradeoffs between additional data and matching technique. When considering matching functions, the easiest technical change is from ordered to ordered-1, where one statement is removed from the test state before comparison with the MDP states. In all cases, the benefit of providing these hints is higher than that of providing hints based on unordered matches. This is probably because there is some inherent partial ordering in proofs, so only limited benefit is seen from reordering statements. When an ordered hint cannot be matched, it is perhaps more likely that the student has just performed a step that no one else has done before, rather than generating a new ordering of steps, so the benefit of ordered-1 can exceed that of unordered. Providing the unordered search requires us to maintain two separate MDPs (one ordered and one unordered) to make the search more efficient, so there are both time and space tradeoffs to using unordered matching. However, adding unordered-1 after adding unordered provides a very large difference in our capability to provide hints, with little investment in time.

Table 10. Marginal increases when comparing matching techniques to ordered

Technique	1-sem. ordered	2-sem. ordered	3-sem. ordered
Unordered	6.83%	5.65%	4.94%
Ordered-1	7.09%	8.27%	9.25%
Unordered-1	12.21%	11.93%	11.64%

As part of this study we also compared the unique states across semesters, as shown in Table 11. This gives us a measure of the percent overlap between MDPs. Three semesters of data with ordered matching or one semester of data with unordered-1 matching will give us over 50% matching of states across MDPs.

Table 11. Unique state % matches across semesters and techniques

Test set	1-sem. MDPs	2-sem. MDPs	3-sem. MDPs
Ordered	34.55%	45.84%	51.93%
Unordered	43.62%	55.23%	59.90%
Ordered-1	48.25%	63.07%	71.39%
Unordered-1	57.28%	71.98%	77.87%

Experiment 2: Exploring the "cold start" problem

One critique of using data to generate hints has been the expected time needed for the method to be applied to a new problem, or in other words, the "cold start" issue. Our hypothesis was that a relatively low number of attempts would be needed to build an MDP that could provide hints to a majority of students. One method for building our hint MDP would be to incrementally add MDP states as students solve proofs. This experiment explores how quickly such an MDP is able to provide hints to new students, or in other words, how long it takes to solve the cold start problem. For one trial, the method is given below.

Let Test = {all 523 student attempts}
Randomly choose and remove the next attempt a from the Test set.
Add *a*'s states and recalculate the MDP.
Randomly choose and remove the next attempt *b* from the Test set.
Compute the number of matches between *b* and MDP.
If Test is non-empty, then let *a* = *b* and go to step 3. Otherwise, stop.

For this experiment, we used the ordered and unordered matching functions, and plotted the resulting average matches over 100,000 trials, as shown in Figure 1. These graphs show a very quick rise in ability to provide hints to students, which can be fit using power functions, whether the system uses ordered or unordered MDP states and matching.

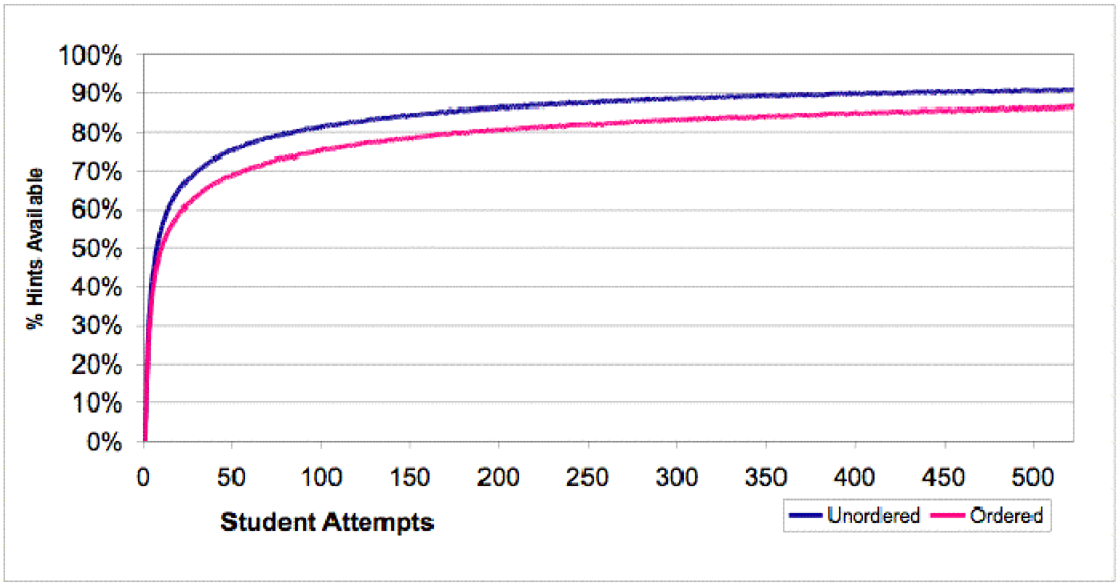

Figure 1. Hints available when the MDP is constructed using a given number of attempts, averaged over 100,000 random orderings of the attempts selected for the MDP.

Clearly, the availability to give hints ramps up very quickly. Table 12 lists the number of attempts needed in the MDP versus target hint percentages. For the unordered matching function, the 50% threshold is reached at just 8 student attempts and the 75% threshold at 49 attempts. For ordered matching, 50% occurs on attempt 11 and 75% on attempt 88. These data are encouraging, suggesting that instructors using our MDP hint generator could seed the data to jump-start new problems. By allowing the instructor to enter as few as 8–11 example solutions to a problem, the method might already be capable of automatically generating hints for 50% of student moves.

Table 12. Attempts needed to achieve threshold % hints levels

	50%	55%	60%	65%	70%	75%	80%	85%	90%
Un-Ordered	8	11	14	20	30	46	80	154	360
Ordered	11	15	22	33	55	85	162	362	N/A

Pilot study on hint generation and availability

We have constructed a hint generator using the methods described herein to add hints to Deep Thought, a visual logic tutor created by Marvin Croy (2000). When a student presses the new hint button, our hint generator searches for the current state in the MDP and checks that a successor state exists. If it does, the successor state with the highest value is used to generate a hint sequence as described above. In our pilot experiment as described in Barnes, Stamper, Lehman, & Croy (2008), hints were added to four logic proof problems in a spring 2008 deductive logic course with 40 students enrolled. MDPs were created for these four problems with 16 to 26 training examples, and the percent hint availability was calculated for all problem states. Based on the results in this study, we predicted that hints (using ordered matching) would be available approximately 56–62% of the time. In the pilot study, if a student had pressed the hint button after every move taken, a hint would have been available about 48% of the time. Although this percentage seems low, we found that when students requested hints, they were available 91% of the time. This suggests that hints are needed precisely where we have data in our MDPs from previous semesters. There are several potential explanations for this: students may be avoiding using hints; hints may be most needed in only a few key steps; or the students may have felt very confident in solving proofs before working these problems. We plan to investigate the reasons for this surprising result in future experiments.

Conclusions and future work

We have proposed and explored the feasibility of an approach to mining Markov decision processes from student work to automatically generate hints. This approach differs from prior work in authoring tutoring systems by mining actual student data, rather than relying on teachers to add examples the system can learn from. Our tutor can already classify many errors students make. Adding the MDP to this tutor enables it to provide hints. This MDP can constantly learn from new student data. We note that on cold start for a new problem that has no student data, the system will still act as a problem-solving environment, but after even one semester of data is collected, a significant number of hints can be generated. As more data are added, more automated hints can be generated. We have implemented this hint system in the Deep Thought logic tutor and are currently testing whether our hints affect overall learning. In our future work, we will continue to explore ways to learn general rules to build intelligent feedback and help with greater coverage and robustness. For instance, we plan to group students according to their proofs behavior and class performance, and create tailored MDPs for each group of students. In an interesting extension, we are investigating ways to determine the usefulness of a problem step based on its frequency in the MDP. This work will allow us to detect and prevent hints that may guide students to perform steps that others have taken, but that did not contribute to the problem solution. This work on measuring the utility of a problem step can also be used to generate help in ill-defined domains.

We believe the feasibility studies presented in this paper provide an important methodology for predicting the reliability of data-driven methods for deriving feedback and hints for students. These methods address teacher concerns regarding the availability of sufficient and accurate help while addressing the need to support student learning with low-cost, scalable methods.

References

Anderson, J., & Gluck, K. (2001). What role do cognitive architectures play in intelligent tutoring systems? In D. Klahr & S. Carver (Eds.), *Cognition & Instruction: 25 years of progress* (pp. 227–262). Mahwah, NJ: Erlbaum.

Baffes, P., & Mooney, R. J. (1996). A novel application of theory refinement to student modeling. *AAAI-96: Proceedings of the 13th National Conference on Artificial Intelligence* (pp. 403–408), Menlo Park, CA: AAAI Press.

Barnes, T. (2006). Evaluation of the q-matrix method in understanding student logic proofs. In G. Sutcliffe & R. Goebel (Eds.), *Proceedings of the Nineteenth International Florida Artificial Intelligence Research Society Conference*, Menlo Park, CA: AAAI Press.

Barnes, T., & Stamper, J. (2007). Toward the extraction of production rules for solving logic proofs, In *Proc. 13th Intl. Conf. on Artificial Intelligence in Education, Educational Data Mining Workshop.* Online proceedings retrieved March 25, 2009, from http://aied.inf.ed.ac.uk/AIED2007/AIED-EDM_proceeding_full2.pdf.

Barnes, T., Stamper, J, Lehman, L., & Croy, M. (2008). A pilot study on logic proof tutoring using hints generated from historical student data. *Proceedings of the 1st Annual International Conference on Educational Data Mining*. Montreal, CA, June 20–21, 2008, Retrieved March 25, 2009, from http://www.educationaldatamining.org/EDM2008/index.php?page=proceedings.

Beck, J., Woolf, B., & Beal, C. (2000). ADVISOR: A machine learning architecture for intelligent tutor construction. In *AAAI-96: Proceedings of the 7th National Conference on Artificial Intelligence* pp. (552–557). Menlo Park, CA: AAAI Press.

Croy, M. J, (2000). Problem solving, working backwards, and graphic proof representation, *Teaching Philosophy 23*, 169–187.

Croy, M., Barnes, T., & Stamper, J. (2007). Towards an intelligent tutoring system for propositional proof construction. In P. Brey, A. Briggle, & K. Waelbers (Eds.), *Proceedings of the 2007 European Computing And Philosophy Conference*, Amsterdam, Netherlands: IOS Publishers.

Joint Task Force for Computing Curricula, ACM, AIS, & IEEE. (2005). Computing curricula 2005. Retrieved March 26, 2009, from http://www.acm.org/education/education/curric_vols/CC2005-March06Final.pdf

Koedinger, K., Aleven, V., Heffernan, N., McLaren, B., & Hockenberry, M. (2004, August). Opening the door to non-programmers: Authoring intelligent tutor behavior by demonstration. *Proceedings of the 7th Intl. Conf. ITS-2004: Intelligent Tutoring Systems.* Maceió, Brazil: Springer.

McKendree, J. (1990, December). Effective feedback content for tutoring complex skills. *Human-Computer Interaction, 5*(4), pp. 381–413.

McLaren, B., Koedinger, K., Schneider, M., Harrer, A., & Bollen, L. (2004, August). Bootstrapping Novice Data: Semi-automated tutor authoring using student log files, In *Proc. Workshop on Analyzing Student-Tutor Interaction Logs to Improve Educational Outcomes*, *Proceedings of the 7th Intl. Conf. ITS-2004: Intelligent Tutoring Systems.* Maceió, Brazil: Springer.

Merceron, A., & Yacef, K. (2005). Educational data mining: A case study. In *12th Intl. Conf. Artificial Intelligence in Education.* Amsterdam, Netherlands: IOS Press.

Mitrovic, A., Koedinger, K. & Martin, B. (2003). A comparative analysis of cognitive tutoring and constraint-based modeling. *User Modeling 2003: Proceedings of the 9th International Conference, UM 2003, Johnstown, PA, USA, June 22–26, 2003, Lecture Notes in Computer Science* (pp. 313–322). Berlin: Springer.

Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *Intl. J. Artificial Intelligence in Education, 10*: 98–129.

Sieg, W. (2007). The AProS Project: Strategic thinking & computational logic. *Logic Journal of IGPL*. Retrieved March 26, 2009, from http://jigpal.oxfordjournals.org/cgi/content/abstract/jzm026v1

Stamper. J. (2006, September). Automating the generation of production rules for intelligent tutoring systems. *Proc. Intl. Conf. Computer-Aided Learning (ICL2006)*, Villach, Austria: Kassel University Press.

Sutton, R., and Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Van Lehn, K. (2006). The behavior of tutoring systems, *International Journal of Artificial Intelligence in Education 16*, 227–265.

Fournier-Viger, P., Faghihi, U., Nkambou, R., & Mephu Nguifo, E. (2010). Exploiting Sequential Patterns Found in Users' Solutions and Virtual Tutor Behavior to Improve Assistance in ITS. *Educational Technology & Society*, *13* (1), 13–24.

Exploiting Sequential Patterns Found in Users' Solutions and Virtual Tutor Behavior to Improve Assistance in ITS

Philippe Fournier-Viger, Usef Faghihi, Roger Nkambou and Engelbert Mephu Nguifo*
Department of Computer Sciences, University of Quebec in Montreal, Montreal, Canada
fournier_viger.philippe@courrier.uqam.ca // faghihi.usef@courrier.uqam.ca // nkambou.roger@uqam.ca
*Department of Computer Sciences, Université Blaise-Pascal, Clermont-Ferrand, France mephu@isima.fr

ABSTRACT

We propose to mine temporal patterns in Intelligent Tutoring Systems (ITSs) to uncover useful knowledge that can enhance their ability to provide assistance. To discover patterns, we suggest using a custom, sequential pattern-mining algorithm. Two ways of applying the algorithm to enhance an ITS's capabilities are addressed. The first is to extract patterns from user solutions to problem-solving exercises for automatically learning a task model that can then be used to provide assistance. The second way is to extract temporal patterns from a tutoring agent's own behavior when interacting with learner(s). In this case, the tutoring agent reuses patterns that brought states of "self-satisfaction." Real applications are presented to illustrate the two proposals.

Keywords

Temporal patterns, Sequential pattern mining, Educational data mining, Intelligent tutoring systems

Introduction

Using knowledge discovery techniques to uncover useful knowledge hidden in a massive amount of educational data has been the subject of much recent research (Baker, Barnes, & Beck, 2008). However, no research has considered mining temporal patterns in Intelligent Tutoring Systems (ITSs) and employed this knowledge to improve their ability to provide assistance. In this paper, we propose two ways of improving the behavior of ITSs by exploiting temporal patterns. Those are to (1) automatically learn task models from recorded novice and expert users' solutions to provide assistance to learners, and (2) building tutoring agents that can adapt their behavior to learners and situations by reusing previously successful patterns of tutoring actions. Our hypothesis is that temporal patterns found in ITSs constitute useful knowledge that can be exploited to improve their ability to provide relevant and adaptive assistance.

The paper is organized as follows. First, it introduces an algorithm for mining temporal patterns. Next, the paper describes two proposals based on this algorithm and describes how they are integrated in an ITS. Finally, the last section presents our conclusions and previews our future work.

Mining temporal patterns from sequences of events

According to Han & Kamber (2006), there are four main kinds of patterns that can be mined from temporal data. These are trends, similar sequences, sequential patterns, and periodical patterns. In this work we chose to mine sequential pattern (Agrawal & Srikant, 1995), as we are interested in finding relationships between occurrences of events that are logged in ITSs. To mine sequential patterns, several algorithms have been proposed (Han & Kamber, 2006). While classical sequential pattern-mining algorithms have for their only goal to discover sequential patterns that occur frequently in several transactions of a sequence database (Agrawal & Srikant, 1995), other algorithms have proposed numerous extensions to the problem of mining sequential patterns (Han & Kamber, 2006). For this work, we chose a sequential pattern-mining algorithm that we have developed (Fournier-Viger, Nkambou, & Mephu Nguifo, 2008a), as it provides several more features than classical sequential pattern algorithms, such as accepting symbols with numeric values, eliminating redundancy, and handling time constraints and contextual information. For a technical description of the algorithm, the reader can refer to Fournier-Viger et al. (2008a). Moreover, a Java implementation of the algorithm can be downloaded from http://www.philippe-fournier-viger.com/spmf/.

The algorithm takes as input a database D of sequences of events. An event $X = (i_1, i_2,\dots i_n)$ contains a set of items $i_1, i_2,\dots i_n$, that are considered simultaneous, and where each item can be annotated with an integer value. Formally, a sequence is denoted $s = <(t_1, X_1), (t_2, X_2), \ldots (t_n, X_n)>$ where each event X_k is associated to a timestamp t_k indicating the time of the event. For example, the sequence *S1* of figure 1 (left) contains two events. It indicates that item *a* appeared with a value of 2 at time 0 and was followed by items *b* and *c* with a value of 0 and 4, respectively, at time 1. An event sequence $sa = <(ta_1, A_1), (ta_2, A_2), \ldots (ta_n, A_n)>$ is said to be contained in another event sequence $sb = <(tb_1, B_1), (tb_2, B_2), \ldots (tb_m, B_m)>$, if there exist integers $1 \leq k1 < k2 < \ldots < kn \leq m$ such that $A_1 \subseteq B_{k1}, A_2 \subseteq B_{k2, \ldots,} A_n \subseteq B_{kn}$, and that $tb_{kj} - tb_{k1}$ is equal to $ta_j - ta_1$ for each $j \in \{1 \ldots m\}$. The relative support of a sequence *sa* in a database *D* is defined as the percentage of sequences $s \subseteq D$ that contain *sa* and is denoted by *supD(sa)*. The problem of mining frequent sequences is to find all the sequences *sa* such that $supD(sa) \geq minsup$ for a sequence database *D*, given a support threshold *minsup*, and optional time constraints. The optional time constraints are the minimum and maximum time intervals required between the head and tail of a sequence and the minimum and maximum time intervals required between two adjacent events of a sequence.

As an example, figure1 illustrates a database of six sequences (left) and the corresponding patterns found for a *minsup* of 33% (right). Consider pattern M5. This pattern appears in sequences S4 and S5, respectively. It has a support of 33% (two out of six sequences) and is frequent. Now consider patterns M1 and M2. Because item *a* appears in sequences S1, S2, S3, and S4, with values 2, 2, 5, and 6, respectively, the algorithm separated these values into two groups to create patterns M1 and M2 instead of creating a single pattern with a support of 66 %. For each of these groups, the median (2 and 5) was kept as an indication of the values grouped. This clustering of similar values only occurs when the support is higher or equal to 2 * *minsup* (see Fournier-Viger et al., 2008a, for details).

ID	Sequences
S1	<(0,a{2}), (1,bc{4})>
S2	<(0,a{2}), (1,c{5})>
S3	<(0,a{5}), (1,c{6})>
S4	< (0,f), (1, g),(2,a{6}e)>
S5	<(0, f b{3}), (1,h),(2,ef) >
S6	<(0,b{2}), (1,d)>

→

ID	Mined sequences	Support
M1	<(0,a{2})>	33 %
M2	<(0,a{5})>	33 %
M3	<(0,a{2}), (1, c{5})>	33 %
M4	<(0,c{5})>	50 %
M5	<(0,f), (2, e)>	33 %
M6	…	…

Figure 1. A database of six sequences (left) and mined sequences (right)

First proposal: Automatically learning task models from users' solutions

Our first proposal is for the acquisition of domain knowledge in an ITS. Typically, domain experts have to provide relevant knowledge to an ITS so that it can guide a learner during problem-solving activities. One common way of acquiring such knowledge is to use the method of cognitive task analysis that aims to produce effective problem spaces or task models by observing expert and novice users for capturing different ways of solving problems. However, cognitive task analysis is a very time-consuming process (Aleven, McLaren, Sewall, & Koedinger, 2006), and it is not always possible to define a satisfying complete or partial task model, particularly when a problem is ill-structured. According to Simon (1978), an ill-structured problem is one that is complex, with indefinite starting points, multiple and arguable solutions, or unclear strategies for finding solutions. A domain that includes such problems and in which tutoring targets the development of problem-solving skills is said to be ill-defined (within the meaning of Lynch, Ashley, Aleven, & Pinkwart, 2006). An alternative to cognitive task analysis is constraint-based modeling (CBM) (Mitrovic, Mayo, Suraweera, & Martin, 2001), which consists of specifying sets of constraints on what is a correct behavior, instead of providing a complete task description. Though this approach was shown to be effective for some ill-defined domains, a domain expert has to design and select the constraints carefully, and it cannot support tutoring services such as suggesting next steps to be performed by a learner. Contrarily to these approaches, where domain experts have to provide the domain knowledge, a promising approach is to use knowledge discovery techniques to automatically learn a partial problem space from logged user interactions in an ITS, and to use this knowledge base to offer tutoring services. A few works have been done in this direction in the field of ITS (for example, Riccuci, Carbonaro, & Casadei, 2007; Matsuda, Cohen, Sewall, Lacerda, & Koedinger, 2007; Barnes & Stamper, 2008), but they either (1) require specifying a minimal set of background knowledge, (2) have been applied in well-defined domains, (3) rely on strong assumption such that tasks can be modeled as production rules, or (4) do not take into account learner profiles.

We propose a solution that is not constrained by those limitations. We illustrate this proposal in the context of CanadarmTutor (Kabanza, Nkambou, & Belghith, 2005) (Figure 2) a virtual learning environment for learning how to operate the Canadarm2 robotic arm on the international space station. The main learning activity in CanadarmTutor is to move the arm from one configuration to another. This is a complex task, because the arm has seven joints and the user must chose at any time the three best cameras for viewing the environment from around twelve cameras on the space station and adjust their parameters. We have integrated a tutoring agent in CanadarmTutor to provide assistance to learners during this task. However, there are a very large number of possibilities for moving the arm from one position to another, and because one must also consider the safety of the maneuvers, it is very difficult to define a task model for generating the moves that a human would execute (Fournier-Viger, Nkambou, & Mayers, 2008b). For this reason, instead of providing domain knowledge to the agent, we implemented a learning mechanism that allows the tutoring system to learn by recording the behavior of users performing a task. The tutoring system then uses this knowledge to provide assistance to learners. We describe next the three operation phases of the learning mechanism as they are implemented in CanadarmTutor, and an experiment.

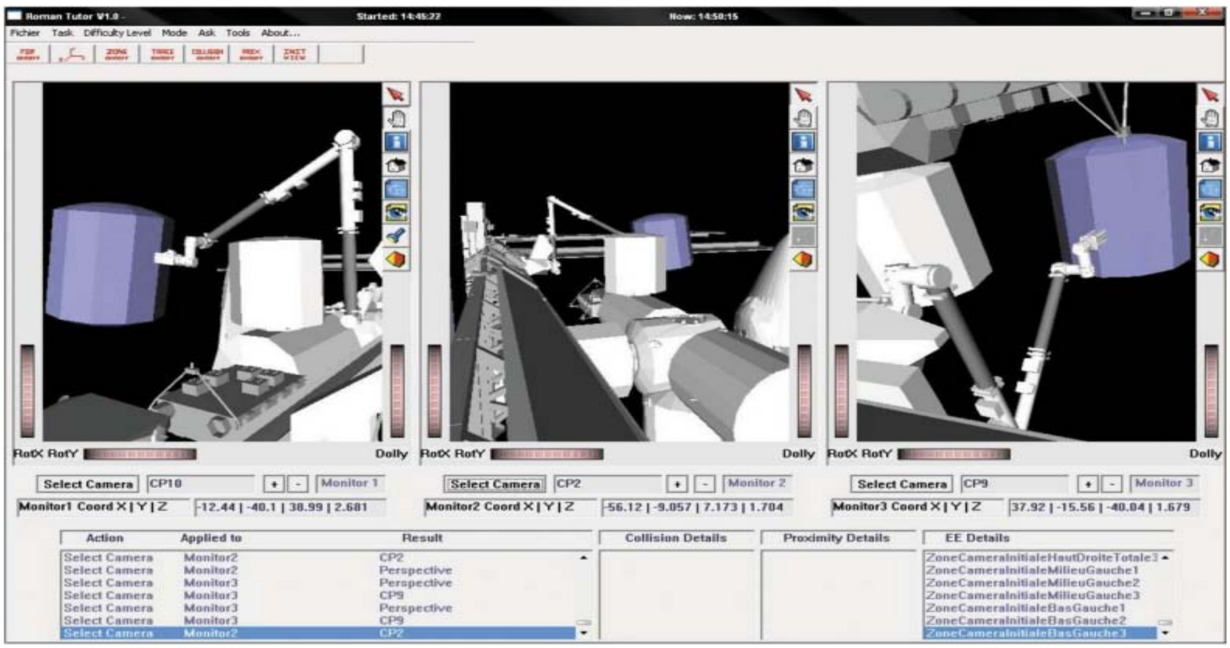

Figure 2. The CanadarmTutor interface

The observing phase

In the first phase, the tutoring system records the solutions of users that attempt an exercise. In CanadarmTutor, an exercise is to move the arm from an initial configuration to a goal configuration. For each attempt, the tutoring system logs a sequence of events. In this context, an event is a set of actions (items) that are considered unordered temporally. We defined 112 primitive actions that can be recorded in CanadarmTutor, which are (1) selecting a camera, (2) performing an increase or decrease of the pan/tilt/zoom of a camera, and (3) applying a rotation value to an arm joint. Actions of types (2) and (3) are annotated with integer values that indicate, respectively, the number of increments/decrements applied to a camera parameter and the number of degrees that a joint is rotated. An example of a partial action sequence recorded for an user in CanadarmTutor is <(0,6{2}),(1,63),(2,53{4}),(3,111{2})> , which represents decreasing the rotation value of joint SP (action 6) by two units, selecting camera CP3 (action 63), and increasing the pan of camera CP2 (action 53) by four units and then its zoom (action 111) by two units.

To annotate sequences with contextual information, we have extended the notion of sequence database with dimensional information, as suggested by Pinto, H., Han, J., Pei, J., Wang, K., Chen, Q., & Dayal, U. (2001). A sequence database having a set of dimensions $D = D_1, D_2, \dots D_n$ is called an MD-Database. Each sequence of an MD-Database (an MD-Sequence) possesses a symbolic value for each dimension or the value "*", which means any value. A set of dimension values is called an MD-Pattern and is denoted *d1, d2, … dn*. An MD-Pattern $Px = \{dx_1,$

dx_2, … dx_n} is said to be contained in another MD-Pattern Py = {dy_1, dy_2, … dy_m} if $dx_1 \subseteq dy_1$, … $dx_n \subseteq dy_m$. The relative support of a sequence (or MD-Pattern) in a sequence database *D* is defined as the percentage of sequences (or MD-Pattern) that contains it. Table 1 shows an example of MD-Database containing six learner plans annotated with five dimensions. The first dimension "Solution state" indicates if the learner plan is a succesful or buggy solution. In the case of CanadarmTutor, values for this dimension are produced by the tutoring system. The four other dimensions of Table 2 are examples of dimensions that can be added manually. Here, whereas the dimension "Expertise" denotes the expertise level of the learner that performed a sequence, "Skill_1", "Skill_2", and "Skill_3" indicate, respectively, if three specific skills were shown by the learner who performed the sequence. This example includes only five dimensions of three main types (skills, expertise level, and solution state). However, our framework can accept any kind of learner information or contextual information encoded as dimensions. In fact, in CanadarmTutor, we used 10 skills and the "solution state" dimension to annotate sequences.

Table 1. An example database containing 6 user solutions

ID	**Dimensions**					**Sequence**
	Solution state	**Expertise**	**Skill_1**	**Skill_2**	**Skill_3**	
S1	successful	expert	yes	yes	yes	<(0, a),(1, bc)>
S2	successful	novice	no	yes	no	<(0, d)>
S3	buggy	expert	yes	yes	yes	<(0, a),(1, bc)>
S4	buggy	intermediate	no	yes	yes	<(0, a), (1, c), (2, d)>
S5	successful	expert	no	no	yes	<(0, d), (1, c)>
S6	successful	novice	no	no	yes	<(0, c), (1, d)>

The learning phase

In the learning phase, the virtual agent applies the algorithm to find all MD-Sequence with a support higher or equal to *minsup*. For mining patterns, we set up the algorithm to mine only sequences of size two or greater, as shorter sequences would not be useful in a tutoring context. Furthermore, we chose to mine sequences with a maximum time interval between two adjacent events of two actions. The benefits of accepting a gap of two is that it eliminates some "noisy" (non-frequent) learners' actions, but at the same time it does not allow a larger gap size that could make it less useful for tracking a learner's actions. As an example, Table 2 shows some patterns that can be extracted from the MD-Database of Table 1, with a *minsup* of two sequences (33%). Consider pattern P3. This pattern represents doing action *b* one time unit (immediately) after action *a*. The pattern P3 appears in MD-sequences S1 and S3. It has thus a support of 33% or two MD-sequences. Because this support is higher or equal to *minsup*, P3 is deemed frequent. Moreover, the dimension values for P3 tell us that this pattern was performed by expert users that possess skills "Skill_1", "Skill_2", and "Skill_3", and that P3 was found in plan(s) that failed, as well as in plan(s) that succeeded.

Another important consideration is that when applying sequential pattern mining, there can be many redundant frequent sequences found. For example, in Table 2, the pattern P1 is redundant because it is included in the pattern P3 and has the same support. To eliminate this type of redundancy, we have adapted our algorithm based on Wang, Han, & Li (2007) and Songram, Boonjing, & Intakosum (2006) to mine closed MD-sequences. Closed MD-sequences are MD-sequences that are not contained in another sequence having the same support. Mining frequent closed MD-sequences has the advantage of greatly reducing the size of patterns found without information loss (Wang et al., 2007). Once patterns have been mined by our sequential pattern-mining algorithm, they form a partial problem space that can be used directly to provide tutoring services. However, one can also edit the patterns or annotate them with tutoring resources, such as textual hints.

Table 2. Some frequent patterns extracted from the dataset of Table 2 for *minsup* = 33%

ID	**Dimensions**					**Sequence**	**Support**
	Solution state	**Expertise**	**Skill_1**	**Skill_2**	**Skill_3**		
P1	*	expert	yes	yes	yes	<(0, a)>	33%
P2	*	*	*	yes	yes	<(0, a)>	50%
P3	*	expert	yes	yes	yes	<(0, a), (1, b)>	33%
P4	successful	*	no	*	*	<(0, d)>	50%
P5	successful	expert	*	*	yes	<(0, c)>	33%
P6	successful	novice	no	*	no	<(0, d)>	33%

The application phase

In the third phase, the tutoring system provides assistance to the learner by using the knowledge learned in the second phase. The basic operation that is used for providing assistance is to recognize a learner's plan. In CanadarmTutor, this is achieved by the plan recognition algorithm RecognizePlan, which is executed after each student action. When RecognizePlan is called for the first time, it iterates on the whole set of patterns found during the learning phase to note all the patterns that include the sequence of actions performed by the learner. If no pattern is found, the algorithm ignores the last action performed by the learner and searches again. This is repeated until the set of matching patterns is not empty or the size of the sequence of student actions is smaller than 2. In our test, removing user actions has shown to improve the effectiveness of RecognizePlan significantly. The next time it is called, it will be called with the set of matching patterns found by its last execution. This ensures that the algorithm will not consider patterns that have been previously rejected.

After performing preliminary tests with RecognizePlan, we noticed that, in general, after more than six actions performed by a learner, it becomes hard to tell which pattern the learner is doing. For that reason, we improved how the CanadarmTutor applies the sequential pattern-mining algorithm to extract a knowledge base. Originally, it mined frequent patterns for a whole problem-solving exercise. We modified our approach to add the notion of "problem states." In the context of CanadarmTutor, where an exercise consists of moving a robotic arm to attain a specific arm configuration, the 3D space is divided into 3D cubes, and the problem state at a given moment is defined as the set of 3D cubes containing the arm joints. An exercise is then viewed as going from a problem state P_1 to a problem state P_f. For each attempt at solving the exercise, CanadarmTutor logs (1) the sequence of problem states visited by the learner $A= P_1, P_2, \dots P_n$ and (2) the list of actions performed by the learner to go from each problem state to the next visited problem state (P_1 to P_2, P_2 to P_3, … $P_{(n-1)}$ to P_n). After many users perform the same exercise, CanadarmTutor extracts sequential patterns from sequences of problem states visited and from sequences of actions performed for going from a problem state to another. To take advantage of the added notion of problem states, we modified RecognizePlan so that every time the problem state changes, RecognizePlan will be called with the set of patterns associated to the new problem state. Moreover, at a coarse grain level, a tracking of the problem states visited by the learners is also achieved by RecognizePlan. This allows connecting patterns for different problem states. We describe next the main tutoring services that a tutoring agent can provide based on the plan-recognition algorithm.

First, a tutoring agent can assess the profile of the learner by looking at the patterns applied. If, for example, a learner applies 80% of the time patterns with value "intermediate: for dimension "expertise," then CanadarmTutor can assert with confidence that the learner expertise level is "intermediate." In the same way, CanadarmTutor can diagnose mastered and missing/misunderstood skills for users who demonstrated a pattern by looking at the "skills" dimensions of patterns applied, and can estimate other aspects of a learner's profile. This results in rich information that can be used in various ways by a tutoring system. An example is given by the next tutoring service.

Second, a tutoring agent can guide the learner. This tutoring service consists of determining the possible actions from the current problem state and proposing one or more actions to the learner. In CanadarmTutor, this functionality is triggered when the student selects "What should I do next?" in the interface menu. CanadarmTutor then identifies the set of possible next actions according to the matching patterns found by RecognizePlan. The tutoring service then selects the action among this set that is associated with the pattern that has the highest relative support and that is the most appropriate for the estimated expertise level and skills of the learner. If the selected pattern contains skills that are not considered mastered by the learner, CanadarmTutor can use tutoring resources to explain them. If no actions can be identified, CanadarmTutor can rely on a special path planner to generate approximate solutions (see Kabanza et al., 2005 for details). In this current version, CanadarmTutor interacts with the learner only upon request. But it would be possible to program CanadarmTutor so that it can intervene if the learner is following an unsuccessful pattern or a pattern that is not appropriate for its expertise level. Testing different tutorial strategies with learners is part of our current work.

Finally, a tutoring service that has been implemented in CanadarmTutor is to let learners explore patterns to learn about possible ways of solving problems. Currently, the learners can explore patterns with a very simple interface. However, the learner could be assisted in this exploration by using an interactive dialog with the system that could prompt them on their goals and help them go through the patterns to achieve these goals.

Experiment

We conducted a preliminary experiment in CanadarmTutor with two exercises to qualitatively evaluate the virtual agent's capability to provide assistance. The two exercises each consist of moving a load with the Canadarm2 robotic arm to one of the two cubes (figure 3a). We asked 12 users to record plans for these exercises. The average length of a plan was 20 actions. From this data, CanadarmTutor extracted a partial problem space. In a subsequent work session, we asked users to evaluate the tutoring services provided by this version of CanadarmTutor. All users agreed that the assistance provided was helpful. We also observed that CanadarmTutor correctly inferred the expertise level of all the learners and thus provided hints that were adapted to the user profile. As an example of interaction with a learner, Figure 3b shows a hint message given to a learner upon request during scenario 1. The guiding tutoring service selects the pattern that has the highest support value, matches the last student actions, is marked "successful," and corresponds with the estimated expertise level of the learner. The given hint is to select camera CP4 on Monitor3, decrease the rotation value of the joint WP, and increase the rotation value of joint WE. The values on the right column indicate the values associated to the action. In this context, values 1 and 3 mean to rotate the joints 10º and 30º, respectively (1 unit equals 10º). By default, three steps are showed to the learners in the hint window depicted in figure 3b. However, the learner can click on the "more" button to ask for more steps or click on the "another possibility" button to ask for an alternative.

It should be noted that, although we applied the sequential pattern algorithm only one time after recording the learners plan, it would be possible to make CanadarmTutor apply it periodically to update its knowledge base, while interacting with learners.

*Figure 3*a. The two scenarios; 3b. A hint generated by the virtual agent

Second proposal: A tutoring agent that learns from its own behavior

Our second proposal is to build tutoring agents that can learn from their own behavior by reusing previously successful patterns of tutoring actions. We illustrate this proposal with a virtual agent named CTS (Dubois, Poirier, & Nkambou, 2007) that we have also tested in CanadarmTutor to provide assistance to learners. The following subsections describe CTS, the three operation phases of the new learning mechanism that was integrated in CTS, and two experiments carried in CanadarmTutor to validate (1) the behavior of the new CTS and (2) the behavior of our sequential pattern-mining algorithm with large data sets.

The CTS cognitive agent

The Conscious Tutoring System (CTS) is a generic cognitive agent whose architecture (fig. 4) is inspired by neurobiology and neuropsychology theories of human brain function. It relies on the functional "consciousness" (Franklin & Patterson, 2006) mechanism for much of its operations. It also bears some functional similarities to the physiology of the nervous system. Its modules communicate with one another by contributing information to its working memory (WM) through information codelets. Based on Hofstadter et al's idea (Hofstadter & Mitchell, 1992), a codelet is a very simple agent, "a small piece of code that is specialized for some comparatively simple task." As in Baars's theory (Baars, 1997), these simple processors do much of the processing in the CTS architecture.

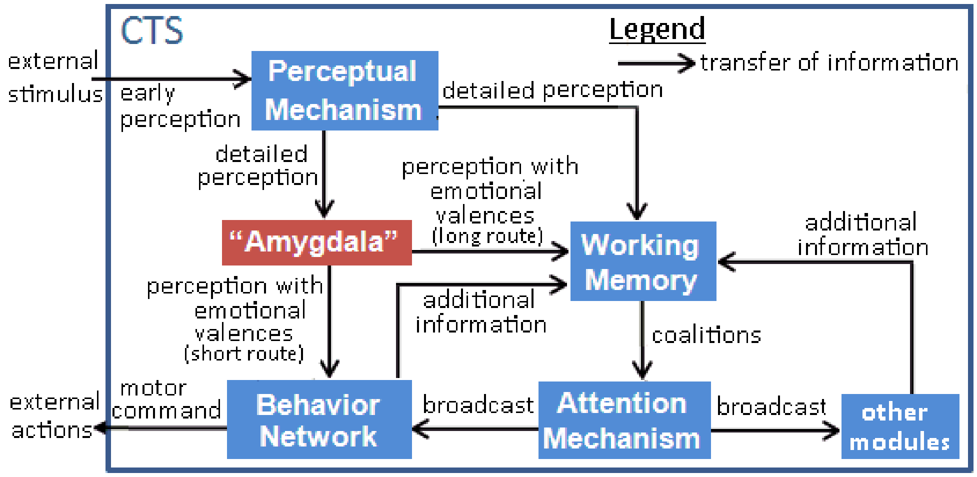

Figure 4. A simplified view of the CTS architecture (see Faghihi et al., 2008 for more details)

CTS possesses two routes for processing external stimuli (cf. fig. 4). Whereas the "long route" is the default route, the "short route" (which will not be described here) allows quick reactions when received information is deemed important by the pseudo-amygdala, the module responsible for emotional reactions in CTS (Faghihi, Poirier, Dubois, Gaha, & Nkambou., 2008). In both cases, the stimuli processing begins with percept codelets (Hofstadter & Mitchell, 1992) that perform collective interpretations of stimuli. The active nodes of the CTS's perception network constitute percepts. In the long route, these percepts enter WM as a single network of codelets, annotated with an activation value. These codelets create or reinforce associations with other already present codelets and create a coalition of information codelets. In parallel, the emotional codelets situated in the CTS's pseudo-amygdala inspect the previously mentioned coalition's informational content, and if is deemed important, infuse it with a level of activation proportional to its emotional valence. During every cognitive cycle, the coalition in the WM that has the highest activation is selected from the WM by the "attention mechanism" and is broadcast to all the modules in the CTS architecture. This selection process ensures that only the most important, urgent, or relevant information is broadcast in the architecture. Following a broadcast, every subsystem (module or team of codelets) that recognizes the information may react to the coalition by appending additional information to it. This process of internal publications (as suggested by Baars, 1997) can continue for many cognitive cycles before an action is executed by CTS. The module responsible for action planning, selection, and execution is the behavior network (BN) (Maes, 1989). When the BN receives a broadcast coalition, it selects the appropriate action to execute. In the current CTS version, we have designed the BN using a graphical authoring tool. We have implemented in CTS, the second proposal that we consider in this article. This learning mechanism is implemented in CTS by the three operation phases, described next.

The observation phase

In the first phase, the observation phase, CTS records a sequence of events (as defined in the second section of this article) for each of its executions. Each event $X = (t_i, A_i)$ represents one cognitive cycle. While the timestamp t_i of an event indicates the cognitive cycle number, the set of items A_i of an event contains (1) an item that represents the coalition of information-codelets that was broadcast during the cognitive cycle and (2) four optional items with numeric values indicating the four emotional valences (high threat, medium fear, low threat, compassion) associated with the broadcast coalition. CTS actually incorporates four emotions inspired by the OCC model of emotions (Ortony, Clore, & Collins, 1988). See Faghihi et al. (2008) for in-depth details about the emotional mechanism of CTS. An example of partial sequence recorded during our experiment was <(1, *c2*), (2, *c4*), (3, *c8* *e2*{-0.4})>. This sequence shows that during cognitive cycle 1 the coalition *c2* was broadcast, followed by the broadcast of *c4* during cognitive cycle 2. Furthermore, it indicates that coalition *c8* was broadcast during the third cognitive cycle and that it generated a negative emotional valence of −0.4 for emotion *e2* (medium fear).

The learning phase

The second operation phase consists of mining frequent patterns from the sequences of events recorded for all executions of CTS by applying our sequential pattern-mining algorithm. This process is executed at the end of each CTS execution, from the moment where five sequences are available (five CTS executions). Currently, we have set up the sequential pattern-mining algorithm to mine only closed sequences with more than three events and with a support higher than 25%. Applying the algorithm results in a set of frequent sequential patterns.

The application phase

The third operation phase consists of improving CTS behavior by making CTS reuse relevant patterns that carry positive emotions. This is done by intervening in the coalition selection phase of CTS. The idea is here to find, during each cognitive cycle, the patterns that are similar to CTS's current execution, then to select as the next coalition to be broadcast the one most probable of generating positive emotions for CTS according to these patterns. Influencing the coalitions that are broadcast will then directly influence the actions that will be taken by the CTS behavior network. This adaption of CTS could be implemented in different ways. We used the SelectCoalition algorithm (figure 4), which takes as parameters (1) the sequence of previous CTS broadcasts (Broadcasts), (2) the set of frequent patterns (Patterns), and (3) the set of coalitions that are candidates to be broadcast during a given cognitive cycle (CandidateCoalitions). This algorithm first sets to zero a variable *min* and a variable *max* for each coalition in CandidateCoalitions. Then, the algorithm repeats the following steps for each pattern *p* of Patterns. First, it computes the strength of *p* by multiplying the sum of the emotional valences associated with the broadcasts in *p* with the support of *p*. Then, it finds all the coalition c $\in$ CandidateCoalitions that appear in *p* after the sequence of the last *k* broadcasts of Broadcasts for any $k \geq 2$. For each such coalition *c*, if the strength of *p* is higher than *c.max*, *c.max* is set to that new value. If that strength is lower than *c.min*, *c.min* is set to that new value. Finally, when the algorithm finishes iterating over the set of patterns, the algorithm returns to CTS's working memory the coalition *c* in CandidateCoalitions having the highest positive value for the sum *c.min* + *c.max* and where *c.max* > 0. This coalition will be the one that will be broadcast next by CTS's attention mechanism. In the case of no coalition meeting these criteria, the algorithm will return a randomly selected coalition from CandidateCoalitions to CTS's working memory.

```
SelectCoalition(Patterns, Broadcasts, CandidateCoalitions)
   FOR each pattern c ∈ CandidateCoalitions
         c.min := 0. c.max := 0.
   FOR each pattern P of Patterns.
         Strength := CalculateSumOfEmotionalValences(P) * Support(P).
         FOR k := 2 to |P|.
               Sa := last k Broadcasts of Broadcasts.
               IF (Sa ⊆ P)
                  FOR each coalition c ∈ CandidateCoalitions appearing
                              after Sa in P
                        c.max := maxOf(Strength, c.max).
                        c.min := minOf(Strength, c.min).
   RETURN c ∈ CandidateCoalitions with the largest positive
               (c.max + c.min) and such that c.max > 0.
```

Figure 5. The SelectCoalition algorithm

The *c.max* > 0 criterion is included to ensure that the selected coalition appears in at least one pattern having a positive sum of emotional valences. Moreover, we have added the *c.min* + *c.max* criterion to make sure that the patterns with a negative sum of emotional valences will decrease the probability of selecting the coalitions that it contains. In our experiments, this criterion has proved to be very important as it can cause CTS to quickly stop selecting a coalition appearing in positive patterns if it becomes part of negative patterns. The reader should note that algorithms relying on other criteria could be used for other applications.

Testing the new CTS in CanadarmTutor

To test CTS's new learning mechanism, users were invited to perform arm manipulations using CanadarmTutor with integrated CTS. These experiments aimed at validating CTS's ability to adapt its behavior to learners. During these experiments, we qualitatively observed that CTS adapted its behavior successfully to learners. Two experiments are described here. The first describes in detail a situation that occurred with User 3 that illustrates well how CTS adapts its behavior thanks to the new learning mechanism. The second experiment describes how our sequential pattern-mining algorithm behaves when the number of recorded sequences increases.

User 3 tended to make frequent mistakes when he was asked to guess the arm's distance from a specific part of the space station. Obviously, this situation caused collision risks between the arm and the space station and was thus a very dangerous situation. This situation was implemented in the CTS's Behavior Network. In this situation, CTS has to make a decision between (1) giving a direct solution such as "You should move joint SP" (Scenario 1) or giving a brief hint such as "This movement is dangerous. Do you know why?" (Scenario 2).

During the interaction with different users, the learning mechanism recorded several sequences of events for that situation, each of them carrying emotional valences. The average length of the stored sequences was 26 events. For example, one partial trace saved when CTS gave a hint (scenario 2) to User 2 was <(13, *c11*), (14, *c14*), (15, *c15*), (16, *c18*), (17, *c19 e4*{0.8})>. In this trace, the positive valence 0.8 for emotion *e4* (compassion) was recorded because the learner answered an evaluation question correctly after receiving the hint. In another partial trace saved by CTS <(16, *c11*), (17, *c14*), (18, *c16*), (19, *c17*), (20, *c20 e2*{−0.4})>, User 2 received a direct solution from CTS (Scenario 1), but failed to answer correctly an evaluation question. This resulted in the valence −0.4 being associated to emotion *e2* (medium fear). After five executions, the learning mechanism extracted ten frequent sequences from the recorded sequences, with a minimum support (*minsup*) higher than 0.25.

Now turning back to User 3, during the coalition selection phase of CTS, the learning mechanism evaluated all mined patterns to detect similar patterns having ended by self-satisfaction. The learning mechanism chose the pattern <(0, *c11*), (1, *c14*), (3, *c18*), (4, *c19 e4*{0.8})> because it contained the most positive emotional valence, had the highest frequency, and the events (0, *c11*), (1, *c14*) matched with the latest events executed by CTS. Therefore, CTS decided that it was better to give a hint (Scenario 2) than to give the answer (Scenario 1) to User 3. This was achieved by broadcasting coalition *c18* (Scenario 2) instead of coalition *c16* (Scenario 1). If the emotional valence had not been as positive as was the case for previous users, CTS might have chosen Scenario 1 rather than Scenario 2. It should be noted that because the set of patterns is regenerated after each CTS execution, some new patterns can be created, while others can disappear, depending on the new sequences of events that are stored by CTS. This ensures that CTS behavior can change over time if some scenarios become less positive or more negative, and more generally that CTS can adapt its behavior to a dynamic environment. In this experiment, the learning mechanism has shown to be beneficial by allowing CTS to adapt its actions to learners by choosing between different scenarios based on its previous experience. This feature is very useful, as it allows the designers to include many alternative behaviors but to let CTS learn by itself which ones are the most successful.

We performed a second experiment with the learning mechanism, but this time to observe how our sequential pattern-mining algorithm behaves when the number of recorded sequences increases. The experiment was done on a 3.6 GHz Pentium 4 computer running Windows XP, and consisted of performing 160 CTS executions for a situation similar to the previous one where CTS had to choose between scenario 1 and scenario 2. In this situation, CTS conducts a dialogue with the student that includes from two to nine messages or questions (an average of six) depending on what the learner answers and the choices CTS makes (similar to choosing between scenarios 1 and 2). During each trial, we randomly answered the questions asked by CTS, and took various measures during CTS's learning phase. Each recorded sequence contained approximately 26 broadcasts.

Figure 6 presents the results of the experiment. The first graph shows the time required for mining frequent patterns after each CTS execution. From this graph, we see that the time for mining frequent patterns was generally short (less than 6 seconds) and increased linearly with the number of recorded sequences. In our context, this performance is very satisfying. However, the performance of the data-algorithm could still be improved as we have not yet fully optimized all of its processes and data structures. In particular, in future works we will consider modifying the algorithm to perform incremental mining of sequential patterns as some other sequential pattern-mining algorithms do. This would improve performance, as it would not be necessary to recompute from scratch the set of patterns for each new added sequence.

The second graph shows the average size of patterns found for each execution. The size ranges from 9 to 16 broadcasts. The third graph depicts the number of patterns found. It remains low and stabilized at around 8.5 patterns during the last executions. The reason why the number of patterns is small is that we mined only closed patterns (see definition in the third section). If we had not mined only closed patterns, all the sub-sequences of each pattern would have been included in the results. Mining closed patterns is also much faster because, during the search for patterns, large parts of the search space that are guaranteed not to lead to close patterns are pruned. For example, for mining non-closed patterns from the first four sequences only, it took more than one hour (we stopped the algorithm after one hour), while mining closed patterns took only 0.558 seconds. The reason for this is that the four sequences share more than 15 common broadcasts. Therefore, if the pruning of the search space is not done, the algorithm has to consider all combinations of these broadcasts, which is computationally very expensive. This demonstrates that it is beneficial to mine closed patterns. Finally, the average time for executing the SelectCoalition algorithm at each execution was always less than 5 milliseconds. Thus, the costliest operation of the learning mechanism is the learning phase.

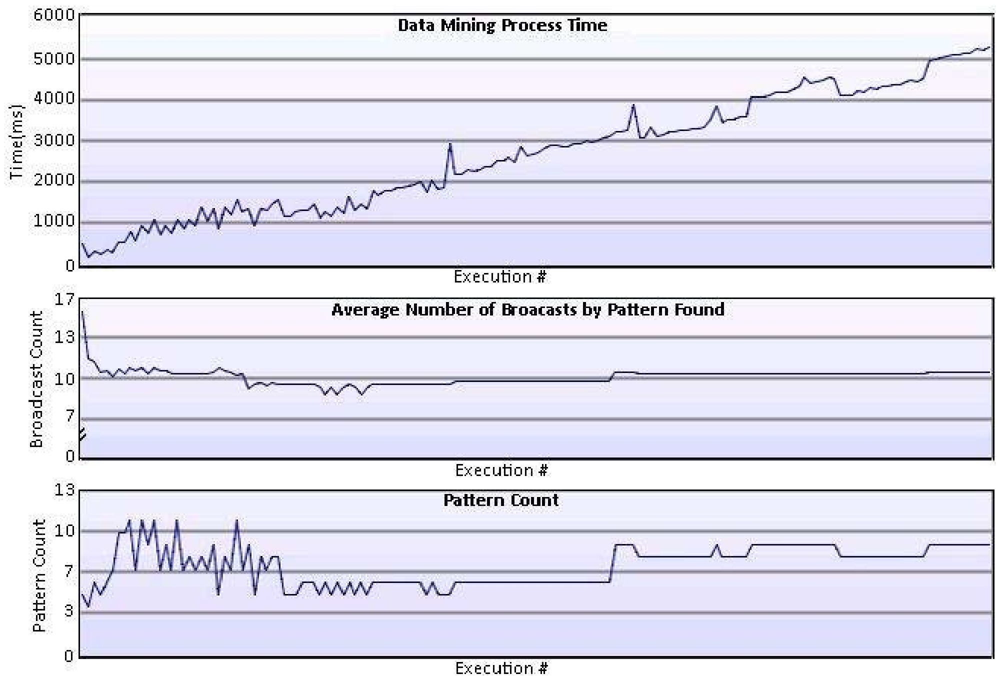

Figure 6. Results from the second experiment

Conclusion

In this article, we presented the idea of exploiting temporal data found in intelligent tutoring system logs to improve their capability to provide relevant and adaptive assistance. To demonstrate this idea, we described two proposals. While the first one is designed to learn task models from recorded novice and expert solutions to provide assistance to learners in problem-solving activities, the second one is aimed at building tutoring agents that can adapt their behavior to learners and situations by reusing previously successful patterns of tutoring actions. The two proposals should be reusable in other tutoring agents and domains, as the format for encoding behaviors is fairly generic.
In future work, we will perform further experiments to measure empirically how the different versions of CanadarmTutor influence the learning of students. We will investigate different ways of improving the performance of our sequential pattern-mining algorithm, including modifying it to perform an incremental mining of sequential patterns. We also plan to integrate the two proposals into other tutoring systems.

Acknowledgements

The authors thank the Canadian Space Agency, Fonds Québécois de la Recherche sur la Nature et les Technologies, and the Natural Sciences and Engineering Research Council for their logistic and financial support. The authors also thank members of GDAC/PLANIART teams who have participated in the development of CanadarmTutor.

References

Agrawal, R., & Srikant, R. (1995). Mining Sequential Patterns. *Proceedings of the International Conference on Data Engineering* (pp. 3–14), Los Alamitos, CA: IEEE Computer Society Press.

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. (2006). The Cognitive Tutor Authoring Tools (CTAT): Preliminary evaluation of efficiency gains. *Proceedings the 8th International Conference on Intelligent Tutoring Systems* (pp. 61–70), Berlin: Springer.

Baars, B. J. (1997). *In the theater of consciousness.* New York: Oxford University Press.

Baker, R., Barnes, T., & Beck, J. E. (2008). *Proceedings of Educational Data Mining 2008: 1st International Conference on Educational Data Mining*. Montreal, Quebec, Canada. June 20-21, 2008.

Barnes, T. & Stamper, J. (2008). Toward automatic hint generation for logic proof tutoring using historical student data, *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (pp. 373–382), Berlin: Springer.

Dubois, D., Poirier, P., & Nkambou, R. (2007). What does consciousness bring to CTS? *Proceedings of the 2007 AAAI Fall Symposium* (pp. 55–60), Menlo Park, CA: AAAI Press.

Faghihi, U., Poirier, P., Dubois, D., Gaha, M., & Nkambou, R. (2008). How emotional mechanism learn and helps other types of learning in a cognitive agent, *Proceedings of the 2009 IEEE/WIC/ACM Conference on Intelligent Agent Technology*, Los Alamitos, CA: IEEE Computer Society Press.

Fournier-Viger, P., Nkambou, R., & Mephu Nguifo, E. (2008a). A knowledge discovery framework for learning task models from user interactions in intelligent tutoring systems. *Proceedings of the 7th Mexican Conference on Artificial Intelligence* (pp. 765–778), Berlin: Springer.

Fournier-Viger P., Nkambou, R., & Mayers, A. (2008b). Evaluating spatial representations and skills in a simulator-based tutoring system. *IEEE Transactions on Learning Technologies, 1*(1), 63–74.

Franklin, S., & Patterson, F.G.J. (2006). The LIDA architecture: Adding new modes of learning to an intelligent, autonomous, software agent. *Proceedings of 9th World Conference on Integrated Design & Process Technology*. San Diego: Society for Design and Process Science.

Han, J., & Kamber, M. (2006). *Data mining: Concepts and technique*s (2nd ed.). San Francisco: Morgan Kaufmann.

Hofstadter, D. R., & Mitchell, M. (1992). The copycat project: A model of mental fluidity and analogy-making. In K. Holyoak, J. & Barnden, J. A. (Eds.) *Advances in connectionist and neural computation theory 2* (pp. 31–113). Norwood, NJ: Ablex.

Kabanza, F., Nkambou, R., & Belghith, K. (2005). Path-planning for autonomous training on robot manipulators in space. *Proceedings of the 19th International Joint Conference on Artificial Intelligence* (pp. 1729–173), Denver, CO: Professional Book Center.

Lynch, C., Ashley, K., Aleven, V. & Pinkwart, N. (2006). Defining ill-defined domains: A literature survey. *Proceedings of the Intelligent Tutoring Systems for Ill-Defined Domains Workshop at the 8th International Conference on Intelligent Tutoring Systems* (pp. 1–10), Jhongli, Taiwan, June 27, 2006.

Maes, P. (1989). How to do the right thing. *Connection Science, (1),* 291–323.

Matsuda, N., Cohen, W. Sewall, J., Lacerda, G., & Koedinger, K. (2007). Predicting students' performance with SimStudent: Learning cognitive skills from observation. *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 467–478), Amsterdam: IOS Press.

Mitrovic, A., Mayo, M., Suraweera, P., & Martin, B. (2001). Constraint-based tutors: A success story. *Proceedings of the 14th Industrial & Engineering Application of Artificial Intelligence & Expert Systems* (pp. 931–940), Berlin: Springer.

Ortony, A., Clore, G., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge University Press.

Pinto, H., Han, J., Pei, J., Wang, K., Chen, Q., Dayal, U. (2001). Multi-Dimensional Sequential Pattern Mining, *Proceedings of the International Conference of Information and Knowledge Management* (pp. 81–88), New York: ACM.

Riccuci, S., Carbonaro, A., & Casadei, G. (2007). Knowledge acquisition in intelligent tutoring system: A data mining approach. *Proceedings of the 6th Mexican Conference on Artificial Intelligence* (pp. 1195–1205), Berlin: Springer.

Simon, H. A. (1978). Information-processing theory of human problem solving. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 5. Human information*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Songram, P., Boonjing, V., Intakosum, S. (2006). Closed multidimensional sequential pattern mining. *Proceedings of the 3rd International Conference on Information Technology: New Generations* (pp. 512–517), Los Alamitos, CA: IEEE Computer Society Press.

Wang, J., Han, J., Li, C. (2007), Frequent closed sequence mining without candidate maintenance, *IEEE Transactions on Knowledge and Data Engineering, 19*(8), 1042–1056.

Chi, M., & VanLehn, K. (2010). Meta-Cognitive Strategy Instruction in Intelligent Tutoring Systems: How, When, and Why. *Educational Technology & Society, 13* (1), 25–39.

Meta-Cognitive Strategy Instruction in Intelligent Tutoring Systems: How, When, and Why

Min Chi and Kurt VanLehn*

Learning Research and Development Center & Intelligent System Program, University of Pittsburgh, PA, USA // mic31@cs.pitt.edu

*Department of Computer Science & Engineering, Arizona State University, AZ, USA // Kurt.Vanlehn@asu.edu

ABSTRACT

Certain learners are less sensitive to learning environments and can always learn, while others are more sensitive to variations in learning environments and may fail to learn (Cronbach & Snow, 1977). We refer to the former as high learners and the latter as low learners. One important goal of any learning environment is to bring students up to the same level of mastery. We showed that an intelligent tutoring system (ITS) teaching a domain-independent problem-solving strategy indeed closed the gap between high and low learners, not only in the domain where it was taught (probability) but also in a second domain where it was not taught (physics). The strategy includes two main components: one is solving problems via backward chaining (BC) from goals to givens, called the BC strategy, and the other is drawing students' attention to the characteristics of each individual domain principle, called the principle-emphasis skill. Evidence suggests that the low learners transferred the principle-emphasis skill to physics while the high learners seemingly already had such skill and thus mainly transferred the other skill, the BC strategy. Surprisingly, the low learners learned just as effectively as the high learners in physics. We concluded that the effective element of transfer seemed not to be the BC strategy, but the principle-emphasis skill.

Keywords

Intelligent tutoring systems, Meta-cognitive skills, Domain-independent problem-solving strategies

Introduction

Certain learners are less sensitive to learning environments and can always learn; while others are more sensitive to variations in learning environments and may fail to learn (Cronbach & Snow, 1977). We refer to the former as high learners and the latter as low learners. Bloom (1984) argued that human tutors not only raised the mean of test scores, but also decrease the standard deviation of scores. That is, students generally start with a wide distribution in test scores but as they are tutored, the distribution becomes narrower: the students at the low end of the distribution begin to catch up with those at the high end. Another way to measure the same phenomenon is to split students into high and low groups based on their incoming competence then measure the learning gains of both groups. According to Bloom, a good tutor should exhibit an aptitude-treatment interaction: both groups should learn, and yet the learning gains of the low students should be so much greater than those of the high ones that their performance in the post-test ties with that of the high ones. That is, one benefit of tutoring is to narrow or even eliminate the gap between high and low. In order to fully honor the promises of learning environments, an effective system should narrow the gap as much as possible without pulling the high learners down. Many preexisting systems can decrease such differences but not eliminate them. This is due in part to the fact that we do not fully understand why such differences exist.

One of many hypotheses is that low learners lack certain specific skills about how to think, including general problem-solving strategies and meta-cognitive skills. If this hypothesis is true, we expect that teaching students an effective problem-solving strategy would not only improve students' learning gains but also decrease the gap between the low and the high learners. Furthermore, if such problem-solving strategy is domain independent, we expect that learners would learn how to apply the strategy and seek to transfer it to new learning environments. Past research has indicated that these skills can be transferred across domains (Lehman, Lempert, & Nisbett, 1988; Lehman & Nisbett, 1990). However, few studies have investigated transfer of problem-solving strategy across domains.

In this paper, we investigate these questions in a special class of learning environments, intelligent tutoring systems (ITSs) (VanLehn, 2006). We present a study in which two groups of college students studied probability first and then physics. The experimental group studied probability with Pyrenees, an ITS that explicitly taught and required students to employ a general problem-solving strategy (VanLehn et al. 2004); while the control group studied probability with Andes, an ITS that did not teach or require any particular strategy (VanLehn et al. 2005). During subsequent physics instruction, both groups used Andes, which also did not teach or require students to employ any particular strategy.

As reported earlier (Chi & VanLehn, 2007), we found that the experimental group out-performed the control group not only in probability, where the strategy was taught and forced upon the participants, but also in physics where it was not forced upon the participants. Furthermore, the strategy seemed to have lived up to our expectations and transferred from probability to physics. In this paper, we determine whether explicit strategy instruction exhibits an aptitude-treatment interaction, that is, whether it narrows or even eliminates the gap between high and low and, moreover, whether both high and low indeed transfer the strategy to the second domain.

Background

A task domain is deductive if solving a problem requires producing an argument, proof, or derivation consisting of one or more inference steps, and each step is the result of applying a domain principle, operator, or rule. For instance, solving algebraic equations is a deductive domain, and in particular, $2x + 5 = 21$ can be done via two steps: 1) subtract the same term 5 from both sides of the equation; and 2) divide both sides by the non-zero term 2. Proving a geometry theorem is deductive, as is solving quantitative physics problems. Deductive task domains are common parts of mathematical and scientific courses such as probability and physics. Two common problem-solving strategies in deductive domains are forward chaining (FC) and backward chaining (BC) (Russell & Norvig, 2003).

In FC, the solver starts with the set of given propositions, applies a principle to some subset of them (which produces a new proposition), adds the new proposition to the known propositions, and repeats this process until the problem's goal is met or no new proposition can be produced. BC is goal-directed in that the goal is progressively broken down into sub-goals and sub-sub-goals, etc. This constructs a partial plan in the form of a goal stack, which the solver uses to guide its forward application of principles. This is easier to explain with the aid of an example, so we will combine that explanation with an introduction to one of the task domains, probability. A portion of the probability task domain is described in first row in Table 1.

Table 1. A subset of the probability principle and an example that can be solved by these rules

Rules:	R1: For any event E, $P(E) + P(\sim E) = 1$ R2: If events A and B are independent, $P(A \cap B) = P(A)*P(B)$. R3: If events A and B are independent, then A and ~B, ~A and B, and ~A and ~B are all independent events.
Problem:	Events A and B are independent, and $P(\sim A) = 0.9$, $P(\sim B) = 0.8$. Compute $P(A \cap B)$.

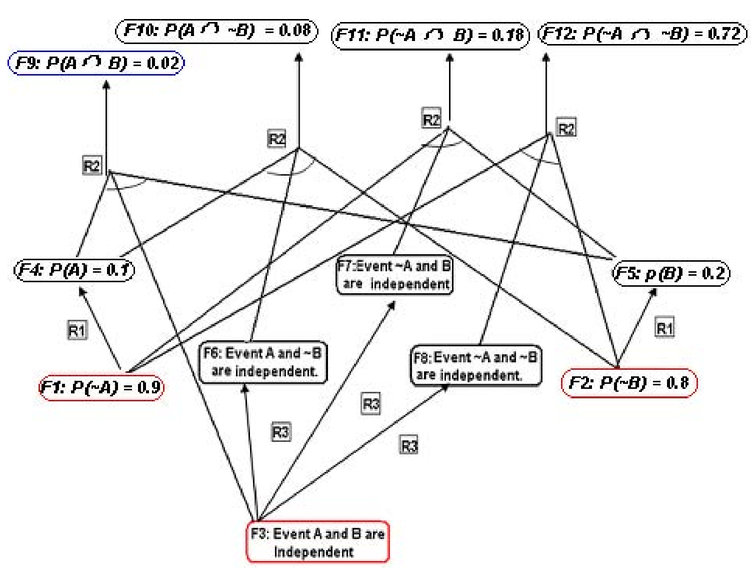

Figure 1. A solution graph using forward chaining

An example of solving a problem using forward chaining is shown in Figure 1 and the problem is listed in the last row of Table 1. One reads this graph starting at the bottom and working upward. The lowest propositions are given. Rules are applied to produce the propositions above. Although forward chaining stops once the goal of the problem, F9, is found, a few more propositions (F10, F11, and F12) are also shown in the graph because it is possible that they may be produced before F9.

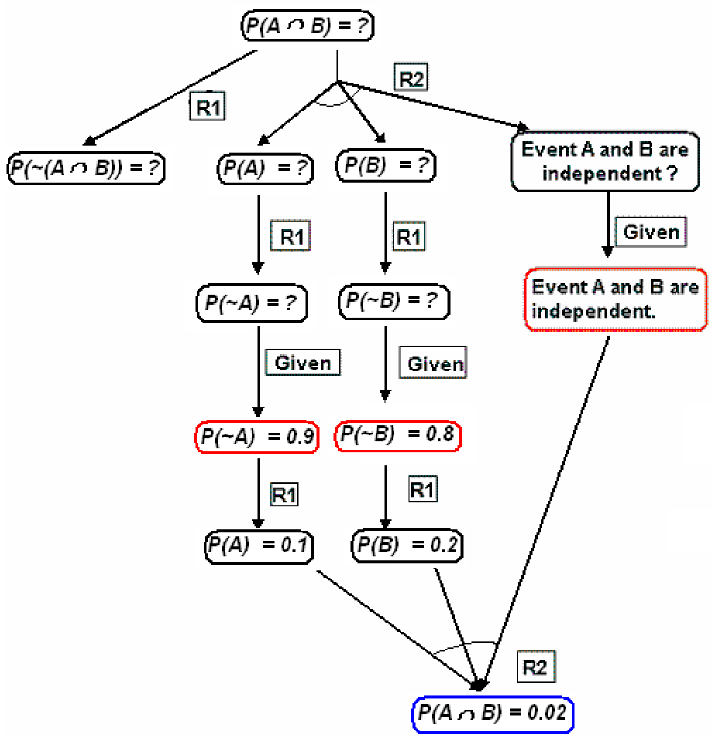

Figure 2. A solution graph using backward chaining

Figure 2 presents solving the same problem via BC. It is read from the top down. The problem's goal is decomposed by R1 into a sub-goal $P(\sim(A \cap B))$, but doesn't look promising because no other rules can be fired further. Thus, BC tries decomposing via R2. That yields three sub-goals, which look promising, so it continues decomposing each of them until they all eventuate in sub-goals that match given propositions. It then applies the rules in the forward direction, guided by the goal and sub-goal links that it has saved, doing computations as it goes and eventually calculating the problem's answer.

FC is complete (Russell & Norvig, 2003) but can be inefficient because its inference process is not directed toward solving the problem's goal. In Figure 1, FC did a lot of irrelevant work. Of the nine facts inferred, only three were needed for solving the problem. BC, on the other hand, is focused on achieving the problem's goals but can meet dead ends. For instance, $P(\sim(A \cap B))$ in Figure 2, is one such dead end because it triggers no more rules in Table 1.

Studies comparing strategy instruction with no-strategy instruction

Although FC and BC are widely used in computer science, they are seldom observed in a pure form in natural human problem solving. In Newell and Simon's (1972) seminal study of logic problem solving, *none* of the subjects used FC or BC in a pure form. Early studies of expert and novice physics problem solvers suggested that novices used BC and experts used FC (Larkin, McDermott, Simon, & Simon, 1980; Simon & Simon, 1978), but later studies showed that both used a mixture and, in fact, used fairly similar mixtures (Priest & Lindsay, 1992). Eventually, work in this area diminished, perhaps because it appeared that most human solvers used a mixture of strategies, analogies, heuristics, and many other kinds of knowledge during their problem solving.

Although neither experts nor novices seem to use FC and BC in their pure form, the strategies' success in guiding computer problem solvers suggests that teaching students to use pure FC or BC might improve their problem solving. Several studies were conducted to test this hypothesis. Next, we will give a brief review.

Sweller and his colleagues conducted a series of studies comparing goal-free problem solving to ordinary problem solving (Owen & Sweller, 1985; Sweller, 1989; Tarmizi & Sweller, 1988). In the goal-free problem solving condition, students were not told a specific goal of a problem. Instead, they were asked to derive everything they could from the given facts. Students could only use FC on these goal-free problems, as BC requires a goal to start from. In the ordinary problem-solving condition, students were given problems with a specific goal, so they could use FC, BC, or a mixture. In all of these studies, the goal-free group learned more than the ordinary problem-solving group, thus suggesting that "teaching" a single problem-solving strategy in the form of pure FC improves learning. However, the number of inferences made to solve a goal-free problem is generally much larger than the number of inferences made to solve a goal-specific problem. Thus, the goal-free students could have benefited simply from having more practice in applying the domain principles. Indeed, when the experimenters modified the study so that students in both conditions applied the same number of domain principles, the difference between conditions disappeared (Owen & Sweller, 1985; Sweller, 1988). Thus, although these studies are consistent with the hypothesized benefits of explicit strategy instruction, there are other explanations for the results as well.

Trafton and Reiser (1991) tested the benefits of explicit strategy instruction in a sub-domain of computer programming, wherein students had to compose primitive functions in order to produce composite goal function. Three forms of instruction were compared based on the way in which the goal function could be assembled: forward-only, backward-only or freely. After completing 13 training problems in less than an hour, all three groups achieved the same learning gains. Although it is always hard to interpret a null result, it could be that the task domain was too simple to allow an explicit instruction on problem-solving strategies to demonstrate benefits.

Scheines and Sieg (1994) gave students over 100 training problems in sentential logic over a five-week period. One group of students was taught to use FC; a second group was taught to use BC; while a third group, the unconstrained group, was not taught any strategy and operated freely. After five weeks of instruction, no significant differences were found among the three groups on the mid-term exam (post-test). When the FC and BC groups were aggregated as a one-way strategy condition, there were still no significant differences between them and the unconstrained group on post-test scores. However, contrary to our hypothesis, the unconstrained students gained more than the one-way strategy students on difficult problems, where one would expect an explicit search strategy to be especially helpful. The experiment suggested that constraining students to use just one strategy may actually harm their performance.

VanLehn et al. (2004) compared an explicitly taught version of backward chaining to unconstrained problem-solving. Students who had not taken college physics were taught elementary mechanics over several multiple-hour training sessions. On some post-test measures, the students who were explicitly taught a strategy scored higher than those who were not taught a strategy and could solve problems in any order. However, on other measures, the two groups did not differ. Overall, performance on the post-test was quite poor, suggesting a floor effect — the post-test was too difficult for both groups.

In summary, most studies above were conducted in a single domain and contrasted students who were taught a strategy and those who were not. In this paper, we investigate the impact of explicit strategy instruction on eliminating the gap between high and low across two unrelated domains and two different ITSs. The problem-solving strategy chosen is the target variable strategy (TVS), a domain-independent BC strategy (VanLehn et al., 2004), and the two selected domains were probability and physics. During probability instruction, students in the experimental group were trained on an ITS, Pyrenees, that explicitly taught the TVS; while students in the control group were trained on another ITS, Andes, without explicit strategy instruction. During subsequent physics instruction, both groups were trained on the same ITS, which did not teach any strategy. On both probability and physics post-tests, we expect the following: high-experimental = low-experimental = high-control > low-control. That is, for both task domains, the low students should catch up with the high students, but only if they were taught the TVS.

Methods

Participants

Participants were 44 college students who received payment for their participation. They were required to have a basic understanding of high-school algebra, but not to have taken college-level statistics or physics courses. Students were randomly assigned to the two conditions. Two students were eliminated: one for a perfect score on the probability pre-test and one for deliberately wasting time.

Target variable strategy (TVS)

The TVS consists of three phases: 1) translating the problem statement, 2) applying principles and generating equations, and 3) solving equations. The central part of TVS happens in the second phase: applying principles and generating equations, in which students follow:

(1) Choose one of the *sought* (unknown) variables as the target variable
(2) Select a principle application that will generate an equation containing the target variable
(3) Define new variables as necessary to ensure that every quantity in the equation has a variable
(4) Write the equation in terms of the defined variables
(5) Mark the target variable *known*
(6) Mark the unknown variables in the equation *sought*

This procedure is repeated until there is no variable marked sought anymore, and then students go to the final phrase, solving equations. More details on TVS are described in (VanLehn et al., 2004). To illustrate, we will compare two example solutions for a probability problem. Table *2* contains a TVS solution by following the TVS while Table 3 contains a non-TVS solution. Note that $P(A)$, $P(A \cap B)$, $P(\sim A \cup \sim B)$, etc. are algebraic variables even though their names make them look like functions.

Table 2. Solving a problem by following the TVS

Problem: **Given $P(A) = 1/3$, $P(B) = 1/4$, $P(A \cap B) = 1/6$, find the probability: $P(\sim A \cup \sim B)$.**		
Step	Proposition	Justification
Phase 1: Translating the problem statement		
1	$P(A) = 1/3$	Given
2	$P(B) = 1/4$	Given
3	$P(A \cap B) = 1/6$	Given
4	$P(\sim A \cup \sim B)$	Sought
Phase 2: Applying principles and generating equations		
5	$P(\sim A \cup \sim B) = P(\sim(A \cap B))$	To find $P(\sim A \cup \sim B)$, apply De Morgan's theorem. Delete *sought* from $P(\sim A \cup \sim B)$ and mark $P(\sim(A \cap B))$ as sought
6	$P(A \cap B) + P(\sim(A \cap B)) = 1$	To find $P(\sim(A \cap B))$, apply the complement theorem. Delete *sought* from $P(\sim(A \cap B))$
Phase 3: Solving equations.		
7	$P(\sim(A \cap B)) = 5/6$	Solve 6.
8	$P(\sim A \cup \sim B) = 5/6$	Solve 5

Table 2 shows that in the first phase the student defines four variables, gives values to three of them, and marks $P(\sim A \cup \sim B)$ the sought variable. Then the student moves to the second phase. During each cycle of the second phase, the student must make two decisions. One is which sought variables to select as the target variable if there is more than one variable marked sought. Since all sought variables must eventually be selected as target variables, the order in which they are chosen does not affect the solvability of the problem. The other decision is which principle application to use if there happen to be several that contain the target variable. If the student makes an unlucky selection, the problem will be unsolvable. If so, the student must back up and make a different selection. In this example, during the first cycle, only $P(\sim A \cup \sim B)$ was marked as sought, so it is selected as the target variable. Then the student chooses De Morgan's theorem as the principle to solve for $P(\sim A \cup \sim B)$. To do so, the student defines a variable $P(\sim(A \cap B))$ and then enters the equation $P(\sim A \cup \sim B) = P(\sim(A \cap B))$. Then the student removes the sought mark from $P(\sim A \cup \sim B)$, and marks the other variable in the equation, $P(\sim(A \cap B))$, as sought. This ends the first cycle. On the next cycle, again, only one variable is marked as sought, $P(\sim(A \cap B))$, so the student selects it as the target variable, applies the complement theorem to it, and removes the sought mark from it. At this point, no variables are marked sought, so the second phase ends. During the third and final phase, the student solves the equations in reverse chronological order.

Table **3.** ***A non-TVS solution of the same problem***

Step	Proposition	Justification
1	$P(A \cup B) = P(A)+P(B)- P(A\cap B)$	Addition theorem for two events: A and B
2	$P(\sim B)+p(B) = 1$	Complement Theorem
3	$P(A \cap B)+ P(\sim A \cup \sim B) = 1$	De Morgan's Law and Complement Theorem
4	$P(A \cap B) = 1/6$	Given
5	$P(\sim(\sim A \cup \sim B)) = 1/6$	Solve 3
6	$P(\sim A \cup \sim B) = 5/6$	Solve 4

Table 3 presents a solution to the same problem derived without the TVS. The solution is neither FC nor BC. It includes an equation (Step 2) that is not necessary for solving the problem and another equation (Step 3: $P(A \cap B) + P(\sim A \cup \sim B) = 1$) that is a combination of two principle applications: $P(\sim(A \cap B)) = P(\sim A \cup \sim B)$ and $P(A \cap B) + P(\sim(A \cap B)) =1$. This is typical of the solutions by students who were not taught the TVS.

Students can use the TVS to solve problems in probability, physics, and many other tasks domains (VanLehn et al., 2004). In general, the TVS applies in task domains where solving a problem consists of generating a solvable set of equations. The TVS can be proved complete in that it will generate a set of equations that solves the problem if such a set exists. So far, we have described only the tedious procedural aspects of the TVS. There is only one key aspect of the TVS that has not yet been described, and we will describe it next.

A key detail and its implications for learning

In the second phase of the TVS, applying principles and generating equations, students select a principle application whose equation will contain the target variable. To do this, the tutoring system has students first pick a principle by name from a list of the domain principles that have been taught. It then has students specify how to apply that principle, again by making menu selections. For example, in step 6 of Table 2, after students pick $P(\sim(A \cap B))$ as the target variable and select the complement theorem to apply, but before they input the correct equation, $P(A \cap B) + P(\sim(A \cap B)) = 1$, the tutoring system would say:

> *You have chosen the complement theorem to apply for the target variable. To apply the principle, you must have noticed that there is a set of events that are mutually exclusive and collectively exhaustive. What are these events?*

Students should input the two events, $\sim(A \cap B)$ and $(A \cap B)$. In AI terms, the students are being asked to supply values for the arguments of the principle, thus establishing which of many possible instances of the principle should be applied. Only when an instantiation has been selected is the equation generated by the principle application completely determined and the tutoring system able to ask the student to enter it.

Therefore, the TVS is not simply a BC strategy, like the one used by Bhaskar and Simon (1977), which has students simply enter an equation. It instead has them specify both a principle and its arguments. Even if students know what equation they want, they have to figure out which principle and which arguments will generate it. Thus, they have to learn the principles deeply instead of simply learn a syntactic version of the available equations. As for the example above, students following the TVS are more likely to learn that the complement theorem should only apply in the events that are mutually exclusive and collectively exhaustive instead of in normal events. The TVS taught the students to focus their attention on acquiring a deep understanding of the principles, as that was all they needed in order to run the TVS. To summarize, the TVS includes two main components: one is to solve problems via BC from goals to givens, called the BC-strategy, and the other is to focus attention to the domain principles, called the principle-emphasis skill.

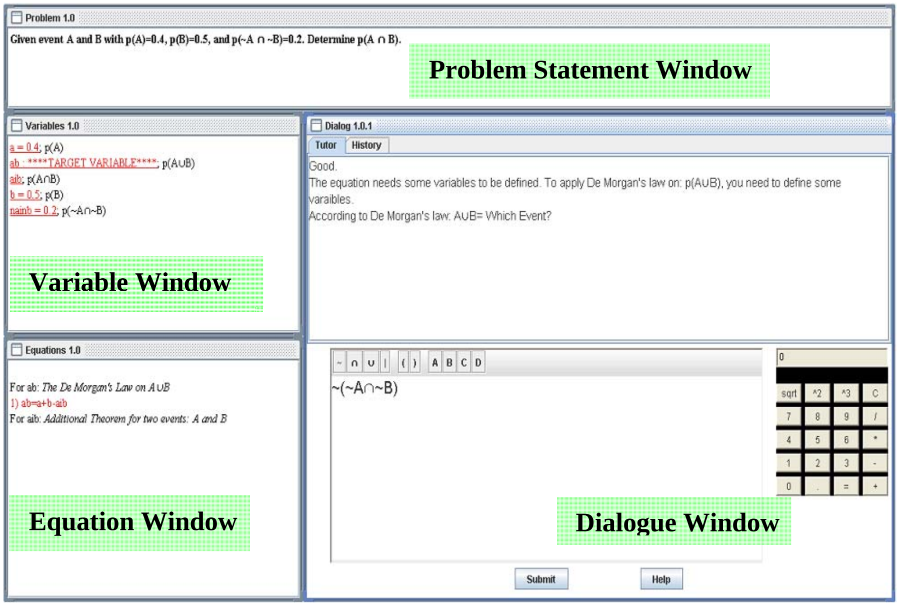

Figure 3. Pyrenees's interface

Three ITSs

The three ITSs involved in this study were Pyrenees, Andes probability, and Andes physics. Their corresponding screen shots are shown in Figure 3, *Figure 4*, and *Figure 5*, respectively. The first two taught probability, whereas the third taught physics. Apart from their domain knowledge, Andes probability and Andes physics were identical, so we use "Andes" to refer to both. Pyrenees required students to follow the TVS, while Andes did not require students to follow any problem-solving strategy. In this study, students in the experimental group learned probability in Pyrenees, and then learned physics in Andes; while student in the Control group learned both probability and physics in Andes. Next, we will compare Pyrenees and Andes from the perspectives of both the user interface and students' behaviors.

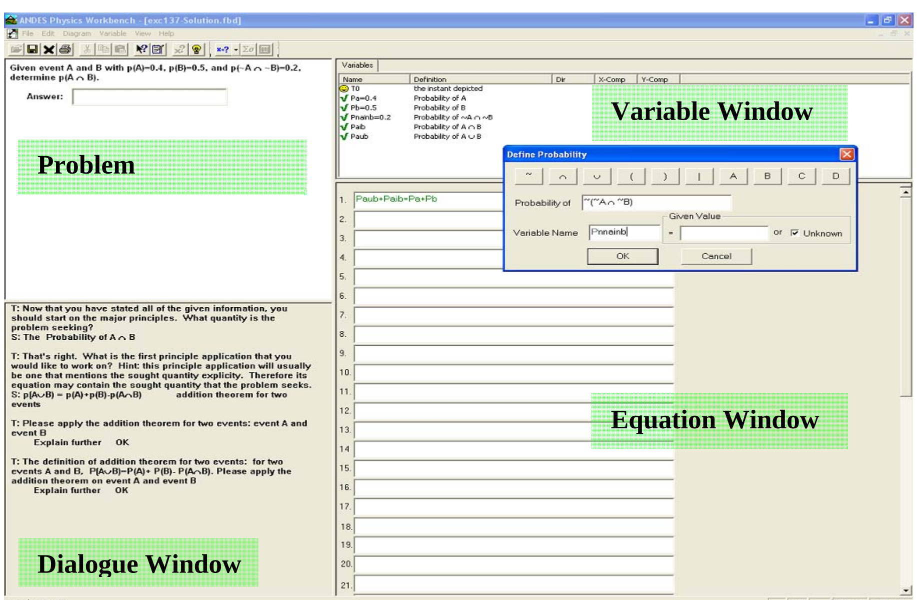

Figure 4. Andes probability's interface

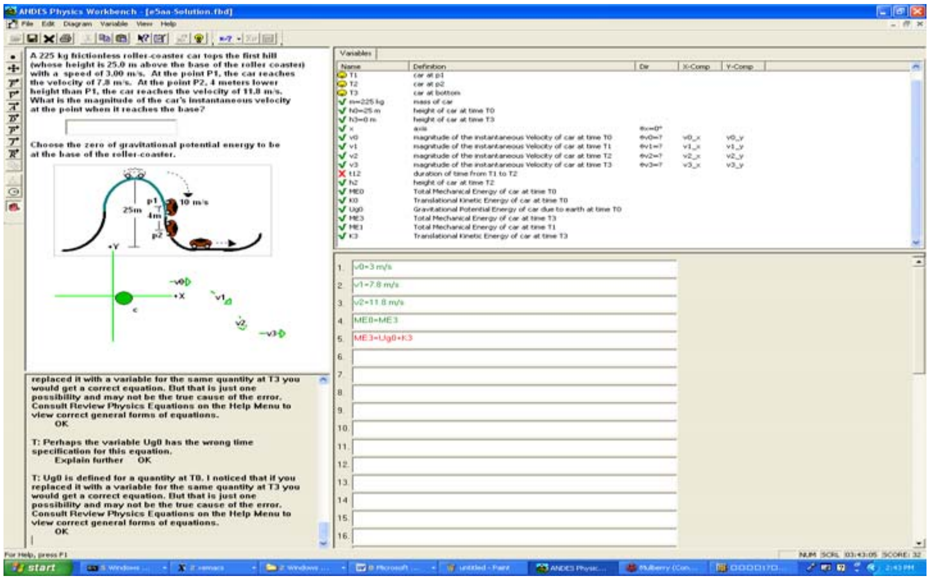

Figure 5. Andes physics' interface

User Interfaces Perspectives

Both Pyrenees and Andes provide a multi-paned screen that consists of a problem-statement window, a variable window for listing defined variables, an equation window, and a dialog window (see Figures 3–5). However, the computer-student interactions were quite different for each system. Pyrenees guided students in applying the TVS by prompting them to take steps dictated by the TVS. For example, when the TVS determined that it was time to define a variable, Pyrenees popped up a tool for that purpose (see Figure 3). Thus the interaction with Pyrenees was a turn-taking dialogue, where the tutor's turns always ended with a question to which the student must reply. All interaction with Pyrenees took place in the dialogue window. In Andes, on the other hand, students used GUI tools to construct and manipulate a solution. Thus the interaction with Andes was open-ended and event-driven. Students could edit or interact with any of the four windows by drawing vectors in the top left window, writing or editing equations in any row of the equation window, and so on. Once an entry or edit was made successfully, Andes provided no further prompting for the next step. If students didn't know what to do next, they could ask for a hint by clicking on the next-step help button.

Interactive behaviors perspectives

Both Andes and Pyrenees provide immediate feedback. However, their standard of correctness differs. Andes considers an entry correct if it is true, regardless of whether it is useful for solving the problem. On Pyrenees, however, an entry is considered correct if it is true and strategically acceptable to the TVS. Moreover, students can enter an equation that is the algebraic combination of several principle applications on Andes but not on Pyrenees because the TVS requires students to apply one principle at a time.

Both systems provide hints when students ask. When an entry is incorrect, students can either fix it independently or ask for what's-wrong help. When they do not know what to do next, they can ask for next-step help. Both next-step help and what's-wrong help are provided via a sequence of hints that gradually increase in specificity. The last hint in the sequence, called the bottom-out hint, tells the student exactly what to do. Pyrenees and Andes give the same what's-wrong help for any given entry, but their next-step help differs. Because Pyrenees requires students to follow the TVS, it knows what step they should be doing next so it gives specific hints. In Andes, however, students can always enter any correct step, so Andes does not attempt to determine their problem-solving plans. Instead, it asks students what principle they are working on. If students indicate a principle that is part of a solution to the problem, Andes provides as a hint an uncompleted step from the principle application. If no acceptable principle is chosen, Andes picks an unapplied principle from the solution that they are most likely to be working on.

Two domains

Two deductive domains, probability and physics, were involved in this study as the initial and transfer domain, respectively. Each domain contained ten major principles. Probability included the complement theorem, Bayes rule, and so on; while physics included the definition of kinetic energy, conservation of total mechanical energy, and so on.

Procedure

The procedure in this study had four main parts: background survey, probability instruction, Andes interface training, and physics instruction (shown in the left column of Table 4). All materials were online. The background survey asked for high school GPA, SAT scores, experience with algebra, and other information.

Table 4**.** Experiment procedure

<table>
<tr><th>Part</th><th>Experimental</th><th>Control</th></tr>
<tr><td>Survey</td><td colspan="2">Background survey</td></tr>
<tr><td rowspan="4">Probability instruction</td><td colspan="2">Pre-training</td></tr>
<tr><td colspan="2">Pre-test</td></tr>
<tr><td>Training on Pyrenees</td><td>Training on Andes probability</td></tr>
<tr><td colspan="2">Post-test</td></tr>
<tr><td>Andes interface training</td><td>Solve a probability problem on Andes probability</td><td></td></tr>
<tr><td rowspan="4">Physics instruction</td><td colspan="2">Pre-training</td></tr>
<tr><td colspan="2">Pre-test</td></tr>
<tr><td colspan="2">Training on Andes physics</td></tr>
<tr><td colspan="2">Post-test</td></tr>
</table>

The probability and physics instruction each consisted of the same four phases: 1) pre-training, 2) pre-test, 3) training on the tutoring system, and 4) post-test. We describe each phase in turn, pointing out relevant differences, if any, between the two task domains.

Pre-training

During pre-training all students studied the domain principles. For each principle, they read a general description, reviewed some examples, and solved a series of single-principle and multi-principle problems. After solving a problem, the answer was marked correct or incorrect, and the correct solution was displayed. If the answer was incorrect, the students were asked to solve another problem isomorphic to the one that they had just failed to solve; this repeated until they either succeeded in solving a problem or failed three times. On multiple-principle problems, students had only one chance to solve the problem and were not asked to solve an isomorphic problem if their answer was incorrect.

Pre-tests

During the pre-tests, after an answer was submitted, students automatically proceeded to the next question without any feedback on the correctness of the answer. Students were not allowed to go back to earlier questions. This was the procedure for the post-tests as well. All students took the same pre- and post-tests. All test problems were open-ended and required students to derive an answer by writing and solving one or more equations.

Training on ITSs

In Phase 3, students first watched a video that demonstrated problem solving in the corresponding ITS. During probability instruction, the strategy students also read a text description of the TVS. Then, all students solved the same twelve probability problems or eight physics problems in the same order. More specifically, students in the experimental group solved all twelve probability problems in Pyrenees and students in the control group solved them in Andes probability. Both conditions solved the eight physics problems on Andes physics. Students could also access the domain textbook at any time during training. During the probability training, students in the experimental group were able to access a description of the TVS. Each main domain principle was applied at least twice in both trainings.

Post-tests

Finally, all students took a *post-test.* Five problems on both post-tests were isomorphic to training problems in Phase 3. In addition, there were five non-isomorphic, novel, multiple-principle problems in the probability post-test and eight in the physics post-test. Table 5 shows the distribution of single-principle and multiple-principle problems in the experiment.

Table 5. Number of various problems during pre-training, pre-test, training, and post-test

		Single-principle	Multiple-principle	Total
Probability	Pre-training	14	5	19
	Pre-test	10	4	14
	Training	–	12	12
	Post-test	10	10	20
Physics	Pre-training	11	3	14
	Pre-test	9	5	14
	Training	–	8	8
	Post-test	5	13	18

Only the students in the experimental group took the third part, Andes interface training. Its purpose was to familiarize them with the Andes GUI without introducing any new domain knowledge. The problem used was one of the 12 probability training problems that they had previously solved on Pyrenees. Pilot studies showed that one problem was sufficient for most students to become familiar with Andes GUI.

To summarize, the procedural difference between the two conditions were: 1) during probability instruction, students in the experimental group trained on Pyrenees while students in the control group trained on Andes probability; 2) students in the experimental group learned how to use the Andes user interface before they received physics instruction.

Grading criteria

We used two scoring rubrics: binary and partial credit. Under the binary rubric, a solution is worth 1 point if it is completely correct or 0 if not. Under the partial credit rubric, each problem score is a proportion of correct principle applications evident in the solution. If they correctly applied four of five possible principles they would get a score of 0.8. Solutions were scored by a single grader blind to conditions.

Results

In order to measure aptitude-treatment interaction, we needed to define high and low groups based on some measure of incoming competence. We chose to use MSAT scores because probability and physics are both math-like domains. Our split point was 640, which divide into high (n = 20) and low (n = 22). Except for the MSAT scores and high-school GPA, no significant difference was found between high and low on other background information such as age, gender, VSAT scores, and so on. As expected, the high group out-performed the low group during the probability pre-training and the probability pre-test under the binary scoring rubric: $t(40) = 3.15$, $p = 0.003$, $d = 0.96$, $t(40) = 2.15$, $p = 0.038$, $d = 0.66$, and $t(40) = 2.27$, $p < 0.03$, $d = 0.70$ on single-principle, multiple-principle problems during probability pre-training, and overall in probability pre-test, respectively. The same pattern was found under partial rubric in the probability pretest. Thus, the MSAT score successfully predicted the incoming competence of the students, which justifies using it to define our high vs. low split.

Incoming competence combined with conditions partitioned the students into four groups: high-experimental (n = 10), low-experimental (n = 10), high-control (n = 10), and low-control (n = 12). Fortunately, random assignment balanced the experimental vs. control conditions for ability, and this balance persisted even with the groups subdivided into high and low via MSAT score. On every measure of incoming competence, no significant difference was found between the experimental and control groups, the low-experimental and low-control ones, or the high-experimental and high-control ones. These measures were the background survey, the probability pre-test, probability pre-training scores, the time spent reading the probability textbook, and the time spent solving the pre-training problems. Averaged over all students, the total time for each training phase were 2.4 hrs and 2.7 hrs for probability pre-training and training and 1.5 hrs and 3.0 hrs for physics pre-training and training, respectively. No significant differences were found among the four groups on any of these times.

Test scores

Error! Reference source not found. shows that the test score results are consistent with our hypothesis. After training on Pyrenees, the low-experimental students scored significantly higher than their low-control peers on all three assessments: probability post-test, physics pre-test and physics post-tests: $t(20) = 4.43$, $p < 0.0005$, $d = 1.90$; $t(20) = 3.23$, $p < 0.005$, $d = 1.34$; and $t(20) = 4.15$, $p < 0.0005$, $d = 1.84$, respectively. More importantly, the low-experimental students even seemed to catch up with the high ones: no significant difference was found among the high-experimental, low-experimental, and high-control on all three assessments, even though the two experimental groups seemed to out-perform the high-control group in Figure 6.

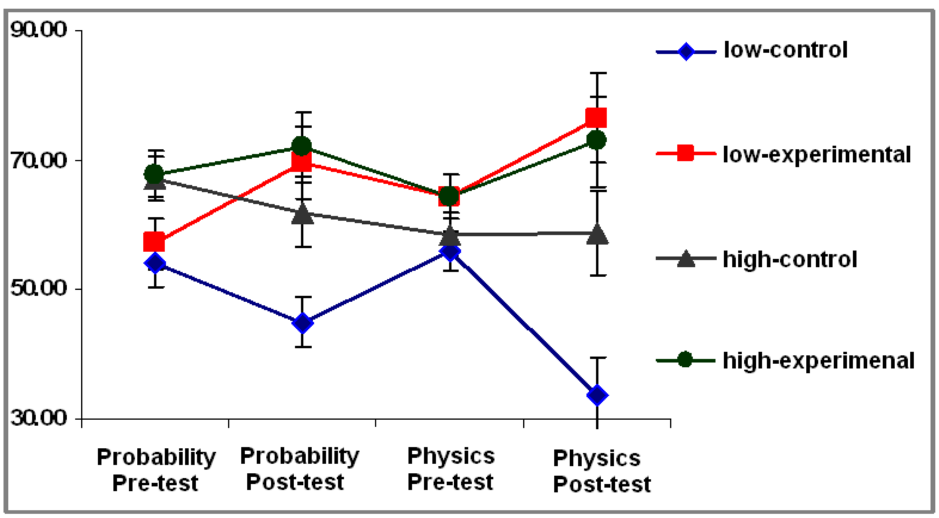

Figure 6. Comparison of four groups on four tests (maximum score = 100)

Thus, the Pyrenees instruction in probability caused the low-experimental group to learn more effectively than those in the low-control group during probability training, physics training, and even physics pre-training. They seemed to have caught up to the high ones while the low-control ones did not. Moreover, while the high-experimental group didn't benefit much from the TVS, they were not harmed either.

Dynamic assessments

While test results are the most common assessment of learning performance, one can also compare students' behaviors as they learn. Such comparisons are called dynamic assessments (Haywood & Tzuriel 2002). In performing dynamic assessments, we can identify students who are effective learners even though their test scores may be equal to or even lower than those of others. Here we investigated students' interactive behaviors on Andes during physics training, as all students received the identical procedure during that period.

Frequency of help requests

Andes physics logs every user's interface action performed, including help requests, tool usage, and equation entries. We first tried to characterize the overall difference in students' solutions via the amount of help they requested. On each of eight physics training problems, the low-experimental students made significantly fewer next-step help requests than the low-control ones. No significant difference was found among the low-experimental, high-experimental and high-control groups. This suggests that the low-experimental students may have transferred the TVS. However, there are other possible explanations, so we conducted several other analyses.

Triage of logs

Solution logs were grouped into three categories: smooth, help-abuse, and rocky. Smooth solutions included no help requests, except on problems that required more than eight principle applications. Students were permitted up to two what's-wrong help requests. Help-abuse solutions were produced when every entry was derived from one or more next-step helps. Otherwise, the solution was categorized as rocky because students appeared capable of solving part of the problem on their own, but needed help on the rest.

Figure 7 shows a significant difference among the four groups on the distribution of the three types of solutions. While no significant difference was found between the high-experimental and low-experimental groups, there was a significant difference between the low-control and the high-control groups: $\chi^2(2) = 11.33$, $p(\chi^2) = 0.003$. Furthermore, a significant difference was found between the low-experimental and high-control groups: $\chi^2(2) = 15.322$, $p(\chi^2) < 0.001$, and between the high-experimental and high-control groups: $\chi^2(2) = 11.585$, $p(\chi^2) < 0.005$. Qualitatively, the results appear to be as follows: high-experimental = low-experimental > high-control > low-control.

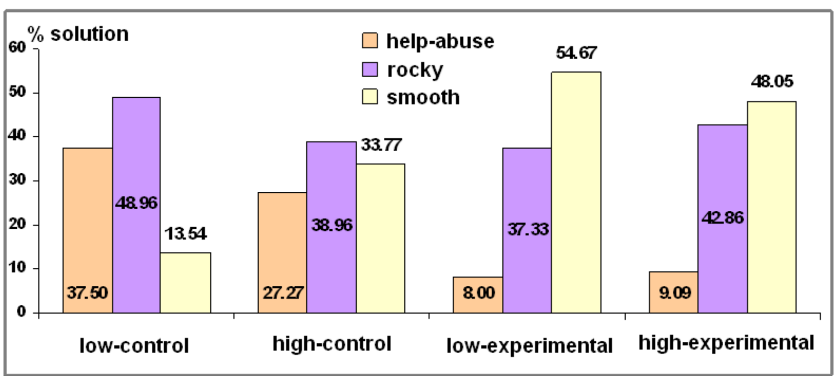

Figure 7. Solution percentage by type

For a more quantitative measure, we used a smaller unit of analysis, individual equations. We coded each correct equation entry in the solution logs with 3 features:

- *Relevance*: The equation was labeled relevant or irrelevant based on whether it contributed to the problem solution.
- *Help:* The equation was labeled "Help" if it was entered after the student asked for help from Andes physics. Otherwise, it was labeled "No-help."
- *Content*: The equation's content was coded as either "a correct equation with new physics content" or "others."

We sought to find out how frequently students made progress toward solving a problem without asking for any help from Andes. In terms of the three-feature coding mentioned above, such a "desirable" equation would be coded as "relevant," "no-help," or "correct equation with new physics content." We called these desirable equations *desirable* steps and defined the desirable steps ratio (DSR):

$$DSR = \frac{Desirable\ \ steps\ \ in\ the\ solution}{All\ \ steps\ \ in\ the\ solution}$$

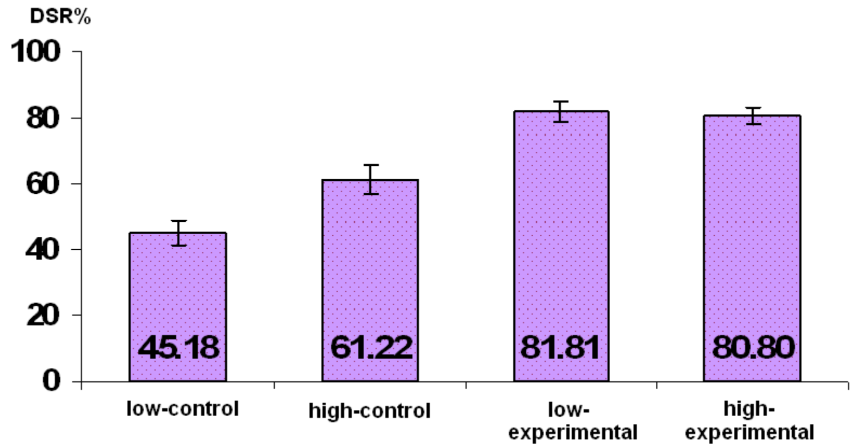

Figure 8. DSR on overall solutions

As shown in Figure 8, the low-experimental students had significantly higher DSR than the low-control ones: $t(169) = 7.50$, $p < 0.0001$. In fact, the former even made significantly more progress than the high-control group: $t(150) = 3.84$, $p < 0.001$. While there is a significant difference between the low-control and high-control groups: ($t(171) = 2.83$, $p < 0.01$), there is no such difference between the two experimental groups. In short, this dynamic assessment showed that the following: high experimental = low experimental > high-control > low-control.

To summarize, both test scores and dynamic assessments show that the low students catch up with the high ones in the experimental condition but not in the control condition. On some measures, the low-experimental students even surpass the high-control ones. Next, we'll investigate what was transferred from probability to physics that made the low experimental students so successful.

Transferring the two cognitive skills of the TVS

As we described above, the TVS includes two main components: solving problems via backward-chaining (BC) from goals to givens, called the BC-strategy, and drawing students' attention to the characteristics of each individual domain principle, called the principle-emphasis skill. In the following, we will investigate whether either or both skills were transferred by the two experimental groups to physics. In order to determine the BC-strategy usage, we analyzed students' logs to see whether the order of equations in their solutions followed the BC strategy. For the principle-emphasis skill, we used the single-principle problems as our litmus test. Students who had applied the BC-strategy would have no particular advantage because solving these single-principle problems need to apply only one principle. On the other hand, students who had learned the idea of focusing on domain principles should be at an advantage.

Transferring the BC strategy

If students engaged in the BC-strategy, we expect they would apply the BC strategy when they had difficulties, that is, on rocky solutions. On smooth solutions, students don't have any difficulties since they may solve problems mainly based on existing schemas (Sweller, 1989). Thus, we subcategorized each desirable step in the logs as BC or non-BC, where non-BC included FC, combined equations, and so on. We then defined BC% as the proportion of desirable steps that were coded as BC. Figure 9 showed that on Rocky solutions the high-experimental group applied BC significantly more frequently than the other three groups: $t(40) = 2.25$, $p = 0.03$ while the low-experimental group used the BC as frequently as the two control groups. Thus, apparently it was the high-experimental group alone who transferred the BC-Strategy to physics.

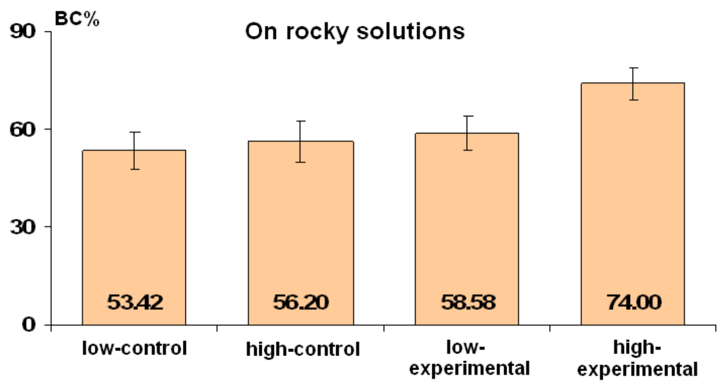

Figure 9. BC usage on rocky solutions

Transfer of the principle-emphasis skill

The low-experimental students scored just as high as the high-experimental ones even though they used BC no more frequently than the students in the two control groups. Our hypothesis is that they transferred the principle-emphasis skill. We divided both post-tests into single-principle and multiple-principle problems. Furthermore, we divided the multiple-principle problems into those that were isomorphic to a training problem and those that were not. If students in the low-experimental group applied the principle-emphasis skill, we expected them to out-perform students in the low-control group on all three types of problems in both post-tests. This turned out to be the case (see Table 6). In Table 6, the third and fourth columns list the means of test scores of the low-experimental and low-control groups. The low experimental group had reliably higher means than the low-control group in both probability and physics post-tests across three types of problems: simple-principle, isomorphic multiple-principle and non-isomorphic multiple-principle. This suggests that a main effect of teaching the TVS to the low students was to get them to focus on the domain principles. Further analysis showed no significant difference among the students in high-control, low-experimental, and high-experimental groups on any types of problems, which indicates that high students may already have such skill.

Table 6. Scores on three types of problems in both probability and physics post-tests

Test	Problem type	Mean (low-experimental)	Mean (low-control)	Statics
Probability post-test	Single	0.93	0.70	$t(20) = 3.62$, $p = 0.002$, $d = 1.58$
	Multiple, isomorphic	0.48	0.23	$t(20) = 3.71$, $p = 0.001$, $d = 1.55$
	Multiple, non-isomorphic	0.44	0.17	$t(20) = 3.734$, $p = 0.013$, $d = 1.15$
Physics post-test	Single	0.93	0.8	$t(20) = 4.33$, $p < 0.001$, $d = 1.85$
	Multiple, isomorphic	0.60	0.18	$t(20) = 4.55$, $p < 0.001$, $d = 1.93$
	Multiple, non-isomorphic	0.70	0.19	$t(20) = 3.734$, $p < 0.001$, $d = 2.10$

Conclusions

Overall, our instructional manipulation indeed exhibited an aptitude-treatment interaction: the gap between high-experimental and low-experimental students seemed to be eliminated in both probability and physics, whereas it remained between the high-control and low-control groups. More detailed analyses of the training behavior and post-test results suggest that students in the low-experimental group transferred the principle-emphasis skill to physics while those in the high-experimental group apparently already possessed it. On the other hand, students in the high-experimental group transferred the BC strategy.

These results suggest that it is not the BC strategy that is most important to teach low learners. Instead, one should teach the meta-cognitive skill of focusing on individual principle applications. It could be that low and high learners differed initially in that low students lacked this “how to learn” meta-cognitive knowledge for a principle-based domain like probability or physics. Such results suggest building an ITS that does not teach the TVS explicitly, but instead just teaches to focus on principle applications in deductive domains. Perhaps it would be just as effective as Pyrenees. Indeed, because its students need not learn all the complicated bookkeeping of the BC strategy, which may cause cognitive overload (Sweller, 1989), it might even be more effective than Pyrenees not only for an initial domain where the ITS was used but also for subsequent domains where it is not used.

References

Bhaskar, R., & Simon, H. A. (1977). Problem solving in semantically rich domains: An example from engineering thermodynamics. *Cognitive Science, 1,* 193–215.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13,* 4–16.

Chi, M. & VanLehn, K. (2007). Accelerated future learning via explicit instruction of a problem solving strategy. In R. Luckin, K. R. Koedinger, & J. Greer (Eds.) *The13th International Conference on Artificial Intelligence in Education* (pp. 409–416). Amsterdam, Netherlands: IOS Press.

Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington.

Haywood, H.C. & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education, 77(2),* 40–63.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, 208,* 1335–1342.

Lehman, D. R., Lempert, R. O., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist, 43,* 431–442.

Lehman, D. R., & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Developmental Psychology, 26,* 431–442

Newell, A., & Simon, H. A. (1972). *Human problem-solving.* Engelwood Cliff, NJ: Prentice-Hall.

Owen, E. & Sweller, J., (1985) What do students learn while solving mathematics problems? *Journal of Educational Psychology 77(3),* June 1985, 272–284.

Priest, A. G., & Lindsay, R. O. (1992). New light on novice–expert differences in physics problem-solving. *British Journal of Psychology, 83,* 389–405.

Russell, Stuart J. & Norvig, Peter (2003), *Artificial Intelligence: A Modern Approach* (2nd Ed.), Upper Saddle River, NJ: Prentice Hall.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), *Children’s thinking: What develops?* Hillsdale, NJ: Lawrence Erlbaum Associates.

Sweller, J. (1988). Cognitive load during problem-solving: Effect on learning. *Cognitive Science, 12,* 257–285

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem-solving in mathematics and science. *Journal of Educational Psychology, 8 (4),* 457–466.

Tarmizi R. A., & Sweller, J. (1988). Guidance during mathematical problem-solving, *Journal of Educational Psychology 80(4),* 424–436.

Trafton, J. G., & Reiser, B. J. (1991). Providing natural representations to facilitate novices’ understanding in a new domain: Forward and backward reasoning in programming. *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society* (pp. 923–927). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

VanLehn, K. (2006) The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education. 16* (3) 227–265.

VanLehn, K., Bhembe, D., Chi, M., Lynch, C., Schulze, K., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2004). Implicit versus explicit learning of strategies in a non-procedural cognitive skill. In J. C. Lester, R. M. Vicari, & F. Paraguacu (Eds.), *Intelligent Tutoring Systems: 7th International Conference* (pp. 521-530). Berlin: Springer-Verlag.

VanLehn, K., Lynch, C., Schulze, K. Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence and Education, 15(3),* 1–47.

McQuiggan, S. W., Robison. J. L., & Lester, J. C. (2010). Affective Transitions in Narrative-Centered Learning Environments. *Educational Technology & Society, 13* (1), 40–53.

Affective Transitions in Narrative-Centered Learning Environments

Scott W. McQuiggan[1], Jennifer L. Robison[2] and James C. Lester[3]
[1]Education Practice, SAS Institute Inc. // scott.mcquiggan@sas.com
[2]Department of Computer Science, North Carolina State University // jlrobiso@ncsu.edu
[3]Department of Computer Science, North Carolina State University // lester@csc.ncsu.edu

ABSTRACT

Affect has been the subject of increasing attention in cognitive accounts of learning. Many intelligent tutoring systems now seek to adapt pedagogy to student affective and motivational processes in an effort to increase the effectiveness of tutorial interaction and improve learning outcomes. To this end, recent work has begun to investigate the emotions experienced during learning in a variety of environments. In this paper we extend this line of research by investigating the affective transitions that occur throughout narrative-centered learning experiences. Further analysis differentiates the likelihood of affective transitions stemming from pedagogical agent empathetic responses to student affect.

Keywords

Affective transitions, Narrative-centered learning environments, Empathetic pedagogical agents

Introduction

Affect has begun to play an increasingly important role in intelligent tutoring systems (ITS). The ITS community has seen the emergence of work on affective student modeling (Conati & Mclaren, 2005), detecting frustration and stress (Burleson, 2006; McQuiggan, Lee, & Lester, 2007; Prendinger & Ishizuka, 2005), modeling student uncertainty (Forbes-Riley & Litman, 2007), modeling agents' emotional states (André & Mueller, 2003; Gratch & Marsella, 2004; Lester, Towns, & FitzGerald, 1999), devising affectively informed models of social interaction (Johnson & Rizzo, 2004; Paiva et al., 2005; Porayska-Pomsta & Pain, 2004; Wang et al., 2008), detecting student motivation (de Vicente & Pain, 2002), and diagnosing and adapting to student self-efficacy (Beal & Lee, 2005; McQuiggan, Mott, & Lester, 2008). All of this work seeks to increase the fidelity with which affective and motivational processes are understood and utilized in intelligent tutoring systems in an effort to increase the effectiveness of tutorial interactions and, ultimately, learning.

Recent work seeking to characterize the affective experience of learners interacting with intelligent learning environments has considered student affective trajectories occurring during learning. D'Mello, Taylor, & Graesser (2007) studied the likelihood of affective transitions among six affective states (boredom, flow, confusion, frustration, delight, and surprise) that were found to be relevant to complex learning (Craig, Graesser, Sullins, & Gholson, 2004). In general, learners are likely to persist in the same affective state (e.g., transitioning from a state of boredom to boredom is likely, and in some cases, significantly more likely than transitioning to another affective state). This analysis was conducted in the AutoTutor learning environment (Craig et al., 2004; D'Mello et al., 2007). Baker, Corbett, Koedinger, & Wagner (2004) were able to replicate many of D'Mello et al.'s (2007) findings when they calculated the likelihood of affective transitions in the Incredible Machine: Even More Contraptions, a simulation-based learning environment (2007). Baker et al. (2004) extend their analyses to investigate how usage choices affect emotion transitions. This work found that confused learners are likely to game the system. Further, it was found that students who game the system are unlikely to transition into a confused state (Baker, Rodrigo, & Xolocotzin, 2007).

In this article we investigate the likelihood of affective transitions in a narrative-centered learning environment, CRYSTAL ISLAND. The CRYSTAL ISLAND environment uses narrative as a mechanism to contextualize learning, making the experience meaningful. Contextualized learning experiences are known to encourage regulated learning behavior (Perry, 1998) and influence student learning and motivation (Linnenbrink & Pintrich, 2001). Because CRYSTAL ISLAND incorporates an engaging storyline into the learning experience, we supplement the known relevant emotions to learning used by D'Mello et al. (2007) and Baker et al. (2007) with affective states that may be relevant to the story (anger, anxiety, boredom, confusion, delight, excitement, fear, flow, frustration, and sadness). We extend our analysis of affective transitions to evaluate the impact of character empathetic responses (parallel vs. reactive empathy) to student affect and the relative impact on transitions. We further extend our analysis to investigate whether additional factors may affect the frequency of transitions between affective states, turning our attention to characteristics of the students. We have chosen four characteristics to examine based on their potential influence on learning and reaction to the learning environment: gender, personality, goal orientation, and presence.

The article is organized as follows. First we describe CRYSTAL ISLAND, the narrative-centered learning environment that has been developed in our lab for the domains of microbiology and genetics. Next, we present the experimental method utilized to study student affective experiences. We then report findings on probable transitions in narrative-centered learning and present analyses of the impact of empathy on such transitions. We discuss results and note limitations, then provide conclusions and an indication of future work.

CRYSTAL ISLAND

The CRYSTAL ISLAND environment (Figure 1 and Figure 2) is being created for the domains of microbiology and genetics for middle-school students. It features a science mystery set on a recently discovered volcanic island, where a research station has been established to study the unique flora and fauna. The user plays the protagonist, Alex, who attempts to discover the genetic makeup of the chickens at the research station whose eggs carry an unidentified infectious disease. The story opens by introducing the student to the island and the members of the research team for which the student's father serves as the lead scientist. As members of the research team fall ill, it is the student's task to discover the cause and the specific source of the outbreak. She is free to explore the world and interact with other characters while forming questions, generating hypotheses, collecting data, and testing her hypotheses. Throughout the mystery, she can walk around the island and visit various locations. She can pick up and manipulate objects and talk with characters to gather clues about the source of the disease. In the course of her adventure she must gather enough evidence to correctly choose which breeds of chickens need to be banned from the island. The virtual world of CRYSTAL ISLAND, the semi-autonomous characters that inhabit it, and the user interface were implemented with Valve Software's Source engine, the 3D game platform for Half-Life 2. The Source engine also provides much of the low-level (reactive) character behavior control. The character behaviors and artifacts in the storyworld are the subject of continued work.

Figure 1. Overview of CRYSTAL ISLAND

Figure 2. The user, Alex, with Jin, the camp nurse on CRYSTAL ISLAND

The following scenario illustrates a student's interactive narrative experience in CRYSTAL ISLAND. In the course of having members of her research team become ill, she learns that an infectious disease is an illness that can be transmitted from one organism to another. As she concludes her introduction to infectious diseases, she learns from the camp nurse that the mystery illness seems to be coming from eggs laid by certain chickens and that the source of the disease must be identified. The student discovers through a series of tests that the bad eggs seem to be coming from chickens with white feathers. The student then learns that this is a co-dominant trait and determines that any chicken containing the allele for white feathers must be banned from the island immediately to halt the spread of the disease. The student reports her findings back to the camp nurse.

Method

After describing the participants, we introduce the experimental design. We then present the results and discuss the affective transitions observed in narrative-centered learning experiences.

Participants

The subjects of the study were 35 graduate students ranging in age from 21 to 60 (M = 24.4, SD = 6.41) including 9 females and 26 males. Among these students, 60% were Asian (n = 21) and approximately 37% were Caucasian (n = 13). One participant chose not to respond.

Procedure

Participants entered the experiment room where they completed informed-consent documentation. They were randomly assigned to either the control condition or the empathy condition and were seated in front of a laptop computer. They were then given an overview of the experiment agenda, and they completed the pre-experiment questionnaires including the demographics survey, the interpersonal reactivity index survey (Davis, 1994), the goal orientation survey (Elliot & McGregor, 2001), and the personality questionnaire (McCrae & Costa, 2003).

Upon completing the pre-experiment questionnaires, participants were instructed to review CRYSTAL ISLAND instruction materials. These materials consisted of the backstory and task description, character overviews, a map of the island, a control sheet, and a definition sheet of the self-report emotions. Participants were then further briefed on the controls via a presentation summarizing the task and explaining each control in detail. Participants maintained access to the materials, including the definition sheet of the self-report emotions, throughout their interaction. Participants were given 35 minutes to solve the mystery. Solving the mystery consisted of completing 15 goals including learning about various diseases, compiling the symptoms of the sickened researchers, testing a variety of possible sources, and reporting the solution (cause and source) back to the camp nurse.

Six CRYSTAL ISLAND characters (Audrey, Elise, Jin, Quentin, Robert, and Teresa), each play distinct roles in the CRYSTAL ISLAND environment. When subjects decided to interact with these particular characters, they were greeted with empathetic reactions to their expressed affective state, which they communicated through self-reports via an in-game dialog. The self-report dialog asked participants to select the affective state that best described their feelings at that time from a set of 10 affective states (anger, anxiety, boredom, confusion, delight, excitement, fear, flow, frustration, and sadness). This set of emotions was comprised of emotions identified with learning (Craig et al., 2004; D'Mello et al., 2007; Kort et al., 2001), together with basic emotions (Ekman & Friesen, 1978) that may play a role in students' experience of the CRYSTAL ISLAND narrative.

Immediately after solving the science mystery of CRYSTAL ISLAND (or after 35 minutes of elapsed interaction time for subjects who had not solved the mystery), subjects completed a post-experiment questionnaire. This researcher-designed questionnaire assessed perceptions of individual CRYSTAL ISLAND characters. The results of this instrument are outside the scope of this discussion. Additionally, participants' presence experience was captured with the presence questionnaire (PQ), which was developed and validated by Witmer and Singer (1998). The PQ contains several subscales, including involvement/control, naturalism of experience, and quality of the interface scales. The PQ accounts for four categories of contributing factors of presence: control, sensory, distraction, and realism.

Results

In this section, we first present findings regarding common affective transitions observed in CRYSTAL ISLAND. These findings are followed by an analysis comparing and contrasting likely affective transitions stemming from parallel and reactive empathetic reactions by CRYSTAL ISLAND characters.

To compute transition likelihoods, we adopt D'Mello et al.'s *L* (2007), which is based on Cohen's Kappa (1960), and has been used by Baker et al. for affective transition analysis in their simulation learning environment (2007). *L* computes the probability that a transition between two affective states (*Current* → *Next*) will occur, where *Current* refers to a reported emotion at time *t*, while *Next* refers to the next reported emotion at time *t* + 1. D'Mello et al.'s *L* accounts for the base frequency of the *Next* affective state in assessing the likelihood of a particular transition (2007). Formally,

$$L\,(\text{Current} \rightarrow \text{Next}) = \frac{P(\text{Next} \mid \text{Current}) - P(\text{Next})}{1 - P(\text{Next})}$$

Table 1. Likelihoods for all transitions *Current* → *Next* for the affective states:
Frustration, **Fl**ow, **Co**nfusion, **De**light, **Bo**redom, **Anx**iety, **Ex**citement, **Ang**er, **Sa**dness, and **Fe**ar

					Next					
Current	***Fr***	***Fl***	***Co***	***De***	***Bo***	***Anx***	***Ex***	***Ang***	***Sa***	***Fe***
Fr	0.28	−0.19	0.10	−0.05	−0.07	−0.15	−0.10	−0.02	−0.01	0.09
Fl	−0.04	0.19	0.04	0.02	−0.01	0.03	−0.07	0.01	0.00	0.00
Co	0.04	0.04	0.16	−0.03	0.05	−0.04	0.10	−0.01	−0.01	−0.03
De	0.01	0.10	−0.13	0.21	−0.03	−0.05	−0.33	−0.02	0.00	0.00
Bo	0.13	−0.03	−0.03	−0.08	0.13	−0.04	−0.04	0.00	−0.03	0.04
Anx	−0.08	0.06	0.04	−0.07	−0.01	0.14	−0.19	0.09	0.00	0.00
Ex	−0.05	−0.11	0.06	−0.03	−0.03	0.03	0.24	−0.01	0.01	−0.02
Ang	0.00	−0.07	0.09	−0.39	0.00	0.23	0.01	0.00	0.00	0.00
Sa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fe	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

L's numerator is divided by 1−*P(Next*) to normalize scores between -∞ and 1 (2007). A result of *L* equal to 1 translates to emotion *Next* always following the *Current* emotion; an *L* value equal to 0 means the likelihood of transitioning to emotion *Next* is equal to chance, that is, the probability of experiencing *Next* (the base rate) regardless of the *Current* emotion. An *L* value less than 0 translates to the likelihood of transitioning to emotion *Next* being less than chance (the probability of experiencing *Next* regardless of the *Current* emotion).

To characterize affective transitions we first compute *L* for each transition (*Current* → *Next*), for each student. We then use mean *L* values across students to determine the likelihood of transitioning from each emotion *Current* to each emotion *Next*. The results of ANOVAs determine whether the differences in likelihoods of transitioning to each *Next* emotion are significantly different for particular *Current* emotions.

Affective transitions

Aggregating self-reported affective states across the 35 participants, we find flow to be the most frequently reported state (42%), followed by excitement (14%), confusion (13%), delight (11%), anxiety (8%), frustration (6%), boredom (3%), sadness (2%), anger (1%), and fear (1%).

ANOVAs indicated that six affective states had statistically significant differences among the likelihoods of transitions. Affective transitions were statistically significantly different transitioning from frustration ($F(9, 340) = 2.06$, $p = .03$), flow ($F(9, 340) = 18.3$, $p < .0001$), confusion ($F(9, 340) = 1.79$, $p = .06$), delight ($F(9, 340) = 5.22$, $p < .0001$), anxiety ($F(9, 340) = 2.98$, $p = .002$), and excitement ($F(9, 340) = 2.62$, $p = .006$).

Notably, frustrated learners are most likely to remain frustrated (Mean L = .28) followed by transitions to confusion (0.10) and fear (0.09). The remaining transitions were below chance levels (i.e., flow (−0.19, $t(34) = -4.24$, $p < .0001$) and excitement (−0.10)).

Table 2. Interesting likelihood for transition differences by empathetic response type (parallel or reactive)

Current	***Transition state (Next)***	***Parallel empathy likelihood***	***Reactive empathy likelihood***
Boredom	Boredom	.35	−.04
	Confusion	0	−.41
	Flow	−.13	.32
	Frustration	−.08	.26
Anxiety	Anxiety	.33	.05
	Frustration	−.20	.17
Frustration	Frustration	.57	−.13
	Flow	.10	−.25
	Confusion	−.17	.15
Flow	Flow	.11	−.05
	Confusion	.04	.08
Delight	Delight	.21	.21
	Flow	.07	.17

Learners in the state of flow were most likely to remain in flow (0.19) followed by confusion (.04, $t(34) = -3.09$, $p = .003$), anxiety (0.03), and delight (0.02). Both frustration (−0.04, $t(34) = -7.91$, $p < .0001$), and excitement (−0.07) were below chance levels. The remaining transitions did not occur or occurred at chance levels.

Confused students were likely to remain in a confused state (0.16) followed by excitement (0.10), boredom (0.05), frustration (0.04), and flow (0.04). The likelihood of these and all remaining conditions is summarized in Table 1.

Affective transitions by empathy

Empathy is the expression of emotion based on another's situation and not merely one's own (Davis, 1994). Its expression can demonstrate that the feelings of the target (the recipient of empathetic expression) are understood or shared. In the case of parallel empathy, an individual exhibits an emotion similar to that of the target (Davis, 1994). This is typically based on an understanding of the target's situation and shows the empathizer's ability to identify with the target. Reactive empathy, in contrast, focuses on the target's affective state, in addition to her situation (Davis, 1994). Reactive empathizers will display emotions that are different from the target's, often in order to alter or enhance the target's own affective state. This type of empathy is focused on the target, whereas parallel empathy is more self-oriented. As such, reactive empathy can be viewed as a higher level of empathetic behavior.

Recent research with the characters of CRYSTAL ISLAND has investigated the merit of providing characters with empathetic capabilities to effectively respond to unfolding student experiences (McQuiggan, Robison, Phillips, & Lester, 2008; McQuiggan, Rowe, & Lester, 2008). In CRYSTAL ISLAND, empathetic responses are short, text-based responses consisting of one or two sentences. Parallel responses consist of the character expressing the same emotion as the user through text responses. On the other hand, reactive responses demonstrate advanced cognitive processing on the character's part by providing responses designed to be more motivating, thus revealing the character's desire for the user to be in a positive emotional state. The results below investigate the likelihood of affective transitions based on empathetic expressions by CRYSTAL ISLAND characters in response to student *Current* emotions. The findings suggest that in certain situations, parallel and reactive empathy have significant differences in the affective transitions (*Next* emotions) that are likely to occur.

While the relatively low frequencies of some transitions prevent many of the visible differences from being statistically significant, interesting patterns do emerge. Figure 3 and Figure 4 present the transitions from the state of flow and frustration by empathetic reaction type (parallel or reactive). Analyzing the transitions from the state of flow, we find that parallel empathy (0.11) is somewhat significantly more likely to support students' remaining in the state of flow than reactive empathy (−.05), $t(12) = -2.08$, $p = .06$. Similarly, we find that the likelihood of transitioning to frustration from a frustrated state is significantly greater when characters' empathetic reactions are more parallel in nature (0.57) than reactive (−0.13), $t(12) = -2.09$, $p = .059$. Other patterns with visible differences emerging from this analysis of affective transitions are summarized in Table 2. Although the transition frequencies were not sufficiently high for the differences to be statistically significant, they merit discussion.

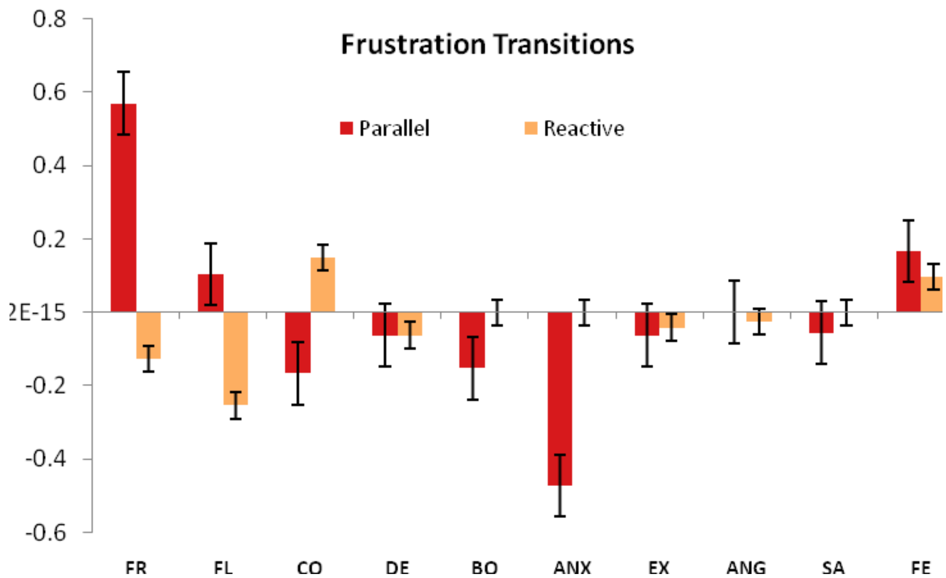

Figure 3. Transitions from frustration to **FR**ustration, **FL**ow, **CO**nfusion, **DE**light, **BO**redom, **ANX**iety, **EX**citement, **ANG**er, **SA**dness, and **FE**ar

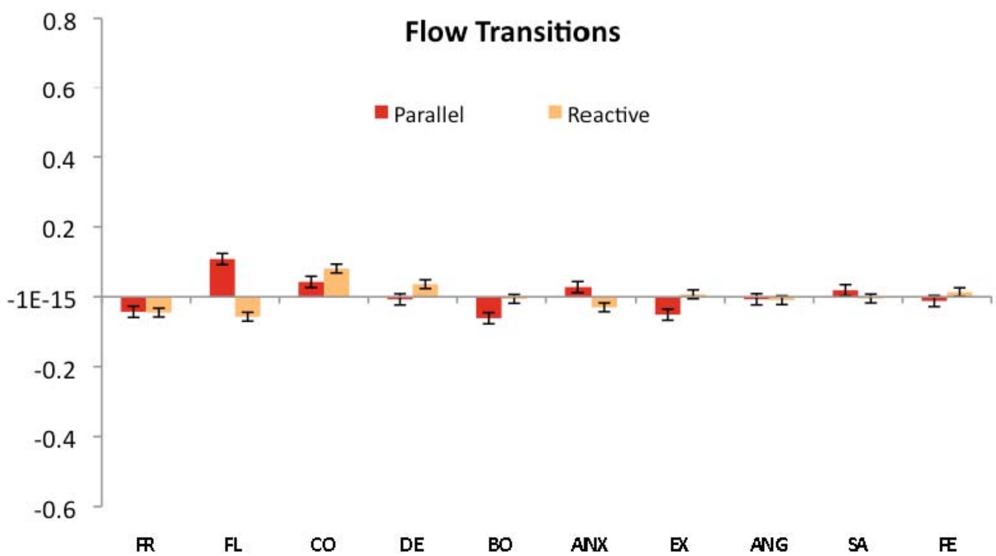

Figure 4. Transitions from flow to **FR**ustration, **FL**ow, **CO**nfusion, **DE**light, **BO**redom, **ANX**iety, **EX**citement, **ANG**er, **SA**dness, and **FE**ar

Student characteristics results

In this section we analyze the individual differences with which affective states are reported. This examination includes demographics, personality, goal orientation, and presence. These findings are followed by a summary of individual differences in affective transitions.

Gender refers to an individual's identification as male or female. Interestingly, significant differences have been found in how male and female students approach learning tasks. For example, women are more likely than men to perceive intelligence as an immutable entity that cannot be improved with increased focus on learning tasks (Lips, 2007). This belief may mean that women are more likely to experience negative emotions such as frustration and confusion, and also experience vicious cycles (D'Mello et al., 2007). In this case, intervention would be necessary to break students of this cycle and encourage a more dynamic approach to learning.

Personality is an individual's disposition over a long duration of time, distinguishing itself from emotions or moods that are more limited in their duration (Rusting, 1998). Using the Big 5 Personality Questionnaire (McCrae & Costa, 2003), personality is divided into five main categories: openness, conscientiousness, extraversion, agreeableness, and neuroticism. Of particular interest among these are openness, conscientiousness, and neuroticism, as these characteristics are likely to affect emotion and learning. Additionally, since information on affective states was obtained through self-report, we expect to find that individuals who score high on openness will display genuine emotions, while others may limit themselves to what they feel comfortable reporting.

Goal orientation reflects a student's primary objective when engaged in learning activities. Students may either view learning in relation to performance or mastery (Elliot & McGregor, 2001). A performance approach would result in a student's wishing to prove his or her competence and achieve better results than other students. Students with a mastery approach, however, view learning as an attempt to gain a skill, regardless of how their ability compares to others. In addition, students may have avoidance strategies in relation to their goals. For example, students with a performance-avoidance approach would simply try to not overtly fail, rather than try to top their fellow students.

Presence relates to the level of student involvement within the system (Witmer & Singer, 1998). Students who experience high levels of presence will be very engaged with the activity, focusing solely on the task while neglecting their external environment. We expect that these students will experience more salient affective states and have more intense reactions to events within the system. Additionally, significant differences in transitions between students who are and are not present may be able to serve as an indicator of presence.

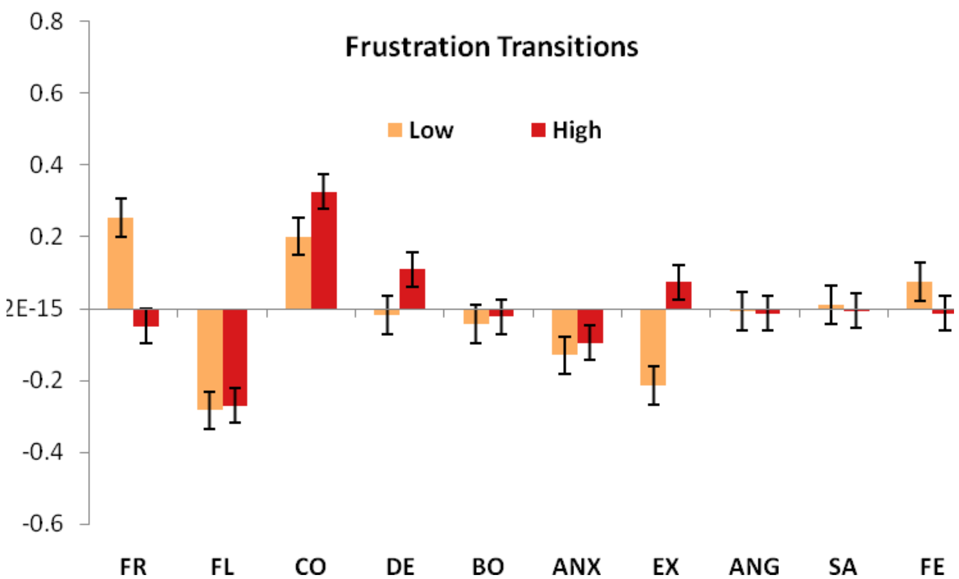

Figure 5. Transitions from frustration to **FR**ustration, **FL**ow, **CO**nfusion, **DE**light, **BO**redom, **ANX**iety, **EX**citement, **ANG**er, **SA**dness, and **FE**ar by level of agreeableness

There were significant differences in the frequencies with which male and female participants reported emotions of boredom. Females (n = 9) did not report feeling bored while the males did, leading to a marginally significant difference, $t(34) = 1.87$, $p = .07$. There were no other significant differences across gender. Student personalities also affected the frequency with which certain affective states were reported, namely, anger, boredom, confusion, delight, and flow. There was a significant difference in the frequency of reported states of flow along the extraversion dimension. Students who were more extraverted reported affective states of flow less frequently than less extraverted students, $t(34) = 2.14$, $p = .04$. Also along the extraversion dimension were differences in the frequencies of delight and anger. Marginally significant was the frequency of which the more extraverted students reported delight than did the less extraverted students, $t(34) = 1.82$, $p = .07$. The more extraverted students reported delight approximately five times per interaction compared to just two times for the less extraverted students. Anger was reported more frequently by the more extraverted students than by the less extraverted students, $t(34) = 2.77$, $p = .009$.

There were significant differences across the personality dimensions of agreeableness (Figure 5 and Figure 6), conscientiousness, and neuroticism in reports of confusion. The less agreeable students reported confusion more frequently ($M = 6.06$, $SD = 1.5$) than the more agreeable students ($M = 2.36$, $SD = 1.4$), $t(34) = 1.77$, $p = .08$. Similarly, the less conscientious students reported confusion more frequently ($M = 6.0$, $SD = 1.43$) than the more conscientious students ($M = 2.0$, $SD = 1.47$), $t(34) = 1.94$, $p = .06$. Students with greater emotional stability (neuroticism dimension) reported confusion more frequently ($M = 7.93$, $SD = 1.48$) than the less emotionally stable students ($M = 1.47$, $SD = 1.2$), $t(34) = 3.37$, $p = .001$. The final significant difference in emotion frequencies along personality dimensions is reports of boredom across student agreeableness. The more agreeable students reported being bored less frequently ($M = 0.1$ $SD = 0.4$) than the less agreeable students ($M = 2.2$ $SD = 0.44$), $t(34) = 3.45$, $p = .001$.

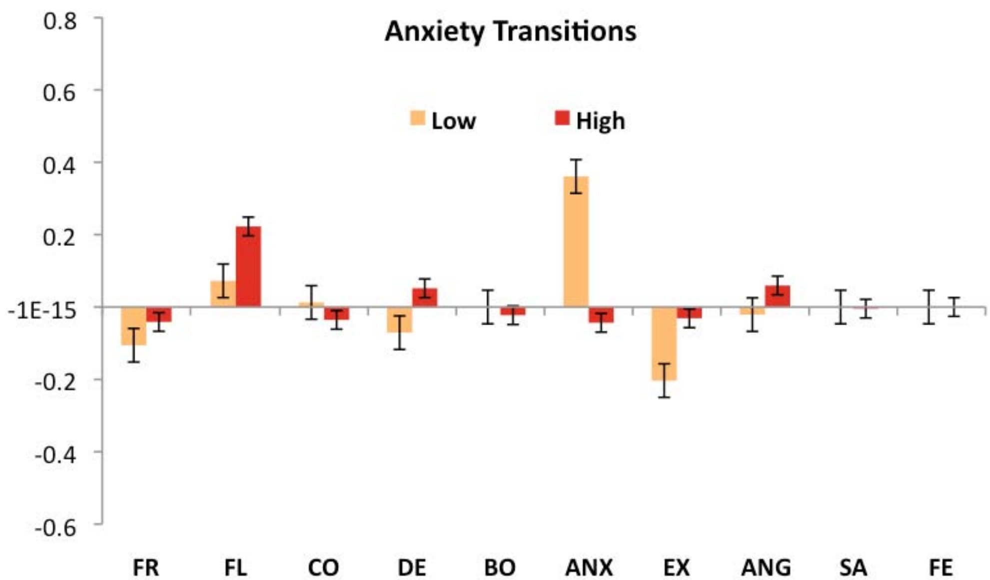

Figure 6. Transitions from anxiety to **FR**ustration, **FL**ow, **CO**nfusion, **DE**light, **BO**redom, **ANX**iety, **EX**citement, **ANG**er, **SA**dness, and **FE**ar by level of agreeableness

Student goal orientation (Figure 7 and Figure 8) also affected the frequency of which students reported anger, anxiety, and flow. Anger was reported more frequently by students scoring higher on the performance approach subscale than students scoring below the performance approach population mean, $t(34) = 2.28$, $p = .03$. Marginally significant was the increased frequency with which students who were dominantly performance oriented reported feeling anxious ($M = 3.62$, $SD = 0.89$) than students who were dominantly mastery oriented ($M = 1.2$, $SD = 1.1$), $t(34) = 1.71$, $p = .09$. Also significant was the frequency with which students scoring high on the performance avoidance subscale reported feeling anxious ($M = 4.05$, $SD = 0.87$) compared to students scoring below the performance avoidance population mean ($M = 0.8$, $SD = 1.01$), $t(34) = 2.43$, $p = .02$. Flow was more frequently reported by students who were dominantly mastery oriented ($M = 18.2$, $SD = 2.8$) than students who were dominantly performance oriented ($M = 10.04$, $SD = 2.2$), $t(34) = 2.25$, $p = .03$. The frequency of flow reports was impacted by students' performance orientations. Students scoring lower on the performance avoidance subscale reported more feelings of flow than students scoring above the performance avoidance population mean, $t(34) = 2.13$, $p = .04$. Comparatively, students scoring lower on the performance approach subscale reported more feelings of flow than students scoring above the population mean for performance approach, $t(34) = 1.87$, $p = .07$.

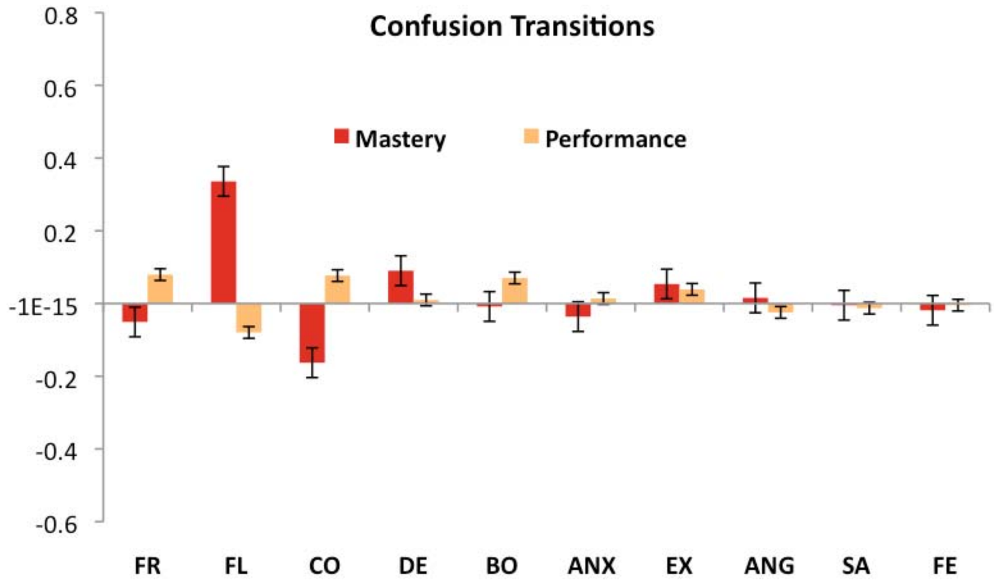

Figure 7. Transitions from confusion to **FR**ustration, **FL**ow, **CO**nfusion, **DE**light, **BO**redom, **ANX**iety, **EX**citement, **ANG**er, **SA**dness, and **FE**ar by dominant goal orientation

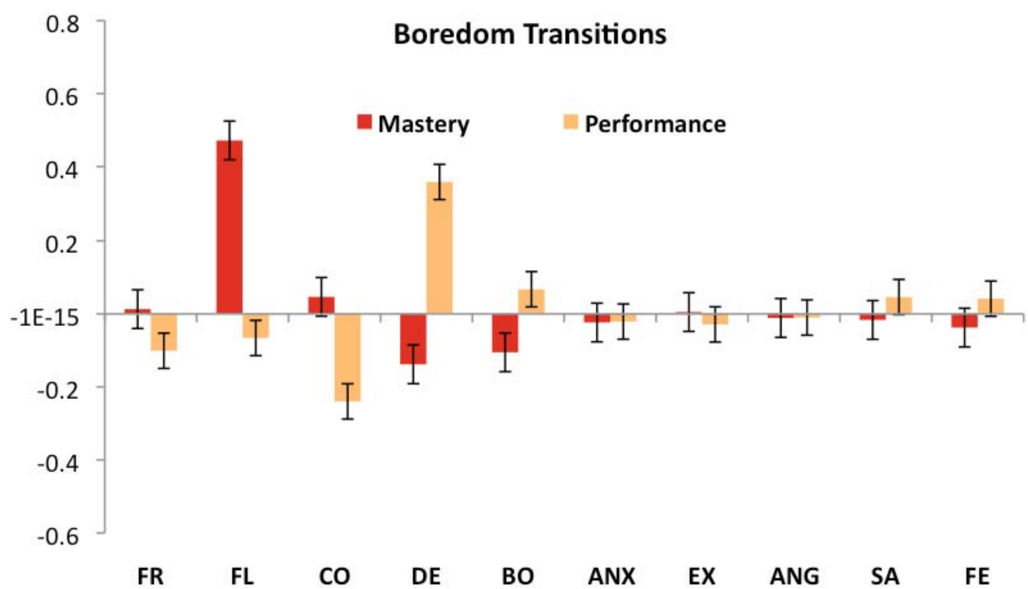

Figure 8. Transitions from boredom to **FR**ustration, **FL**ow, **CO**nfusion, **DE**light, **BO**redom, **ANX**iety, **EX**citement, **ANG**er, **SA**dness, and **FE**ar by dominant goal orientation

Lastly, there were differences in the frequency of reports of frustration and anxiety across students' reported sense of presence (Figure 9 and Figure 10). Students scoring below the population mean of the presence questionnaire reported frustration with greater frequency than students reporting a greater sense of presence with marginal significance, $t(34) = 1.70$, $p = .09$. Anxiety was reported more frequently by students scoring above the population mean on the presence questionnaire than by students reporting lower levels of presence, $t(34) = 2.23$, $p = .03$.

There were few statistically significant differences in affective transitions across individual differences. This is likely due to a small population size (n = 35) resulting in small split population sizes. However, there are noticeable trends that may be concretely uncovered in a large-scale study. We report on several of these trend findings below.

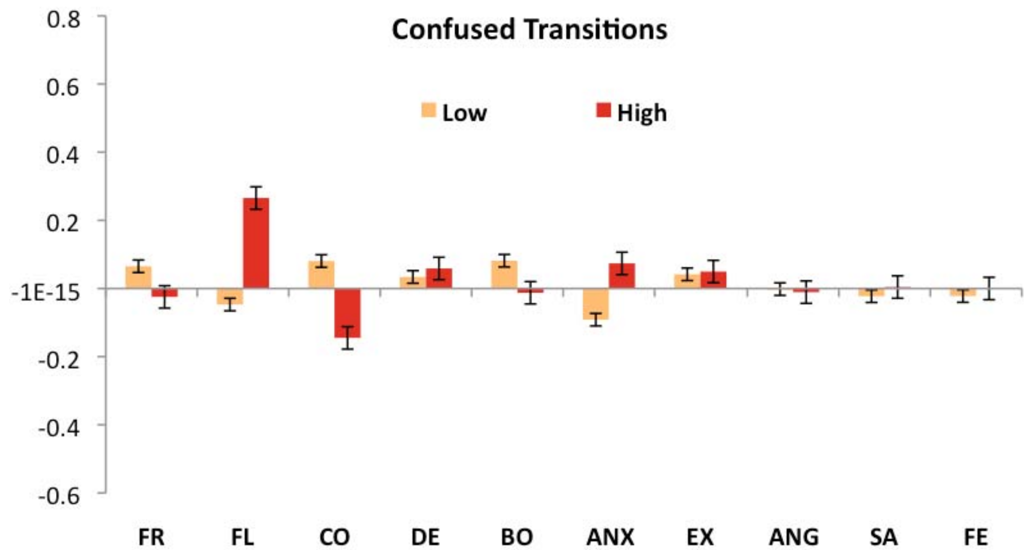

Figure 9. Transitions from confusion to **FR**ustration, **FL**ow, **CO**nfusion, **DE**light, **BO**redom, **ANX**iety, **EX**citement, **ANG**er, **SA**dness, and **FE**ar by level of reported sense of involvement/control (presence subscale)

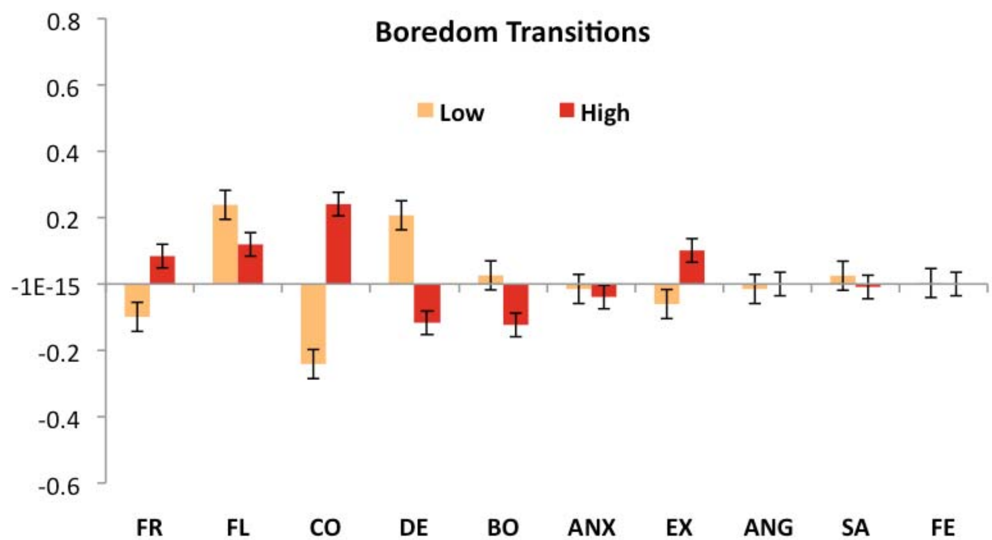

Figure 10. Transitions from confusion and boredom to **FR**ustration, **FL**ow, **CO**nfusion, **DE**light, **BO**redom, **ANX**iety, **EX**citement, **ANG**er, **SA**dness, and **FE**ar by level of reported sense of involvement/control (presence subscale)

For example, there are interesting differences in affective transitions when we consider student dominant-goal orientations. Mastery-oriented students are not likely to stay confused and are most likely to transition to a state of flow, a finding that suggests that mastery-oriented students are engaged or motivated by the cognitive disequilibrium associated with confusion. Being in a confused state is associated with a need to learn, and the CRYSTAL ISLAND environment supports mastery-oriented students' goal of acquiring knowledge. There is a chance that performance-oriented students may stay confused or transition to negative states such as frustration, boredom, or anxiety. Perhaps this is indicative of the fact that CRYSTAL ISLAND is guiding performance-oriented students into situations where they must master content to proceed, thus slowing progress and inadvertently decreasing perceived performance. Also, we notice that bored mastery-oriented students are not likely to remain bored and are more likely to transition to a state of flow or confusion. These emotional states are thought to be preferred for learning (Craig et al., 2004).

Lastly, there are interesting differences in likely transitions when we consider reported student presence as well. The participant population was broken into two groups around the population mean for the involvement/control subscale: *low* and *high*. Here we notice that students reporting high levels of involvement are not likely to stay in a state of confusion and are most likely to transition to a state of flow. On the other hand, students reporting lower levels of involvement in their experience were likely to stay confused or transition to other affective states, such as frustration or boredom. We notice a similar trend in transitions from a state of boredom. Students reporting high levels of involvement are not likely to stay bored and are more likely to become confused, excited, or enter a flow state. Students reporting lower levels of involvement are somewhat likely to stay bored, but are surprisingly more likely to transition to flow or delight. However, the occurrences of the vicious boredom cycles may in part be the cause for lower levels of reported involvement and control due to student disengagement.

Discussion

The analysis of affective state transitions in CRYSTAL ISLAND replicate findings by D'Mello et al. (2007) and Baker et al. (2007). For instance, the state of flow dominated self-reported affect. The dominance of the flow state has been reported in a number of affective studies with intelligent learning environments (Baker et al., 2007; Craig et al., 2004; D'Mello et al., 2007). Frustration and boredom were reported notably less frequently than in D'Mello et al.'s study (2007) and was comparably reported to frequencies found in Baker et al. (2007). Perhaps surprisingly, emotions found to be relevant to learning (boredom, confusion, delight, flow, and frustration) were more prevalent than the narrative affective states (anger, excitement, fear, and sadness) hypothesized to be relevant affective outcomes to experiencing the CRYSTAL ISLAND story.

Among the most likely transitions were transitions where *Next* = *Current*. This was true for the affective states of frustration, flow, confusion, delight, boredom, anxiety, excitement, and anger. This result also replicates the findings of D'Mello et al. (2007) and Baker et al. (2007). D'Mello termed these cycles vicious cycles for negative affective states (similar to Burleson's notion of "state of stuck" [2006]) and virtuous cycles when students are likely to stay in positive states (i.e., flow) (2007).

When we consider affective transitions where *Next* occurs at time $t + 1$ after an empathetic response from a CRYSTAL ISLAND character, we notice differences in the likely affective outcomes. For instance, if a student is in a frustrated state, parallel empathy is likely to elicit a transition in which the student stays frustrated. In contrast, reactive empathy is less likely than chance to prompt the same vicious cycle. Instead, reactive empathy tends to promote transitions to a confused state, which is known to have better correlations with learning (Craig et al., 2004).

When we consider likely transitions from the state of flow, we find that parallel empathy is likely to encourage students to enter a virtuous cycle and remain in the state of flow. Reactive empathy is less likely than chance to produce the flow state and is likely to promote an affective state transition to confusion. Since a flow state is an optimal state of experience (Csikszentmihalyi, 1990), it seems reasonable that reactive empathy cannot motivate students to enter a more engaged state.

Analyzing transition patterns from the state of boredom, we find that parallel empathy is likely to encourage a vicious cycle, whereas reactive empathy is less likely than chance to produce the same cycle. Instead, reactive empathy is most likely to transition to flow, with frustration slightly less likely than flow. In the future, when we can accurately predict when reactive empathy is likely to encourage flow as opposed to when it is likely to promote frustration, this diagnostic information can inform pedagogical agents' empathetic responses to alleviate student boredom and promote a state of flow.

Among the differences between personality traits, those relating to extroversion and conscientiousness are perhaps the most interesting. Highly extroverted individuals were more likely to report narrative-based emotions such as anger and delight and less likely to focus on learning or flow. Perhaps these individuals were more focused on the narrative aspects of the environment, such as interacting with characters, and consequently their attention was drawn away from learning tasks. Additionally, individuals who reported high levels of conscientiousness were less likely to report experiencing confusion. Generally, conscientious individuals are more likely to regulate their own behavior and perhaps this led them to focus on finding solutions to resolve their confusion. This notion was also supported by the increased likelihood of conscientious individuals to transition into flow and the very low likelihood that they remained confused.

Overall, the trend among affective frequencies shows that increased levels of performance orientation leads to reduced levels of flow and increased levels of anxiety. This is true when examining students' dominant orientation as well as their avoidance and approach subscales. This correlates well with understanding the approaches used by these two categories. Individuals who are mastery oriented are focused strongly on learning and may therefore be more likely to immerse themselves in learning-oriented activities in the environment. Similarly, as suggested by the rates of affective transitions, they may return more quickly to flow after experiences of other affective states. Conversely, performance-dominant students are focused on their measures of success. The higher level of anxiety reported by these students may be a direct result of concerns of performance. Because there is no objective measure of performance in the CRYSTAL ISLAND environment, performance-dominant students may become nervous over supposed comparison to others and opinions of the researcher present.

Interestingly, differences were found based on individual reports of presence. Students who reported higher levels of presence were more likely to have been anxious and less likely to have experienced frustration. Perhaps students who became frustrated disengaged themselves from the environment, resulting in lower levels of presence. Also, students who were highly engaged may have felt more salient responses to the narrative aspects of the environment. They may have become more concerned over the wellbeing of the characters and anxious over the outcome of the events. These differences are especially significant as they suggest that anxiety might be used to indicate measures of presence. Similarly, it appears that given an objective of maintaining presence, it would be highly important to avoid frustrating users.

Limitations

It seems likely that the results of this study are influenced by the virtual characters that interacted empathetically with participants. It is possible that the gender, narrative role, and pedagogical role of the characters may affect the likelihood of transitions in addition to the type of empathy. Another limitation is that affective states were solely collected from student self-reports. In contrast, both D'Mello et al. (2007) and Baker et al. (2007) used judged reports of affect in their transition analysis. In the study reported here, participants' faces were videotaped during interactions with the learning environment to permit future work that considers judged reports of affect with this dataset. Finally, to determine how broadly the results hold, the transitions that were found to be likely with this subject population need to be validated with other populations, such as the intended population of middle-school student users.

Conclusion

Given the central of role of affect and motivation in cognitive processes, it is becoming increasingly more important for intelligent tutoring systems to consider the affective experiences of students. The study reported here replicates the findings of studies conducted with AutoTutor (D'Mello et al., 2007) and The Incredible Machine simulation-based learning environment (Baker et al., 2007), including a demonstration of the prominence of the state of flow during learning. By extending our analysis to consider how affective transitions differ given empathetic character responses, the findings can inform the design of heuristics for pedagogical agents to determine when the use of empathy is likely to have desired outcomes and what type of empathy (parallel or reactive) would be best utilized. Such analysis can also inform the utility-induced models of empathy (McQuiggan, Robison, et al., 2008).

The results suggest two directions for future work. First, they call for investigation of what type of feedback pedagogical agents should consider when empathy does not promote desirable affective states for learning. For instance, reactive empathy was likely to encourage transitions to either flow or frustration. In instances where empathy promoted frustration, we should determine why empathy does not work and what type of system response would be more appropriate. Second, analysis of individual differences is necessary to determine the affective transitions common across a variety of demographics such as gender, but also across learning attributes such as efficacy, goal orientation, interest, and abilities to self-regulate both learning and affect.

Acknowledgments

The authors wish to thank the members of IntelliMedia Center for Intelligent Systems, Omer Sturlovich, and Pavel Turzo for use of their 3D model libraries, and Valve Software for access to the Source engine and SDK. This research was supported by the National Science Foundation under Grants REC-0632450, IIS-0757535, DRL-0822200, IIS-0812291, and CNS-0540523. 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 National Science Foundation.

References

André, E., & Mueller, M. (2003) Learning affective behavior. In *Proceedings of the 10th Intl. Conf. on Human-Computer Interaction* (pp. 512–516). Mahwah, NJ: Lawrence Erlbaum.

Baker, R., Corbett, A., Koedinger, K., & Wagner, A. (2004). Off-task behavior in the cognitive tutor classroom: When students "game the system." *Proceedings of ACM SIGCHI Conference on Human Factors in Computing Systems*, (pp. 383–390).

Baker, R., Rodrigo, M., & Xolocotzin, U. (2007). The dynamics of affective transitions in simulation problem-solving environments, *Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interactions* (pp. 666–677). Lisbon, Portugal.

Beal, C. & Lee, H. (2006). Creating a pedagogical model that uses student self reports of motivation and mood to adapt ITS instruction. In *Workshop on Motivation and Affect in Educational Software, in conjunction with the 12th Intl. Conf. on Artificial Intelligence in Education*.

Burleson, W. (2006). *Affective learning companions: Strategies for empathetic agents with real-time multimodal affective sensing to foster meta-cognitive and meta-affective approaches to learning, motivation, and perseverance*. Ph. D. thesis, Massachusetts Institute of Technology.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37−46.

Conati, C., & Mclaren, H. (2005). Data-driven refinement of a probabilistic model of user affect. In *Proceedings of the 10th Intl. Conf. on User Modeling* (pp. 40–49). New York: Springer-Verlag.

Craig, S. D., Graesser, A. C., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media, 29*(3), 241–250.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper and Row.

D'Mello, S., Taylor, R. S., & Graesser, A. (2007). Monitoring affective trajectories during complex learning. In *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (pp. 203−208). Austin, TX.

Davis, M. (1994). *Empathy: A social psychological approach*. Madison, WI: Brown & Benchmark Publishers.

de Vicente, A., & Pain, H. Informing the detection of the students' motivational state: An empirical study. In *Proceedings of the 6th International Conference on Intelligent Tutoring Systems* (pp. 933−943). New York: Springer-Verlag.

Ekman, P., & Friesen, W. (1978). *The facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Elliot, A., & McGregor, H. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501–519.

Forbes-Riley, K., & Litman, D. (2007). Investigating human tutor responses to student uncertainty for adaptive system development. In *Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction* (pp. 678−689). Lisbon, Portugal: Springer-Verlag.

Gratch, J., & Marsella, S. (2004). A domain-independent framework for modeling emotion. *Journal of Cognitive Systems Research, 5*(4), 269−306.

Johnson, L., & Rizzo, P. (2004). Politeness in tutoring dialogs: "Run the factory, that's what I'd do." In *Proceedings of the 7th Intl Conf. on Intelligent Tutoring Systems* (pp 67–76). New York: Springer-Verlag.

Kort, B., Reilly, R., & Picard, R. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy—building a learning companion. In *Proceedings of IEEE Intl. Conf. on Advanced Learning Technology: Issues, Achievements and Challenges* (pp. 43−48). Madison, WI: IEEE Computer Society.

Lester, J., Towns, S., & FitzGerald, P. (1999). Achieving affective impact: Visual emotive communication in lifelike pedagogical agents. *The International Journal of Artificial Intelligence in Education, 10*(3−4), 278−291.

Linnenbrink, E., & Pintrich, P. (2001). Multiple goals, multiple contexts: The dynamic interplay between personal goals and contextual goal stresses. In S. Volet & S. Jarvela (Eds.), *Motivation in learning contexts: Theoretical advances and methodological implications* (pp. 251–269). New York: Elsevier.

Lips, H. (2007). *Sex and Gender: An Introduction,* (6th ed.). McGraw-Hill.

McCrae, R., & Costa, P. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). New York: Guilford Press.

McQuiggan, S., Lee, S., & Lester, J. (2007). Early prediction of student frustration. In *Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction* (pp. 698−709). Lisbon, Portugal: Springer-Verlag.

McQuiggan, S., Mott, B., & Lester, J. (2008). Modeling self-efficacy in intelligent tutoring systems: An inductive approach. *User Modeling and User-Adapted Interaction, 18*(1–2), 81–123.

McQuiggan, S., Robison, J., Phillips, R., & Lester, J. (2008). Modeling parallel and reactive empathy in virtual agents: An inductive approach. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems* (pp. 167–174). Estoril, Portugal: International Foundation for Autonomous Agents and Multiagent Systems.

McQuiggan, S., Rowe, J., & Lester, J. (2008). The effects of empathetic virtual characters on presence in narrative-centered learning environments. In *Proceedings of the 2008 SIGCHI Conference on Human Factors in Computing Systems* (pp. 1511–1520). Florence, Italy: ACM.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Woods, S., Hall, L., & Zoll, C. (2005). Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence, 19*, 235−266.

Perry, N. (1998). Young children's self-regulated learning and the contexts that support it. *Journal of Educational Psychology, 90,* 715−729.

Porayska-Pomsta, K. & Pain, H. (2004). Providing cognitive and affective scaffolding through teaching strategies. In *Proceedings of the 7th International Conference on Intelligent Tutoring* Systems (pp. 77−86). New York: Springer-Verlag.

Prendinger, H., & Ishizuka, M. (2005). The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence, 19*, 267−285.

Rusting, C. (1998). Personality, mood, and cognitive processing of emotional information: Three conceptual frameworks. *Psychological Bulletin, 124*(2), 165−196.

Wang, N., Johnson, W. L., Mayer, R., Rizzo, P., Shaw, E., & Collins, H. (2008). The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human Computer Studies, 66*, 98−112.

Witmer, B., & Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225–240.

ISSN 1436-4522 (online) and 1176-3647 (print).

Yeh, H.-C., Yang, Y.-F., & Wong, W.-K. (2010). Interaction Chain Patterns of Online Text Construction with Lexical Cohesion. *Educational Technology & Society*, *13* (1), 55–68.

Interaction Chain Patterns of Online Text Construction with Lexical Cohesion

Hui-Chin Yeh[1], Yu-Fen Yang[1] and Wing-Kwong Wong[2]
[1]Graduate school of applied foreign languages and [2]Department of Electronic Engineering, National Yunlin University of Science & Technology, Douliu, Yunlin, Taiwan, R. O. C. // hyeh@yuntech.edu.tw // yangy@yuntech.edu.tw // wongwk@yntech.edu.tw

ABSTRACT

This study aims at arousing college students' metacognition in detecting lexical cohesion during online text construction as WordNet served as a lexical resource. A total of 83 students were requested to construct texts through sequences of actions identified as interaction chains in this study. Interaction chains are grouped and categorized as a meaningful entity in order to investigate the students' thinking process and behavior in general and to understand the interaction between the computer and the students in particular. From the interaction chains, it was found that some students revised incorrect sentences to correct ones. In making correct revision, they needed to assess incoming information, interpret and organize textual information, engage in thinking what they know, monitor their own meaning construction process, and take remedial actions to reach comprehension. The rate of correct sentence selection increased from 34.04% to 55.02% in three sequential text construction tasks. The recognition of lexical cohesion was found to be a determining factor for successful construction of a text.

Keywords

Interaction patterns; Lexical cohesion, Reading comprehension, Metacognition, Text construction

Introduction

Lexical cohesion has been widely defined as a property of text to connect sentences. According to Halliday and Hasan (1976), cohesive devices present a necessary semantic continuity between sentences for the purpose of interpreting and comprehending a text. The interpretation of each sentence in a text may depend on both the selection of vocabulary and its cohesive relations with other sentences. Halliday and Hasan further address that "typically, in any text, every sentence except the first exhibits some forms of cohesion with a preceding sentence, usually with the one immediately preceding" (p. 293). They also point out that "cohesive ties between sentences stand out more clearly because they are the only source of texture…..it is the inter-sentence cohesion that is significant, because that represents the variable aspect of cohesion, distinguishing one text from another" (p. 9). As such, lexical cohesion is derived from the selection of vocabulary items. Lexical cohesive ties, namely, pairs of cohesive words, provide a context for identifying the connections between the concepts embedded in a text. Hasan (1984) and Hoey (1991) suggested that 40 to 50 percent of cohesive ties in a text are composed of lexical cohesion. It is worth further investigation to what extent lexical cohesion can help students construct meaning in a text.

Many studies have affirmed the significance of the role that lexical cohesive ties play in the comprehension of a text (Bridge & Winograd, 1982; Chapman, 1982; Rogers, 1974; Staddord, 1991; Nunan, 1993, Nunan, 2004; McCarthy, 1991, Wang, 1998). Nunan (1993) stresses the fact that the ability to detect the cohesive relationships across sentence boundaries is significant for students to comprehend a text. When the sequence of sentences are scrambled or altered, the meaning of the text is surely distorted or even radically changed. According to the results of Bensoussan and Laufer's study (1984), the major reading difficulty that English as a Second Language (ESL) or English as a Foreign Language (EFL) college students encountered was their failure in recognizing the connections among the sentences in a text. Along this line of research, Chu, Swaffar, and Charney (2002) pointed out that most Taiwanese EFL students were found to be less aware of cohesive devices when reading English texts, as they occur less often in Chinese. In other words, Taiwanese EFL students rarely use cohesive devices for integrating textual information (Chen, 2003; Sharp, 2003). Their difficulty in identifying cohesive ties and finding out the relationships among these devices in a text lowers their English proficiency.

Some researchers suggest that learners should take an active role and obtain a metacognitive ability to manage their learning for better effectiveness (El-Hindi, 1997; Yang, 2002). Particularly, computer-assisted language learning (CALL) environment has been reported to have had a positive impact on learners' learning process, because the ease of text manipulation facilitates the revelation of cognitive processes (Dewitt, 1996; Forbes, 1996; Hirvela, 2004, 2005). Some studies also indicate that CALL environment can help explore and facilitate students' thinking and critical learning skills (Dreyer & Nel, 2003; Sinclair, Renshaw, & Taylor, 2004; Yeh & Lo, 2001; Yeh, 2003) because the computer "is a good tool to expand human cognitive development and knowledge construction" (Yeh, 2003, p. 613). In the past, it was extremely challenging to study either a learner's thinking process, except with a few labor-intensive and time-consuming methods: naturalistic observations, interviews, or think-aloud protocols (Schacter, Herl, Chung, Dennis, & O'Neil, 1999). When these methods were used, the learners' thinking and learning processes were often disrupted so that the data obtained might be distorted as well.

This study reports on our design of a text construction system, in which EFL college students actively construct and reconstruct text meanings among sentences. The constructing and reconstructing process also rely much on the use of metacognitive abilities perceived as formative assessment skills or the ability to "think about thinking" (Abromitis, 1994; Underwood, 1997; Kolić--Vehovec & Bajsanski, 2001). Metacognition involves active control over the cognitive processes engaged in learning" (Livingston, 1997, p. 1), and "active monitoring and consequent regulation and orchestration of cognitive process to achieve cognitive goals" (Flavell, 1976, p. 252). By exercising metacogntive skills, learners are in control of their learning process in assessing the new information they read, and retrieving the knowledge needed for understanding the written text (Abromitis, 1994).

For example, during reading, readers are monitoring their comprehension process that activates their prior knowledge, their failures of comprehension, and the strategies that can help them understand the text (Kolić-Vehovec & Bajsanski, 2001; Brown, 1985). They will in turn attempt to modify their reading process to match their purpose of understanding the text. In other words, they have to activate their metacognition by firstly acquiring the concepts from the text while adjusting the concepts in the light of subsequent information (Baker & Brown, 1984a, 1984b; Yang, 2002).

Metacognition refers to the consciousness of people when they are aware of monitoring and regulating their cognitive activities in the process of performing a cognitive task (Baker & Brown, 1984a, 1984b). They should be "monitoring ongoing activities to determine whether comprehension is occurring" and "taking corrective action" (Baker & Brown, 1984a, p. 354). When encountering problems, metacognitive learners can address the problems consciously, and act remedially. As learners "recognize mistakes and inconsistencies in texts and understand that they impair readers' comprehension" (Ruffman, 1996, p. 33), and "take remedial action" (Yang, 2006, p. 67) to resolve the inconsistencies, they are actually engaged in metacognition to improve their proficiency. The learner should be continuously encouraged to detect inconsistency and to make revision in order to achieve full comprehension.

Brown (1987) also specifically characterized four components in metacognition: planning, monitoring, evaluating, and revising. Planning refers to the deliberate activities that organize the entire learning process, including setting the goal, sequence, strategies, and expected time for learning. Monitoring involves the activities that moderate the current progress of learning. Evaluating one's own learning process involves an assessment of the current progress of the activity. This evaluation can assist learners to develop the necessary skills and strategies. Revising one's own learning process refers to the modifications of previous strategies related to goals and other possible learning approaches. These four components of metacognition all lead to the improvement of comprehension.

Based on our design of the computer-based text construction task, students' intensive engagement of metacognition can be expected. At first, students will use tutoring and trial practice to get familiar with both the lexical cohesion and the computer-based learning environment. This is the process of planning in this system for students have to be aware of the ultimate goal of the text construction task and of what they are expected to do next. On the way to insert sentences to construct a text, they will be required to first identify the lexical cohesive items, select types of lexical cohesion, and select inter-sentential relations since these are key elements to establish a paragraph. The selection of lexical cohesion types and sentence relation is defined as the process of monitoring in this system for students do not have to achieve the ultimate goal of inserting a correct sentence. Instead, students merely need to monitor if they are on the right track to approach the ultimate goal. The process of evaluating occurs when students have to evaluate four optional sentences and select a correct one to be related to the previous sentence in a text. This process differs from that of monitoring since correct sentence insertion is students' ultimate and only goal in text construction. After the submission of a sentence, students are allowed to revise their sentences. This is the process of revising in this system. These four components make it possible to investigate the students' comprehension and learning process while students are required to exercise their metacognitive ability to identify lexical cohesive ties and inter-sentential relations and select the correct sentence to complete the text construction task. The system is designed to serve as a cognitive tool for a student to assess incoming information, interpret and organize textual information, engage in thinking what he knows, monitor his own meaning construction process, and take remedial action to reach comprehension. When constructing a text and identifying lexical cohesion in the computer system, the student undergoes a dynamic and recursive process to reestablish the consistency of the text, while simultaneously searching the lexical cohesion between sentences to achieve comprehension. The system aims to enhance his metacognition and make his thinking process visible to both himself and the teacher.

The learning system used in this study is also a computer-supported interaction environment. With the Recording module, the system would trace and document each action a student takes in order to understand his thinking process. The student's metacognition could then be further analyzed as he makes revision while comprehending the text. The ability to determine "what has been done right or wrong," and "to take remedial action when comprehension failures occur" is a self-regulatory behavior (Yang, 2006, p. 67). Based on this definition, the sequences of actions students take in this study are identified as interaction chains. The interaction chains are grouped and categorized as a meaningful entity in order to investigate a student's thinking process and behavior in general and to understand the interaction between the computer and the student in particular. Several patterns are used to explain both the students' monitoring process and what facilitates or hinders their comprehension. The interaction chain patterns can be used by a teacher to examine the students' comprehension process. They also provide the details of how the students utilized the system's scaffolding to improve their metacognition.

Method

Participants

A total of 83 freshman students were recruited from two EFL classes in a technological university in central Taiwan. These two classes were from two different departments, English and Engineering. The distribution of students is shown in Table 1.

Table 1. The distribution of students

	Department	Number of participants	Gender
Class A	English	41	Male: 8 Female: 33
Class B	Engineering	42	Male: 37 Female: 5

A total of 83 students participated in this study. Class A was comprised of eight males and thirty-three females and Class B was comprised of thirty-seven males and five females.

Material

In order to provide appropriate English texts for the EFL college students to construct, the study adopted the following criteria for the selection of the texts for the text construction task. First, the readability level of the English texts should be controlled. The texts were all selected from College Reading Workshop (Malarcher, 2005). Second, the texts should be similar in length. Three texts were chosen: The Best Travel Bargains (Text 1), with 168 words, 8 sentences and 6 lexical cohesive pairs; Traditional Markets VS. Modern Markets (Text 2), with 206 words, 11 sentences and 15 lexical cohesive pairs, and Fat to Store or Fat to Burn (Text 3), with 188 words, 10 sentences and 13 lexical cohesive pairs.

System architecture

The system built for this study includes three modules: user interface, recording module, and feedback module. The relationships among these three modules are presented in Figure. 1. The teacher sets the objectives of the course, selects the appropriate texts and enters the texts into a database through the teacher interface. The recording module documents students' constructive behavior and process. The lexical cohesion and relations of sentences that are identified will be recorded in a database. The feedback module looks up WordNet, matches the lexical cohesive items, and provides candidate words back to students when they have a hard time identifying the lexical cohesion. These modules will be discussed in detail.

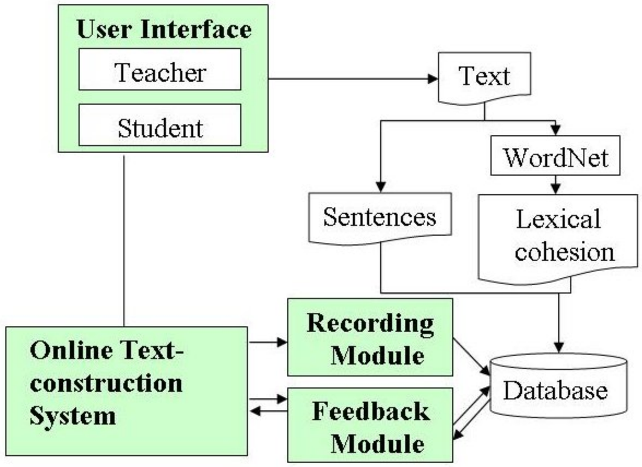

Figure 1. System architecture

User Interface

The user interface includes a teacher interface and a student interface. The teacher interface allows the teacher to manage a course, provide the texts to be constructed, and analyze the students' constructive process and behavior. The student interface is where a student undertakes text construction tasks.

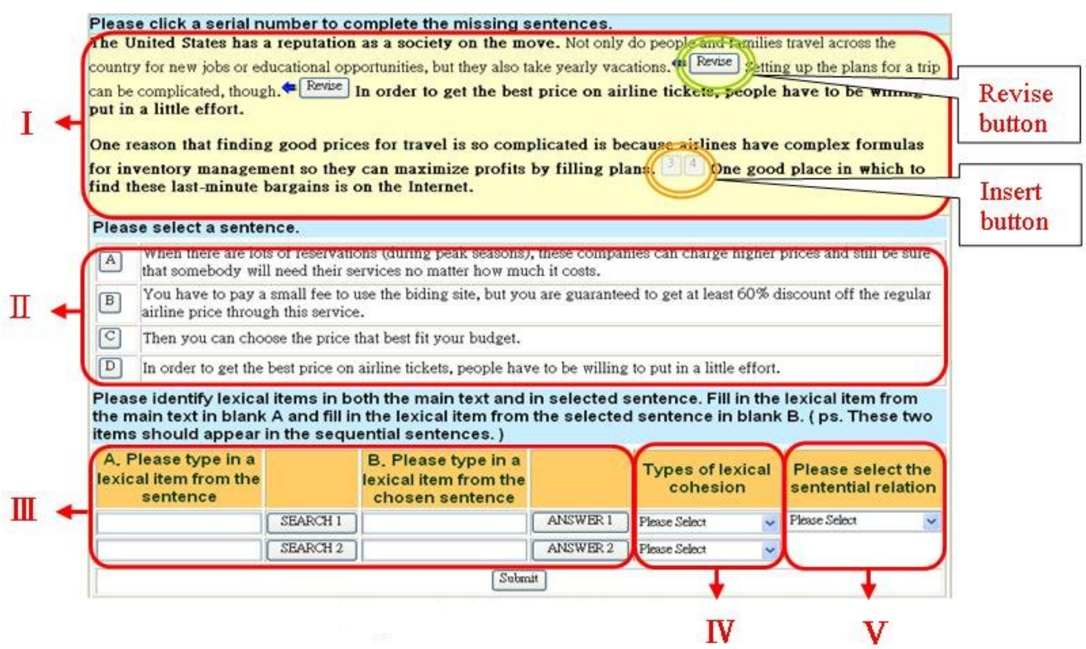

Figure 2. Student interface

As shown in Figure 2, the student interface is divided into five areas. The main text area (I) presents the text that is being constructed. An "insert" button tagged with a serial number represents a missing sentence that a student should identify based on the first and the last sentences (in boldface) of a paragraph. A "Revise" button allows the student to revise before final submission. Next, the area of multiple-choice items (Ⅱ) provides four sentences, tagged A, B, C, and D for the student to select from. Three distracting sentences were randomly picked from the rest of the text except the first and the last sentences. In the area of lexical cohesion (Ⅲ), students are asked to fill in two cohesive items from the current sentence and its preceding sentence. They can get hints of cohesive words by clicking the "Search" buttons or "Answer" buttons that are created based on the number of cohesive word-pairs embedded in the two sentences.

In the area of cohesion types (Ⅳ), students have to fill in the relations between the two cohesive items from the text fields A and B. The lexical cohesion types include: repetition, synonym, hypernym, hyponym, meronym. From the menu of inter-sentential relations, students have to choose the relation between the two sentences. The relations include: addition, clarification, comparison, contrast, example, location or spatial order, cause and effect, summary and time (Sharp, 2003).

Recording Module

The system uses a recording module to trace students' constructive process and behavior, which is presented as interaction chains. The teachers can analyze the interaction chains and identify the difficulties students encounter and the different performance among various proficiency groups. The analysis of interaction chains reveal whether the students make good use of the scaffolding in identifying lexical cohesions during text construction. The records are also helpful for the teacher to modify his instruction according to the demonstrated strengths and weaknesses of the students. The module uses some predicates to record students' behavior data (Table 2).

Table 2. The predicates for recording students' data

Predicates	Description
Search(W, S, t)	use WordNet search word W and type t at sentence S
Insert a sentence [n] (S)	insert a sentence s at sentence n
Identify cohesive words(w1, w2, t)	Identify cohesive words w1 and w2 and cohesive type t
Select relation(T)	Select inter-sentential relation T
Answer(S)	Provide a pair of lexical items as Answer for sentence S
Revise [n] (S, T)	Sentence n is revised from sentence S to sentence T

An example of how the recording module traces students' learning process of using WordNet to search for hypernyms of cost, inserting sentence 1, identifying lexical cohesion and selecting sentence relation is shown in Table 3.

Table 3. Records of a student's construction process

Explanation	Sequence of interaction chain
	1. Search ("cost", "The best way to get a cheap airline ticket used to be by reserving a ticket at least 21 days before you planned to travel.", "Hypernym")
1.Sentence:X	2. Insert a sentence: [The best way to get a cheap airline ticket used to be by reserving a ticket at least 21 days before you planned to travel.]
2.Cohesion:0/2	3. Identify cohesive words_1: [travel, travel, Repetition]
	4. Identify cohesive words_2: [plan, plan, Repetition]
3. Relation:O	5. Select relation: [Addition]

Feedback Module

Among the above lexical cohesion, repetition, synonym, hypernym, hyponym, and meronym are found most commonly in WordNet. WordNet 2.1 (Miller *et al*., 2005) is used in the current system to find these five types of lexical cohesion. It also assists the student to identify the relationships between given lexical cohesive items. WordNet is an online lexical database with a hierarchical structure, where a node is a synset (a set of synonyms) and is linked to other nodes with some relationship. WordNet is thus adopted in the current computer system as the knowledge source for the lexical semantic relationship utilized in matching cohesive words. In the past, WordNet has been utilized as a corpus to analyze discourse automatically. It has been rarely used in a learning systems to assist learners to comprehend a text.

In this study, a "SEARCH" button was designed to scaffold the student's text construction. The candidate lexical cohesive words generated from WordNet are shown in red color in both the main text and the multiple-choice items. For example, when the student enters the word "Brazil" in text field A and chooses hypernym as the type of lexical cohesion, the cohesive words "Brazil" and "country" found in WordNet will be highlighted in the text (Figures 3 and 4).

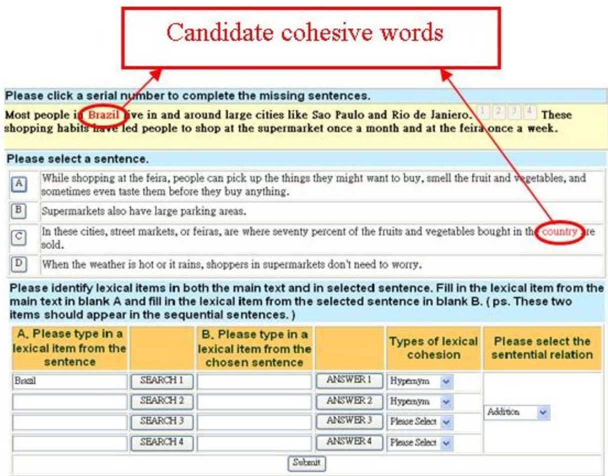

Figure 3. An example of using SEARCH button

Brazil, Federative Republic of Brazil, Brazil – (the largest Latin America country and the largest Portuguese speaking country in the would; located in the central and northeastern part of South America; worlds' leading coffee exporter)
INSTANCE OF
=> South American country, South American nation—(countries occupying the South American continent)
=> country, state, land—(the territory occupied by a nation; "he returned to the land of his birth ";"he visited several European countries")

Figure 4. Search results of "Brazil" (hypernym)

The student can click the "ANSWER" button to get a pair of lexical cohesive items, Brazil and country, when he encounters difficulty in figuring out the cohesive items (Figure 5), but he still needs to decide which type of lexical cohesion these two words form.

A. Please type in a lexical item from the sentence		B. Please type in a lexical item from the chosen sentence		Types of lexical cohesion
Brazil	SEARCH 1	country	ANSWER 1	Please Select

Figure 5. Example of using ANSWER button

After the student finishes constructing the text, the system will match the student's inserted sentences, cohesion, and sentential relation with the correct answers. The system takes a few steps to match the correct answers. First, when the student inserts sentence S, the system would match the answer to the target text. Next, after the student identifies lexical cohesive items (word 1 and word 2), the system would match the items against the lexical cohesion list. Finally, after the student selects the sentence relation R, the system would match the selected relation with the target answers.

Procedures of Data Collection

Eight three students were asked to construct three texts online at three different periods of time with one-month intervals. Before the instructor introduces the strategy, identifying lexical cohesion and sentence relation in text comprehension, the students undertook the text construction task for the first time. They were asked to complete the second text after strategy instruction for a month and to finish the final text after instruction for two months. The system recorded and traced the students' construction process. The text construction tasks were to identify the target sentence from the multiple choices, pairs of lexical devices and their cohesion type, and finally the sentential relation. The online computer system graded the students' text-construction by giving one score point to each correct answer of (1) sentences of the constructed text, (2) pair of cohesive words, (3) type of lexical cohesive ties, and (4) sentence relation.

Procedures of Data Analysis

The collected data were analyzed in terms of the students' text construction product and process. Text construction product refers to the students' overall scores. Text construction process includes the students' constructive process and behavior traced by the system. Revision and no-revision behaviors were categorized into different interaction chain patterns. The trace results revealed the students' difficulty in the constructive process and how they improved their understanding of the text.

Results

Product of Text Construction

The current study focused on understanding the students' cognitive product and process in constructing online texts. The data of 83 students' performances on Text 1 (The Best Travel Bargains), Text 2 (Traditional Markets VS. Modern Market), and Text 3 (Fat to Store or Fat to Burn) were analyzed. The overall reading performance is presented in Table 4. The total score for sentence selection in Text 1 was 4 and that for lexical cohesive ties was 6. The total score for sentence selection in text 2 was 7 and that for lexical cohesive ties was 17. The total score for sentence selection in text 3 was 6 and that for lexical cohesive ties was 13.

Table 4. Results of the Correct Sentence Selection

Participant	Text 1		Text 2		Text 3	
	Css* mean	SD	Css* mean	SD	Css* mean	SD
All students	1.36/4 (34.00%)	1.37	3.80/7 (54.28%)	2.11	3.30/6 (55.00%)	1.69
English major	2.12/4 (53.00%)	1.5	5.39/7 (77.00%)	2.01	3.70/6 (61.67%)	1.45
Engineering major	0.77/4 (19.25%)	1.24	2.77/7 (39.57%)	2.21	2.61/6 (43.5%)	1.93

*Css: Correct sentence selection

As shown in Table 4, almost all the students made progress in three sequential text construction tasks as the percentage of correct sentence selection increased from 34.04% to 55.02%. Specifically, the developmental progress was evident for the engineering students. Their percentage of correct sentence selection increased from 19.25% to 43.5%.

Table 5. Results of the Correct Lexical Cohesive Ties

Participant	Text 1		Text 2		Text 3	
	Clct** mean	SD	Clct** mean	SD	Clct** mean	SD
All students	1.16/6 (19.33%)	1.19	4.00/17 (23.53%)	2.30	4.96/13 (38.15%)	2.98
English major	1.64/6 (27.33%)	1.10	5.52/17 (32.47%)	2.11	5.55/13 (42.69%)	2.66
Engineering major	0.87/6 (14.50%)	1.28	3.26/17 (19.18%)	2.49	4.81/13 (37.00%)	3.3

**Clct: Correct lexical cohesive ties

This was also true for their performance in lexical cohesive ties. The engineering students progressed from 14.5% (Text1) to 37% (Text 3). The English-major students also made progress from 27.33% (Text1) to 42.69% (Text 3).

Process of Text Construction

According to the design of this system, the student must follow the steps to construct and reconstruct the text. One of the major functions in the system is the "Revise" button, which encourages the students to activate their metacognition. The steps of the students' revision and no-revision processes are shown in Figure 6. If students detected an inconsistency between the sentences, they would revise their thinking in order to reach new understanding. Otherwise, they did not undertake revision. For the step of "search cohesive items," students can choose to use or not to use WordNet to search for cohesive items before completing the task.

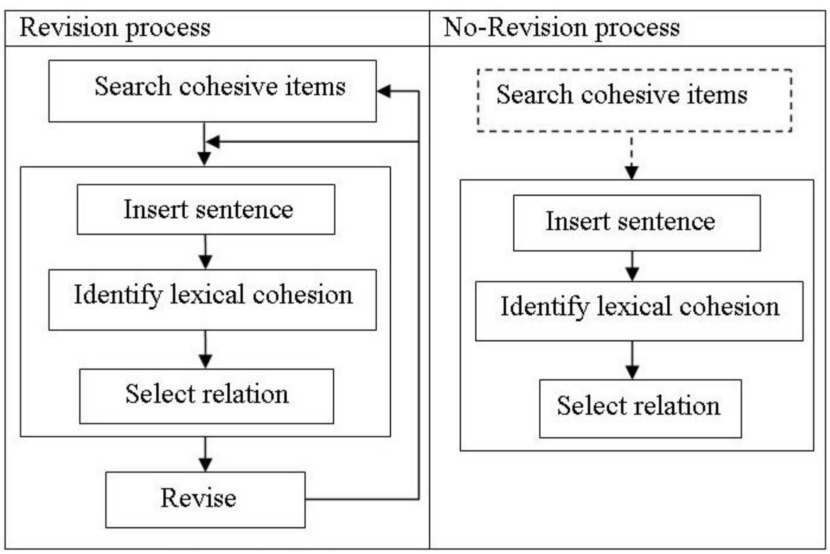

Figure 6. Revision and no-revision processes

Table 6. Three revision patterns

Pattern		Description
A	Incorrect→Correct	An incorrect sentence was revised correctly with correct identification of lexical items and their types of lexical cohesion.
B	Incorrect→Incorrect	An incorrect sentence was revised incorrectly with incorrect identification of lexical items and their types of lexical cohesion.
C	Correct→Incorrect	A correct sentence was revised with incorrect identification of lexical items and the types of lexical cohesion.

Patterns of Revision in the Text-construction Task

The students clicked on the "Revise" button to revise a sentence. The way a sentence could be revised can be grouped into three different interaction chain patterns (Table 6). In a sample text given below, the first and the last sentences of the paragraph were provided initially while sentence 1 and sentence 2 were inserted by a student. Sentence 1 was used to illustrate the three major revision patterns.

Sample text

One reason that finding good prices for travel is so complicated is because airlines have complex formulas for inventory management so they can maximize profits by filling plans. $\boxed{1}$ *When there are lots of reservations (during peak seasons), these companies can charge higher prices and still be sure that somebody will need their services no matter how much it costs.* $\boxed{2}$ *On the other hand, during the off-peak season, demand is low, so companies cut their prices to try and attract people who would normally not travel at that time.* One good place in which to find these last-minute bargains is on the Internet.

In Patten A, the students revised an incorrect sentence to the correct one by detecting the lexical items and the types of lexical cohesion. Figures 8 and 9 shows an example for Pattern A taken from the record of a student.

Sentence 1 **[Incorrect]**	**[13]** The best way to get a cheap airline ticket used to be by reserving a ticket at least 21 days before you planned to travel.		
Lexical cohesion	Word 1	Word 2	Type
	[14] travel	**[14]** travel	**[14]** Repetition
	[15] plan	**[15]** plan	**[15]** Repetition
Sentence relation	**[16]** Addition		

[17, 18]

Sentence 1 **[Correct]**	**[19]** When there are lots of reservations (during peak seasons), these companies can charge higher prices and still be sure that somebody will need their services no matter how much it costs.		
Lexical cohesion	Word 1	Word 2	Type
	[20] *price*	**[20]** *cost*	**[20]** *Hypernym*
Sentence relation	**[21]** Addition		

Figure 7. Interaction chain of Pattern A

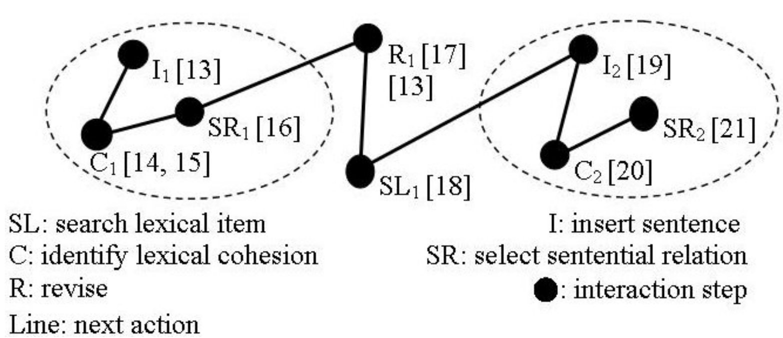

Figure 8. A graphical illustration of Figure 7

In Figure 7 the first three actions the student took for his text-construction activities were to insert sentence (action 13), identify lexical cohesion (actions 14, 15) and select sentence relation (action 16). These three actions constituted a meaningful entity (circled in an oval shape with a dotted line in Figure 9). Before final submission, he reread the sentence and searched for lexical items to confirm his ideas (action 18). Yet, he was aware of the inconsistency between the sentences. In the construction process, the student asked for help in his text comprehension and searching for the candidate lexical items. Thus, he revised both the sentence and the lexical cohesion correctly (actions 19-21) after the search.

As shown in the graphical illustration of Figure 8, before the student took remedial actions 17 and 18 to obtain the correct answers, he underwent a rereading process to reconsider his first answers. It is evident that the ability to identify the key lexical cohesion contributed to the student's comprehension of the text.

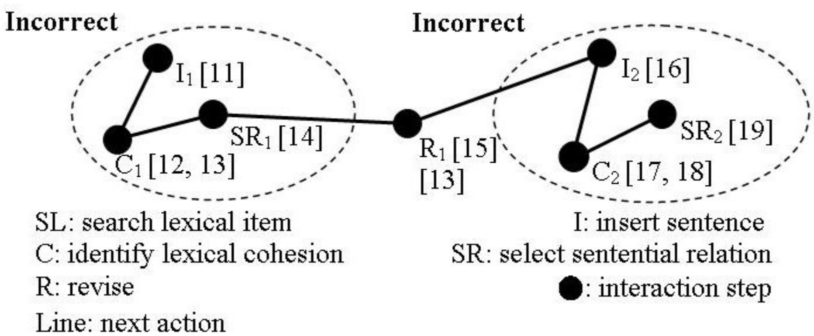

Figure 9. A graphical illustration of Pattern B

In Pattern B, the student revised the incorrect sentence to another incorrect one because he could not identify the lexical items and the types of lexical cohesion correctly (Figure 9). The student did not use the "Search" function. Though he detected the inconsistency between sentences, he did not get the correct sentence. He would have a better chance getting the correct answer if he used the system to help detect the correct lexical cohesion.

With regard to Pattern C, the student replaced the correct sentence with an incorrect one because he could not identify the lexical items and the types of lexical cohesion correctly no matter whether they revised or not (Figure 11). Though the student revised the sentence without asking for the assistance, the student did not fully understand the text or the lexical cohesion.

Table 7. Two no-revision patterns

D	Correct	A student selects the correct sentence and fills in correct lexical cohesion
E	Incorrect	Student selects the incorrect sentence and fills in incorrect lexical cohesion

Among the students who did no revision, Pattern D characterizes those who selected the correct sentence and filled in the correct cohesion while Pattern E characterizes those who selected the incorrect sentence and filled in the lexical cohesion incorrectly (Table 7).

The student in Pattern D selected the sentence and lexical cohesion correctly. The student did not undergo the complex reading process as those in the revision group. The student in Pattern E apparently did not detect the inconsistency between sentences. He neither asked for feedback nor revised the sentence. Only when the student can identify the inconsistency between sentences can they take remedial actions as reading strategies to fix the problem. In summary, five different patterns emerged to classify how students behaved in their learning process and what fostered or hindered their comprehension.

The relationship between metacognition and students' learning outcomes

As Trainin and Swanson (2005) mentioned, the employment of metacognitive strategies is positively linked to students' learning outcomes in academic environments. In this study, all participants' correctness rate in revision is shown in Table 8. It could be seen that students learned to make more revisions from text 1 to text 3. As they made more revisions, their correctness rate in selecting correct sentences of text construction also increased from 34.04 % to 55.02% (see Table 9). The result of the Pearson product-moment correlation coefficient between revision and correct sentence selection is shown in Table 10. The correctness rate of revision has a positive correlation with that of sentence selection. That is, the more the participants revised their sentences, the higher scores they obtained in text construction.

Table 8. Students' correctness rate in revision

	Text 1	Text 2	Text 3
Number of Students	83	83	83
Percentage of correct revisions	26 %	31.5 %	42.6%

Table 9. Students' correctness rate in inserting sentences

Participant	Text 1	Text 2	Text 3
	Css* mean	Css* mean	Css* mean
All students	1.36/4 (34.00%)	3.80/7 (54.28%)	3.30/6 (55.00%)
English major	2.12/4 (53.00%)	5.39/7 (77.00%)	3.70/6 (61.67%)
Engineering major	0.77/4 (19.25%)	2.77/7 (39.57%)	2.61/6 (43.50%)

*Css: Correct sentence selection

Table 10. The correlation between revision and correct sentence selection

	Text 1	Text 2	Text 3
N	PC	PC	PC
83	.70	.76	.81

*N: the number of participants
*PC: Pearson's correlation

All these consciousness-raising requests designed in the system aim to arouse students' metacognition so that they can improve their cognitive process. That is, in each text construction and reconstruction, the previous process served as a cognitive stimulus for the next process. The positive result might derive from whether students can actively take on revision and search for the clues, such as lexical cohesion, in WordNet to facilitate their comprehension. Without metacognition, students might remain in the original status of cognitive process for an extensive time without further improvement.

Conclusion

In this study, the text construction system served as a medium for the students to monitor and regulate their comprehension and for the teacher to understand the students' learning behavior and process. The system is designed to encourage the learners to use their metacognition by identifying the lexical cohesive ties across sentences and the types of lexical cohesion. Some conclusions can be drawn from both the students' text construction product and process. The students made progress in three sequential text construction tasks as the percentage of correct sentence selection increased from 34.04% to 55.02%. The improvement for their performance in lexical cohesive ties is also evident as the percent of correct lexical cohesive ties increased from 19.28% to 38.18%. In the process of text construction, five different interaction patterns emerged to characterize not only how the system supports the student's learning but also what processes the learners undergo while completing the required tasks. For the revision group, three different patterns described how students revised from incorrectness to correctness, from incorrectness to incorrectness, and from correctness to incorrectness. For the no-revision group, two different patterns characterized: (1) how the student selected a correct sentence and filled in the correct lexical cohesion and (2) how the student selected an incorrect sentence and filled in an incorrect lexical cohesion.

In the system, a Recording module traced and documented the student's learning process and behavior in general and their revision actions in particular. Interaction sequences presented in a chain manner illustrate the students' metacognitive process. Different patterns generated from the data informed the teacher of how the students activated their awareness to detect the inconsistency in sentences and how they took action to repair their understanding. Through the construction tasks, the students had to carefully access incoming information by reading the preceding sentence and the options from the multiple choices. The students also had to interpret and organize those incoming information by reading between lines. When they were requested to identify the lexical cohesive items in text, they were engaged in thinking what they knew and building the consistency between sentences. In this engagement, they would try to monitor their constructive process and detect the inconsistency. From the findings, we can also learn that the previous process of either taking on revision or asking for assistance from WordNet served as a cognitive stimulus for the next process. There is a strong correlation between revision and correct sentence selection. In other words, successful comprehension might result from whether students can actively take on revision and search for the clues, such as lexical cohesion, in WordNet to facilitate their comprehension. Without metacognition, students might remain lost in their cognitive process and would not be able to take any step to mend any comprehension breakdown.

Finally, if there were any comprehension failures, the students had the autonomy over taking remedial actions or not. The students' difficulties through the construction process in detecting lexical cohesion were examined and analyzed. It was found that the recognition of lexical cohesion was a determining factor in successful construction of a text. Whether or not the students could detect the lexical cohesion was a crucial factor in facilitating their learning process and comprehension. It was also found that more proficient students tended to make good use of the feedback to understand the connections in the text, and less proficient students seldom used the scaffolding provided by the system. Future study would focus on the different interaction chains generated from students at different proficiency levels in order to understand their text construction behavior and process, and explore the issues of how students at different levels of proficiency benefit from the current system.

Based on the analysis and generalization from the trace results, the teacher could assist the students in overcoming their difficulty not only in text construction per se but the lexical cohesion itself. Different interaction chain patterns can possibly allow the teacher to analyze the students' metacognitive process and understand what factors hinder their successful text construction. The trace results provided by the system would serve as a guide for the teacher to prepare for the instructional materials and effective teaching approaches. Their ability to identify correct lexical cohesion as well as their metacognition were critical for successful text construction and comprehension. It is hoped that this system can be used for a variety of course designs in additional to EFL courses because university courses make extensive use of academic materials written in English (Carrell, 1998).

The teacher can select appropriate texts for the students to construct their meanings through the text construction task. The students are encouraged to read across sentences instead of solely focusing on individual sentences in detecting lexical cohesion. While completing the tasks, the students need to activate their metacognition and detect the lexical cohesion across sentences. In future studies, we can generalize this approach to other types of cohesion mechanisms, such as anaphora. When students encounter any difficulty or confusion, they are encouraged to ask for feedback from the system to scaffold their comprehension. Then the scaffolding provided by the system can guide the students to achieve full reading comprehension.

Acknowledgement

This study was supported by National Science Council in the Republic of China, Taiwan (NSC 97-2410-H-224-017-MY2). The research grant made the continuation of this study possible.

References

Abromitis, B. (1994). The role of metacognition in reading comprehension: implications for instruction, *Literacy Research and Report Number 19*, Dekalb, MI: Curriculum and Instruction Reading Clinic, Northern Illinois University.

Baker, L., & Brown, A. L. (1984a). Cognitive monitoring in reading comprehension. In J. Flood (Ed.), *Understanding reading comprehension*, Newark, DE: International Reading Association, 21-44.

Baker, L., & Brown, A. L. (1984b). Metacognitive skills and reading. In P. D. Pearson (Ed.), *Handbook of reading research*, New York: Longman, 353–394.

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading, 7*, 15-32.

Bridge, C., & Winograd, P. (1982). Readers' awareness of cohesive relationships during cloze comprehension. *Journal of Reading Behavior, 14*, 299-312.

Brown, A. L. (1985). Metacognition: The development of selective attention strategies for learning form texts. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd Ed.), Newark, DE: International Reading Assoc, 501-526.

Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding*, Hillsdale, NJ: Lawrence Erlbaum, 65-116.

Carrell, P. L. (1998). Introduction: Interactive approaches to second language reading. In P. L. Carrell, J. Devine & D. Eskey (Eds.), *Interactive approaches to second language reading*, Cambridge: CUP, 73-100.

Chapman, J. (1982). A Study in Reading Development: A Comparison of the Ability of 8, 10 and 13 Year Old Children to Perceive Cohesion in Their School Texts. *Paper presented at the 19th Annual Conference of the United Kingdom Reading Association*, July 19-23, Newcastle-upon-Tyne, England.

Chen, Y. J. (2003). On-line metacognitive enhancement program for essay writing. In *Selected papers from the 12th International Symposium on English Teaching*, Taipei: The Crane Publishing, 338-347.

Chu, H. C. J., Swaffar, J., & Charney, D. H. (2002). Cultural representations of rhetorical conventions: The effects on reading recall. *TESOL Quarterly, 36* (4), 511-541.

Dewitt, S. L. (1996). The current nature of hypertext research in computers and composition studies: An historical perspective. *Computers and Composition*, *13*, 69-84.

Dreyer, C., & Nel, C. (2003). Teaching reading strategies and reading comprehension within a technology-enhanced learning environment. *System, 31* (3), 349-365.

El-Hindi, A. E. (1997). Connecting reading and writing: College learners' metacognitive awareness. *Journal of Developmental Education, 21* (2), 10-17.

Flavell, J. H. (1976). Metacognition aspects of problem-solving. In L. B. Resnick (Ed.), *The Nature of Intelligence*, Hillsdale, NJ: Erlbaum, 231-235.

Forbes, C. (1996). Cowriting, overwriting, and overriding in portfolio land online. *Computers and Composition, 13*, 195-205.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*, London: Longman.

Hirvela, A. (2004). *Connecting reading and writing in second language writing instruction*, Ann Arbor, MI: University of Michigan Press.

Hirvela, A. (2005). Computer-based reading and writing across the curriculum: Taiwan case studies of L2 writers. *Computers and Composition*, *22* (3), 337-356.

Hoey, M. (1991). *Patterns of lexis in text*, Oxford: Oxford University Press.

Kolić-Vehovec, S., & Bajsanski, I. (2001). *Children's metacognition as predictor of reading comprehension at different developmental levels*, retrieved May 1, 2009, from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED456422.

Livingston, J. A. (1996). Effects of metacognitive instruction on strategy use of college students, *Unpublished manuscript*, Buffalo, NY: State University of New York.

Malarcher, S. (2005). Family Planning Success Stories in Sub-Saharan Africa Technical Brief. *The INFO Project*, Baltimore, MD: Johns Hopkins Bloomberg School of Public Health Center for Communication Programs.

McCarthy, M. (1991). *Discourse analysis for language teachers*, Cambridge: Cambridge University Press.

Miller, G. A., Fellbaum, C., Tengi, R., Wakefield, P., Langone, H., & Haskell, B. R. (2005). *WordNet*, New Jersey: Princeton University.

Nunan, D. (1993). *Introducing discourse analysis*, London: Penguin.

Nunan, D. (2004). *Task-based language teaching*, Cambridge: Cambridge University Press.

Rogers, D. (1974). Which connectives? Signals to enhance comprehension. *Journal of Reading*, *17*, 462-466.

Ruffman, T. (1996). Higher order factors in comprehension disability: Processes and remediation. In C. Cornoldi & J. Oakhill (Eds.), *Reading Comprehension Difficulties*, Mahwah, NJ: Erlbaum, 33-67.

Schacter, J., Herl, H. E., Chung, G. K. W. K., Dennis, R. A., & O'Neil, H. F. Jr., (1999). Computer-based performance assessments: A solution to the narrow measurement and reporting of problem-solving. *Computer in Human Behaviors, 15* (3-4), 403–418.

Sinclair, K. J., Renshaw, C. E., & Taylor, H. A. (2004). Improving computer-assisted instruction in teaching higher-order skills. *Computers and Education*, *42* (2), 169-180.

Sharp, A. (2003). *Reading comprehension and text organization,* New York: Edwin Mellen.

Trainin, G., & Swanson, H. L. (2005). Cognition, metacognition, and achievement of college students with learning disabilities. *Learning Disability Quarterly, 28*, 261-272.

Underwood, T. (1997). On knowing what you know: metacognition and the act of reading. *Clearing Hourse, 71*, 77-80.

Wang, Y. F. (1998). Facilitating EFL reading by teaching text cohesive ties*. Proceedings of the 7th International Symposium on English Teaching,* Taipei: Crane, 855-866.

Yang, Y. F. (2002). Reassessing readers' comprehension monitoring. *Reading in a Foreign Language*, *14* (1), 18-42.

Yang, Y. F. (2006). An investigation of undergraduates' comprehension monitoring knowledge and strategy in EFL reading. *Journal of Science and Technology, 15* (1), 67-82.

Yeh, S. W. (2003). The effects of web page structure and metacognitive knowledge on EFL learners' comprehension in hypermedia CALL. *Proceedings of the 2003 International Conference on English Teaching and Learning in the Republic of China*, Taipei: Crane, 613-624.

Yeh, S. W., & Lo, J. J. (2001). Metacognitive cues: A technique for enhancing EFL reading in hypermedia. *Fu Jen Studies, 34*, 28-40.

Amory, A. (2010). Education Technology and Hidden Ideological Contradictions. *Educational Technology & Society*, *13* (1), 69–79.

Education Technology and Hidden Ideological Contradictions

Alan Amory
Department of Mathematics, Science, Technology and Computer Education, University of Johannesburg, Johannesburg 2006, South Africa // aamory@uj.ac.za

ABSTRACT

This article examined, thought a Cultural Historical Activity Theory lens, how immersive- or pervasive environments and pedagogical agents could more easily support social collaboration as foundation of contemporary learning theory. It is argued that the fundamentalism-liberationism contradiction (learn *from* versus learn *with* technology) is no longer justifiable as contemporary technology tools (pervasive/immersive environments and agent technology), the understanding of social networks, and recent neuro-science discoveries negate instructional design philosophies and innatist positions. The use of an activity lens allowed for identification of a number of educational technology design principles including explication of ideological positions, designs for contradictions, acceptance of a post-modern position, designs to overcome homophilic associations, and use of complex real-world learning activities.

Keywords

Collaboration, Cultural historical activity theory, Mirror-neurons, Education technology, Ideological contradiction

Introduction

This article is premised on the notion that there are hidden ideological contradictions in education technology as a field of practice and also of theory. These contradictions are embedded in the discourses of these fields and are present in, for example, positions about what constitutes learning, what constitutes technology itself, what constitutes theoretical positioning and also what constitutes the design for inquiries about these phenomena. It is important to understand educational ideological positions so to facilitate the development of appropriate education technology design and praxis.

Amory (2007) suggested that much of education technology replicates hegemonic practices that limit educational transformation, have little to do with contemporary learning practices and much more to do with fundamental and totalitarian ideologies of instruction. Similarly, Cohen (1987) argued that fundamentalists ideological beliefs embedded in technological products are incongruent with educational transformation. In addition, Gerardi (2006) suggested that advanced technologies are often tools of an authoritarian state leading to standardization of thought and social conformity. However, not all educational technology is driven by fundamentalist approaches to maintain the status quo. Referring to Cultural-Historical Activity Theory (CHAT), Stensenko (2005, p. 72) wrote:

> "[P]eople not only constantly transform and create their environment: they also create and constantly transform their very lives, consequently changing themselves in fundamental ways and, in the process, gaining self-knowledge. Therefore, human activity – material, practical, and always, by necessity, social collaborative processes aimed at transforming the world and human being themselves with the help of collectively created tools – is the basic form of life of people ."

Individual ideologies therefore operate, as McAllister (2004) suggested, within societal dialectical struggles reflecting the relationship between self and society and are a cultural artifact of mass-market post-modernism production.

This paper explores immersive- or pervasive environments and pedagogical agents that could more easily support social collaboration and individual transformation. In addition, the ideology-technology-learning triad is informed by developing an understanding of social networks and by recent neuro-scientific discoveries. The main argument of this paper is that personal and societal transformation can be cultivated through fostering social collaboration, designing complex learning activities that include contradictions, and make use of education technology in which embedded ideological positions are explicated. However, it is first necessary to position this exploration within an appropriate theoretical framework, as discussed in the next section.

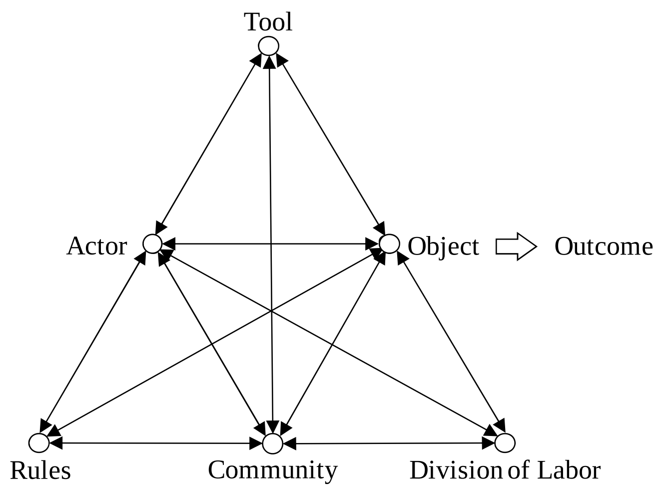

Figure 1. Activity system diagram (from Engeström, 1987)

Theoretical Framing

This paper made use of CHAT both as a way of understanding education technology (tool mediated construction) and as an analytical frame to identify pertinent design principles. A recent contribution to the legacy of Lev Vygotsky's cultural historical theory of learning is Engeström's (1987) broadening and amending of what Leontiev (1978) began with regard to theory of activity. Engeström's now often utilized model of an "activity system" is a helpful tool for work in information technology that includes social mediated activity (Fig. 1). In an activity system *Outcomes* result from *Actors* interrogating *Objects* by means of *Tools* (physical – pencils and technological artifacts; or psychological – signs and symbols). *Tools* mediate interactions through the activity context that includes *Community*, *Division of Labour* and associated *Rules* (Engeström, 2000, 2001; Barab, Evans & Baek, 2004; Roth & Lee, 2007). Internal *Contradictions* create instability and drive the *development of* and *change in* the system (Engeström, 2000, 2001). *Contradictions* are contextualized within dialectical approaches and change over time (Roth & Lee, 2007). *Objects*, as cultural entities, are the prime unit of analysis within an activity system (Engeström, 2001), embody communal social practices that transform and further develop during human activity (Stetsenko, 2005), and, in conjunction with motive, give the system coherence (Engeström, 2000). Socially created *Tools* are inseparable from the associated activity and are part of the purpose, relevance and value appropriated to them by the *Actor* (Robbins, 2005), and may become *Objects*, or *Outcomes*, of activity (Roth & Lee, 2007). It is implicitly understood that socially created *Tools* involved in cognitive mediation may not be ideologically neutral.

Education technological artifacts (*Tools*) contain specific ideological positions and often support fundamentalist world-views (Amory, 2007). The design, development and deployment of technological artifacts (*Objects*) are part of an activity system that includes complex communities (software engineers, interface designers, programmers, public relation officers) who use other technological tools and language (both human and machine) to create new artifacts (*Tools*). Therefore, the work of designers, developers and programmers is underpinned by specific ideologies, which are therefore part of the design of the *Objects*, *Outcomes* and *Tools*.

The paper is written from my own ideological position within a liberationist world-view that supports social freedom and equality. Knowledge is, therefore, viewed as a tool to support social reform (especially for the underprivileged and those facing discrimination) in order to develop individual potential.

Education technology that could inherently support social collaboration is explored in the next section. First a brief synopsis of the role in collaboration in learning is presented to frame the arguments. Thereafter Immersive and Pervasive Environments as Social Spaces, Pedagogical Agents as Social Negotiators, Social Networks and Social Neuro-scientific Discoveries are discussed and analyzed using CHAT.

Education Technology and Social Collaboration

Role of Collaboration in Learning

Collaboration, two or more people work together to realize a common objective, is an important component of contemporary learning theory (Vygotsky, 1933/1978; Piaget, 1977; Duffy and Cunningham, 1996), human development and transformation (Stetsenko, 2004), complex-games (Prensky, 2005), and learning environments that include authentic games (Gee, 2003; Reeves, Herrington & Oliver, 2004; Shaffer, 2005). In Massively Multiplayer Online (MMO) games collaboration played a greater role than did informational content, and such games should be viewed as social practice (Steinkuehler, 2004). In addition, Thomas (2005) found that the use of collaboration, reciprocal teaching, and sustained social and discursive practices in online role-playing environments supported problem solving and learning. However, Puntambekar (2006) showed that in an online graduate course there was little knowledge co-construction and collaboration that resulted in neither new ideas nor the inclusion of group ideas into individual responses. Barab (2003, p. 197) contended that we are yet to understand the difference "between a community of learners" (collaboration) and a "group of students learning collectively" (cooperation). While collaboration is one of the cornerstones of contemporary educational practices, the mechanisms and processes of collaboration in complex virtual worlds are not fully understood. Therefore, the relationships between learning and collaboration in immersive and pervasive systems, roles of pedagogical agents in social environments, social networks and neuro-scientific discoveries that support social interaction at the neurological level are explored to better understand the role of collaboration in cyber environments.

Immersive and Pervasive Environments as Social Spaces

Immersive and pervasive environments are cyberspaces in which individuals need to work together to solve complex problems that cannot be resolved individually. Here immersive and pervasive systems are discussed to discern the role of collaboration in environments specifically designed to solve complex problems.

A network environment that includes collective and political actions is an example of an immersive environment (McGonigal, 2003). In these spaces, which are different to pervasive ones (see below), McGonigal argued that participants need to solve complex real-world problems that include both physical and cyber world interactions, and require many different skills that include programming, translations, specific domain knowledge, and brute force. Young, Schrader and Zheng (2006) argued that in such environments interaction between players, via their avatars, is characterized by agent-environment, perception-action, and affordance-effectivity duals that led to active learning rather than retrieval of information from memory.

Pervasive environments, on the other hand, "extend the gaming experience out into the real world – be it on the streets, in the remote wilderness, or a living room. Players with mobile computing devices move through the world. The game player … experiences a game that is interwoven with the real world and is potentially available at any place and any time" (Benford, Magerkurth & Ljungstrand, 2005, p. 54). Pervasive systems consist of three core technologies that include: content made available through mobile phones, hand-held computers and wearable computers; wireless communication to support communication between participants; and sensing technology that captures participants' positioning. A pervasive learning environment includes a connected community of autonomous players where learning takes place at places meaningful to the learners (Thomas, 2006). It is interesting to note that the on-line community (in which the social interaction takes place) is more important in problem-solving than knowing where participants are located in the real-world (Nova, Girardin, Molinari & Dillenbourg, 2006). However, de Souza e Silva and Delacruz (2006) argued that the benefits of such environments relate to collaboration and bridging players in separate spaces via different forms of technology. While the creation of pervasive environments is not a trivial task (Lankoski, Heliö, Nummela, Lahti, Mäyrä & Ermi, 2004), such environments offer the most interesting artifacts through which technology can support interaction and collaborative problem-solving.

Summary and CHAT Analyses

In immersive and pervasive environments activity centers around complex real-world problems (the *Object*) that are situated in real and cyber space. Communication and work is supported by mobile and other "intelligent" devices that function across dissimilar environments to connect wet (human) and digital (cyber) components to support collaborative problem-solving (*Tools*). Immersive and pervasive systems include *Contradictions* related to Agent–Environment, Perception–Action and Affordance–Effectivity. The wet-digital dichotomy allows multiple social formations between other communities, device, people and situations (*Communities*) that allow various individual identities (*Division of Labour*) to work in self-forming and regulated communities.

The next section investigates the role of software-based pedagogical agents in cyber communities to bring to the fore important social interactions between synthetic characters (agents) and people working and playing in cyberspaces.

Pedagogical Agents as Social Negotiators

Wooldridge and Jennings (1995) argued that intelligent agents will be one of the most important computer systems in the development of complex software - autonomous agents with social ability, can react to their environment, are mobile within a network, benevolent, and rational. However, such a role of agents is not yet realized and this may be due to a number of factors. Artificial Intelligent systems do not include the necessary procedures to support appropriate interactions. In addition, the conceptualization of agents is limited to support current hegemonic instructional approaches to learning. Pedagogical agents could also play a more social role in digital environments. This is of particular interest to designers and developers of learning/instruction who wish to include suitable characters within socially constructed cyber environments. Therefore this section deals with the role of pedagogical agents in virtual environments and in learning.

Pedagogical agents play many different roles within virtual environments as: agents to support learning (mentors, tutors, guides, coaches, learning companion, teacher support, or motivational mediators) (Lester, Converse, Stone, Kahler & Barlow, 1997; Baylor, 2002; Payr, 2005); actors/coaches in interactive drama (Baylor, 2002; Chou, Chan & Lin, 2003); role-play actors that perform and are part of knowledge-based systems (André & Rist, 2001); as characters in content presentation, simulation, navigation, searching, management, and teacher support (Payr, 2005); as empathic companions (Prendinger & Ishizuka, 2005); and as intermediaries to support collaboration (Payr, 2005). The most fully described role of an agent in a learning environment is that of Chou et al (2003) who suggested that a learning companion should be a computer-simulated character that has human characteristics, plays a non-authoritative role, promotes social learning, and provides useful information. Hence, the role of pedagogical agents in constructivist learning environments is of particular interest.

Problem solving skills by middle school students improved when advice was received from pedagogical agents with visual and/or auditory modalities (Lester et al, 1997). Baylor (2002) found that the use of a constructivist agent changed teacher perceptions that led to the production of more constructivist teaching plans. However, it was not apparent if these changes were solely due to the presence of a pedagogical agent, the result of some other interaction such as the learning environment, or the design of the research investigation. Nevertheless, there are a number of research findings that point to the usefulness of pedagogical agents. Software agents that guided learners in ActiveWorlds resulted in participants making better use of advice and the production of more and better quality explanations (Holmes, 2007). However, this research was undertaken with small groups and might involve the repetition of information provided by agents in the answers.

Summary and CHAT Analysis

When pedagogical synthetic characters are used in learning environments they function at many levels including, in activity theory parlance, as *Object* (personalization of agent avatar); as *Tool* (acting as a knowledge agent, simulator, navigator, searching device, manager); as part of the *Community* (socially adept, pro-active, mobile, benevolent, veracious and rational); take on different roles – *Division of Labour* (mentor, tutor, guide, coach, companion, motivator, role-player, and emphatic companion); and function within a set of *Rules*.

The previous section discussed issues related to collaboration in cyberspaces and with intelligent objects (agents) without reference to specific theoretical constructs. The following section of the paper explores concepts related to social networks.

Social Networks

Kossinets and Watts (2006) suggested that social networks are important as they form part of information processing, social influence, and distributed search. The collective value of social networks is social capital forming part of the building and maintenance of democracy supported through information flows, norms of reciprocity, collective action, and broader identities and solidarity (Putnam, 1995). In this section, social capital, social network interactions, and Social Network Analysis are discussed.

Nahapiet and Ghoshal (1998) suggested that social capital includes three dimensions: structural (patterns of social interactions between actors), relational (the relationships between the actors), and cognitive (shared representations, interpretations, and systems of meaning among actors). The structural dimension contributes to a shared understanding and the cognitive dimension supports learning (Sorama, Katajamäki & Varamäki, 2004). Social capital appeared to work when individuals perceive that they have an advantage due to their position within a social structure (Burt, 2004). This is especially relevant to individuals who can act as brokers across groups to develop a shared understanding of similarities and differences between groups, or to provide critical syntheses. Such brokers, Burt argued, are more likely to propose ideas that are creative, openly discussed and are more likely to be accepted. However, the design, flow of information in, and dynamics of social networks is complex.

Kossinets and Watts (2006) suggested that the evolution of a network arises from the topology of the network itself and the organizational structure supporting the network. The movement of simple and complex knowledge within networks is different: simple knowledge diffused equally between distant and close actors while complex knowledge resisted diffusion (Sorenson, Rivkin & Fleming, 2006). The dynamics of social networks is bound by the focus, friendships and homophily (love of the same). In many instances the cognitive needs of a group can be of less importance when the foci of the interaction are highly valued (for example, the completion of a group task) (Feld, 1981). The degree to which two actors interact within a social network appeared to be directly related to the strength of their friendship ties (Granovetter, 1973). McPherson, Smith-Cain and Cook (2001) agreed that actors who share knowledge are more likely to interact. They showed that personal networks characterized by homophily and ethnicity, based on race, created the greatest divides; sex, age, religion and education strongly influenced network interactions and structures; and race and sex homophily appeared to be both a baseline and an inbreeding phenomenon. Punishment-reward systems were not needed in networks of actors that are alike (Durrett & Levin, 2005). In addition, altruism, jealousy and fairness appeared to be the more important reward when playing a dictator computer game (Andreoni & Miller, 2002). However, homophily attitudes can be less important in certain circumstances. Yuan and Gay (2006) found that in computer mediated distributed networks, team members are more sensitive to location than to gender or race homophily, and that social capital significantly influenced performance of both actors from different or within groups. This would be particularly true in cyber communities as actor representation through the use of avatars would allow the reconstruction of identities not based on reality. While the building block of democracy is social capital, the relationships between actors and cognitive shared meaning making dimensions that interact through altruism, or cooperation, are also important and may be more important in task-based social networks.

Social Network Analysis allowed for the visualization of social networks to made explicit the social capital to an actor (indegree – actors build on each other's notes – collaborative writing; outdegree – the number of specific notes a specific actor users; betweeness – whether an actor is a broker of information) (Sha & Aalst, 2003); allowed actors to visualize their own status in a network (Lockerd & Selker, 1999); improved performance (Cho, Gay, Davidson & Ingraffea, 2007); and permitted actors to serendipitously bump into each other thereby extending the social network (Farnham, Kelly, Portnoy & Schwartz, 2004). Therefore, the use of SNA in co-operative learning environments would support the development of social ties within learning communities.

Summary and CHAT Analysis

Social networks are used to develop social capital within democratic communities that are based on altruism and social justice (*Object*). The social network *Community* is multicultural and consists of individuals who are friends, part of a learning network and function as *Actor* and/or broker (*Division of Labour*). Social networks require *Tools* to support collaboration and feedback, and to visualize through the use of SNA individual position within the community (*Tools*). *Rules* of social networks are complex but support information flows, reciprocity, collective actions, broader identities, solidarity, and homophily. Social networks include altruistic–co-operative *Contradictions*.

Recent neurological discoveries using non-invasive technologies to view brain functions offer the first clue to a neurological social dance that appears to influence all aspects of our development and life. In the next part of the paper some aspects of this research are presented to provide evidence to support the role of social interaction in learning.

Social Neuro-scientific Discoveries

The sociological neuro-scientific advances that bind us socially together include the discovery of *spindle cells* and *mirror neurons*. Spindle cells are the most plentiful in humans compared to other mammals and appear to guide snap social decisions. Mirror neurons are involved in sensing the movement that others make to rapidly prepare us to respond to such movements and are involved in language and culture development (Goleman, 2006). Neuro-scientists are now, for the first time due to the development of advanced imaging technology, able to show the social aspects of our brain functions. The following discussion includes an introduction to the phenomenon of mirror neurons, and the role of mirror neurons in language development and collaboration.

Mirror neurons are scattered throughout key parts of the human brain – the promoter cortex and centers for language, empathy and pain – and fire not only as we perform a certain action but also when we watch someone else perform that action (Perrett, Rolls & Caan, 1982; Montgomery, Isenberg & Haxby, 2007). The mirror-neuron system appears to be tightly coupled with action understanding and imitation learning argued to be the basis of human culture (Rizzolatti & Craighero, 2004); involved in the evolution of language, music, art, tool-making and empathy (Azar, 2005; Oztop, Kawato, Arbib, 2006); and appeared to be an essential cognitive skill involved in social groups (Erlhagen, Muskovsky & Bicho, 2006) and emotional awareness (Parr, Waller & Fugate, 2005). However, the link between mirror neurons and language appears to be noteworthy.

Fogassi and Ferrari (2004) argued that gestural communication is the predecessor of human speech and Rizzolatti and Craighero (2004) illustrated a direct link between hand gestures, mouth gestures and the oro-laryngeal movement used in speech production. These works supported the proposal by Rizzolatti and Arbib (1998) that the mirror-neuron system is the neuro-physiological mechanism of language evolution. Bickerton (2007) was critical of these positions and argued that mirror neurons “could hardly be innately programmed to respond to action sequences that nobody has yet produced” and the theory cannot “shed any light on how symbols originated or how syntax originated” (p. 523). However, Arbib (2005) provided a detailed description of the role of mirror-neurons in the development of human language that also included the development of protosign, protospeech and evolution of the brain and body.

Goleman’s (2006) argument that “the major function of the social brain – interaction synchrony, the types of empathy, social cognition, interactive skills, and concern for others – all suggest strands of social intelligence” (p. 329) is built on works such as that of Gallese, Keysers and Rizzolatti (2004) who discussed how mirror neurons provide the first unifying theory of social cognition that afford us with insight into the minds of others. Rizzolatti and Craighero (2004) suggested that mirror neuron development played a role in the development of altruism and allowed individuals to understand the intensions, meanings and emotions of others. Gallese (2007) argued that the premotor system (part of the mirror neuron system) is involved in the mastery of hierarchical structure of language (one of the most important components involved in social cognition) and abstract inference. Therefore, the “circuitry that controls how we move our bodies and enables our understanding of the action of others can, in principle, also structure language and abstract thought” (Gallese, 2007, p. 666). Uddin, Iacoboni, Lange and Keenan (2007) argued that social interactions are dependent on self- and other-representations and that the mirror neuron system supports the physical other-to-self mapping. In addition, the discovery of a human auditory mirror system associated with the left hemispheric temporo-parieto-premotor circuit may be involved in language evolution and was more strongly activated in individuals that scored higher on an empathy scale (Gazzola, Aziz-Zadeh & Keysers, 2006) reinforced the role of language in social networks.

Chernigovskaya (2007) wrote that the mirror neuron hypothesis is of importance “both for explaining the organization of language functions … and for learning in general, as it allows links to be made between the *agens* (who does the action), the *patiens* (the object of the action), and the *instrument* (the means or tool)" – or in Vygotskian terms: the Actor, Object and Mediating Artifact triad. The neurological support for collaboration directly challenges a number of positions including:

- The reductionist cognitivist's agenda emphasizing value-free information processing in individual and isolated minds (Barab, Evans & Baek, 2004; Stetsenko, 2005; Vianna & Stetsenko, 2006);
- Skinner's operant conditioning model previously critiqued and rejected by Chomsky (1967);
- Instructional design, as proposed by Gagné and Merrill, based on behaviorism as frame through which concepts are treated as distinct learning outcomes (objects of instruction) rather than cognitive tools for representing ideas (Jonassen, 2006); and
- Chomsky's innatist "universal grammar" of language development (Holden, 2004).

The acquisition metaphor to describe learning is therefore inappropriate. The associated reductionist approaches are "directly affiliated with positivist, non-dialectical and ultimately conservative approaches in education" (Vianna & Stetsenko, 2006, p. 82) and drive the constant re-invention of education practice that is not substantially different from past practices. The dialectic struggle between perpetuating the past (fundamentalist) and the constant creation and transformation of oneself and thereby the world (liberationist) is unsustainable: fundamentalism offers little except that it drives the neo-liberal agenda that has little to do with learning and everything with maintenance of the past.

Summary and CHAT Analysis

Neuro-scientific discoveries support the notion that learning is situated in the cognition of a social mind. The development of language mediated action (*Tools)* in *Actors* is socially constructed thereby supporting the Actor, Object and Mediating Artifact triad. In addition, Engeström's (2001) notion that capitalism, and thus neo-liberalism, is the primary contradiction within an activity system is part of an educational dialectical struggle to rid the world of instructional design positions that are contrary to the social construction of knowledge.

The final section of the paper explores education technology in relation to CHAT both as argument and as tool of analyses. Thereafter design elements pertinent to the development of education technology artifacts and to the development of social interactions are presented.

Reflection, Documentation and Development of Design Principles

Amory (2007) argued that the design, development, integration and use of classroom technologies support current hegemonic fundamentalist positions maintained through observation and control systems and include Reusable Learning Objects (based on totalitarian ideologies of instruction), Learning Management Systems (which include information redistribution, observation and monitoring), blended learning (the inclusion of technological tools into existing courses with no pedagogical change perpetuating the past), and education games (ideologically suspect simulations based on model-using rather than model-building approaches). However, learning systems and artifacts built to support collaboration (immersive and pervasive environments, and pedagogical agents), as argued here, offer opportunities to create tools to support transformative activity systems and foster liberationist approaches. Stetsenko (2004) suggested that socially constructed tools can overcome the constraints of nature and the environment. This is true when the tools originate as part of a social framework and are not the result of other ideologies. While technological artifacts may include suspect ideological positions, such artifacts can foster social construction when the embedded ideology is explicitly declared, or when the embedded ideological position is used to create a contradiction that could be labeled the 'fundamentalism-liberationism contradiction'. This contradiction is explored in the next section.

Activity systems change and develop over historical time (Roth & Lee, 2007). Also, Vygotsky's ideas are a direct consequence of his historical context (theories were developed in opposition to Freud's psycho-analysis and behaviourist approaches) and are part of a social project that is both a product and an instrument (Stetsenko, 2004). Similarly, educational tools are products and instruments that are part of our globalized neo-liberal world. Klein (2007) shows that neo-liberal economic policies of privatization, free trade, decreased social spending and privatization of government enterprises disenfranchise specific race groups, perpetuate gender exclusivity and support fundamentalist religious belief systems. Education systems are not immune from neo-liberalisation. Neo-managerialism – the use of a market-driven educational system – makes use of neo-conservative standards, traditional hierarchies of class and race, accountability, national curriculum and national assessment policies (Amory, 2008). Learning is therefore driven by and consumes neo-liberalization and learning technologies often promote Fordist assembly line production of learning materials, and teaching and learning.

While, on the one hand, education technologies are often driven by neo-liberals to support a globalized economy that protect fundamentalist hegemonies that perpetuate and support race, gender and religious homophily within social networks. Technology, on the other hand, could support the liberalization of education practices and include immersive and pervasive system, and pedagogical agents that are designed to support collaboration. Activity systems that can lead to the production of *Tools* to support the "collaborative processes of material production of social life – human object-related activity" (Stetsenko, 2005) need to include the collaboration of the future *Actors* in the development process (*Object* activity) and make explicit the fundamentalism-liberationism contradiction. These *Actors* subsequently use the *Tool* to transform both themselves and the world around them in ways that support social reform to support the discriminated and underprivileged in order to develop fully individual potentials.

The final section of the essay uses the previously highlighted CHAT components to bring to the fore the fundamentalism-liberationism contradiction and the software design concepts applicable to the designed technological artifacts to support educational transformations. Design principles identified through CHAT analyses are discussed to make concrete design for future educational technologies. Technological artifacts and social interactions are discussed separately.

Analyses of design, development and use of immersive and pervasive social spaces and pedagogical agents as social negotiators allowed for the identification of a number of concepts:

- Ideological positions of software designers and engineers are integral to any activity system and directly influence the construction and development of the artifact, or *Object*;
- Production of technological artifacts minimize inherent *Contradictions* to improve usability and thereby decreases learning opportunities; and
- Constructed artifacts function as either an *Object* (for example, the deconstruction of an artifact) or as a *Tool*.

With respect to *social interactions* numerous *Actors* involved in software design build and participate in social networks. Activity systems of complex social networks are more complex than those associated with technological artifacts, and include:

- *Actors* playing different roles, for example, to facilitate *Object* development or complex real-world problem solving;
- Task-based *Outcomes* allowing *Actors* to overcome homophilic associations and thereby work together to reach common goals;
- Inherent *Contradictions* supporting disruptions that challenging preconceived notions leading to new understandings; and
- Access into activity systems using technological artifacts is normally via the *Object* or *Tool* component of the system; the *Actor* space is the primary access point is the design of technological systems.

The theory and practical use of education technology needs to acknowledge the fundamentalism-liberationism contradiction. Consequently, technological artifacts as *Objects* perpetuate past practices but when functioning as *Tool* could mediate learning. Second, education technology design needs to include *Contradictions* that challenge existing paradigms and allow for disruption, and therefore learning. Third, learning with technology needs to support a post-modern view that there may be more than one correct solution to a given/existing problem. Lastly, designs for learning technology need to allow for multiple voices to work together to solve complex real-world problems that includes both the digital and wet worlds in order to overcome the powerful race-gender-belief homophily that often dominates social network.

Conclusions

This paper argued that it is not longer tenable to consider the fundamentalist learn *from* technology position. Contemporary technology tools, the functioning of social networks and the findings of neuro-science challenge instructional design and innatist positions. The CHAT lens brings to the fore a number of principles that need to be considered in the design, development and use of technology for teaching and learning and include: explication of ideological positions, design for contradictions, acceptance of a post-modern position, power of social networks and the use of complex-real world learning tasks to overcome homophilic attractions. The development of such systems will be difficult and requires concerted efforts to topple the predominant use of techno-reductionist tools more associated with content and user management than with the transforming people and the world through human actions supported by socially constructed tools.

Acknowledgements

Special thanks to Elizabeth Henning, Duan van der Westhuizen, Geoffrey Lautenbach, Kathy Morgan, Gadija Petker and Nadine Petersen for their critical comments and encouragement.

References

Amory, A. (2007). It's not about the tool, it's about the ideology. *South African Journal of Higher Education*, *22* (6), 655–671.

Amory, A. (2008). Academic administration as the puppet masters (Sicelankobe). The consequences. *South African Journal of Higher Education*, *22* (3), 498-514.

André, E., & Rist, T. (2001). Presenting through performing: on the use of multiple life like characters in knowledge-based presentation systems. *Knowledge-Based Systems*, *14* (1–2), 3–13.

Andreoni, J., & Miller, J. (2002). Giving according to Garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70* (2), 737–753.

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, *28* (02), 105–124.

Azar, B. (2005). How mimicry begat culture: Researchers from varied disciplines look to mirror neurons to explain many aspects of human evolution. *Monitor on Psychology*, *36* (9), 54.

Barab, S. A. (2003). An introduction to the special issue: Designing for virtual communities in the service of learning. *The Information Society*, *19* (3), 197–201.

Barab, S. A., Evans, M. A., & Baek, E. O. (2004). Activity theory as a lens for characterizing the participatory unit. In D.H. Jonassen (Ed.), *Handbook of research on educational communities and technology*, Washington, DC: AECT, 199–214.

Baylor, A. L. (2002). Expanding preservice teachers' metacognitive awareness of instructional planning through pedagogical agents. *Educational Technology Research and Development*, *50* (2), 5–22.

Benford, S., Magerkurth, C., & Ljungstrand, P. (2005). Bridging the physical and digital in pervasive gaming. *Communication of the ACM*, *48* (3), 54–57.

Bickerton, D. (2007). Language evolution: A brief guide for linguists. *Lingua*, *117* (3), 510–526.

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110* (2), 349–99.

Chernigovskaya, T. V. (2007). The mirror brain, concepts, and language: The price of anthropogenesis. *Neuroscience and Behavioral Physiology*, *37* (3), 293–302.

Chou, C., Chan, T., & Lin, C. (2003). Redefining the learning companion: the past, present, and the future of educational agents. *Computers & Education*, *40*, 255–269.

Cho, H., Gay, G., Davidson, B., & Ingraffea, A. (2007). Social networks, communication styles, and learning performance in a CSCL community. *Computers & Education*, *49* (2), 309–329.

Chomsky, N. (1967). Review of Skinner's Verbal Behavior. *In Leon A. Jakobovits &Murray S. Miron (Eds.), Readings in the Psychology of Language*, Prentice-Hall, 142-143.

Cohen, D. K. (1987). Educational technology, policy and practice. *Educational Evaluation and Policy Analysis*, *9* (2), 153–170.

Durrett, R., & Levin, S. A. (2005). Can stable social groups be maintained by homophilous imitation alone. *Journal of Economic Behavior and Organization*, *57* (3), 267–286.

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In D. Johansson (Ed.), *Handbook of Research for Educational Communications and Technology*, New York: Simon & Shuster Macmillan, 170–198.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*, Helsinki: Orienta-Konsultit.

Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*, *43* (7), 960–974.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, *14* (1), 133–156.

Erlhagen, W., Muskovskiy, A., & Bicho, E. (2006). A dynamic model for action understanding and goal-directed imitation. *Brain Research*, *1083* (1), 174–188.

Farnham, S., Kelly, S. U., Portnoy, W., & Schwartz, J. L. K. (2004). Wallop: Designing social software for co-located social networks. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, Los Alamitos, CA: IEEE Computer Society, 107-116.

Feld, S. L. (1981). The focused organization of social ties. *American Journal of Sociology*, *86* (5), 1015–1035.

Fogassi, L., & Ferrari, P. F. (2004). Mirror neurons, gestures and language evolution. *Interaction Studies*, *5* (3), 345–363.

Gallese, V. (2007). Before and below 'theory of mind': Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362* (1480), 659–669.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8* (9), 396–403.

Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current Biology*, *16* (18), 1824–1829.

Gee, J. (2003). *What video games have to teach us about learning and literacy*, New York: Palgrave MacMillan.

Gerardi, S. (2006). Some implications of modern technology: Revisited. *The Social Science Journal*, *43*, 293–295.

Goleman, D. (2006). *Social intelligence: The new science of human relationships*, London: Hutchinson.

Granovetter, M. S. (1973). The strength of weak ties. *The American Journal of Sociology*, *78* (6), 1360–1380.

Holden, C. (2004). Oldest beads suggest early symbolic behavior. *Science*, *304* (5669), 369.

Holmes, J. (2007). Designing agents to support learning by explaining. *Computers & Education*, *48* (4), 523–547.

Jonassen, D. H. (2006). On the role of concepts in learning and instructional design. *Educational Technology Research and Development*, *54* (2), 177–196.

Klein, N. (2007). *The shock doctrine: The rise of disaster capitalism*, New York: Metropolitan.

Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science*, *311*, 88–90.

Lankoski, P., Heliö, S., Nummela, J., Lahti, J., Mäyrä, F., & Ermi, L. (2004). A case study in pervasive game design: the songs of north. *Proceedings of the 3rd Nordic conference on Human-computer interaction*, New York: ACM, 413–416.

Lester, J. C., Converse, S. A., Stone, B. A., Kahler, S. E., & Barlow, S. (1997). Animated pedagogical agents and problem-solving effectiveness: A large-scale empirical evaluation. *Proceedings of the SIGCHI Conference*, New York: ACM, 359–366.

Leontiev, A. N. (1978). *Activity, personality, and consciousness*, Englewoods Cliffs: Prentice-Hall.

Lockerd, A., & Selker, T. (1999). DriftCatcher: Enhancing social networks through email. *Proceedings of the 12th International Sunbelt Social Network Conference*, Retrieved May 1, 2009, from http://pubs.media.mit.edu/pubs/papers/sunbelt_paper.pdf.

McAllister, K. (2004). *Game work. Language, power and computer game culture*, Tuscaloosa: The University of Alabama.

McGonigal, J. (2003). 'This is not a game': Immersive aesthetics and collective play. *Proceedings of the Digital Arts and Culture Conference*, Retrieved June 9, 2009, from http://hypertext.rmit.edu.au/dac/papers/McGonigal.pdf.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27* (1), 415–444.

Montgomery, K. J., Isenberg, N., & Haxby, J. V. (2007). Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Social Cognitive and Affective Neuroscience*, *2* (2), 114.

Nahapiet, J., & Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. *The Academy of Management Review*, *23* (2), 242-266.

Nova, N., Girardin, F., Molinari, G., & Dillenbourg, P. (2006). The underwhelming effects of location-awareness on collaboration in a pervasive game. *Paper presented at the International Conference on the Design of Cooperative Systems*. May 9-12, Carry-le-Rouet, Provence, France.

Oztop, E., Kawato, M., & Arbib, M. (2006). Mirror neurons and imitation: A computationally guided review. *Neural Networks*, *19* (3), 254–271.

Parr, L. A., Waller, B. M., & Fugate, J. (2005). Emotional communication in primates: implications for neurobiology. *Current Opinion in Neurobiology*, *15* (6), 716–720.

Payr, S. (2005). Not quite an editorial: Educational agents and (e-)learning. *Applied Artificial Intelligence*, *19* (3–4), 199–213.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47* (3), 329–342.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*, New York: Viking.

Prendinger, H., & Ishizuka, M. (2005). The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, *19* (3–4), 267–285.

Prensky, M. (2005). *In educational games, complexity matters - Mini-games are trivial - but "complex" games are not - An important way for teachers, parents and others to look at educational computer and video games*, Retrieved August 15, 2009, from http://www.marcprensky.com/writing/Prensky-Complexity_Matters.pdf.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, *6*, 65-65.

Puntambekar, S. (2006). Analyzing collaborative interactions: Divergence, shared understanding and construction of knowledge. *Computers & Education*, *47* (3), 332–351.

Reeves, T. C., Herrington, J., & Oliver, R. (2004). A development research agenda for online collaborative learning. *Educational Technology Research and Development*, *52* (4), 53–65.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21* (5), 188–194.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Robbins, J. (2005). Contexts, Collaboration, and cultural tools: A sociocultural perspective on researching children's thinking. *Contemporary Issues in Early Childhood*, *6* (2), 140–149.

Roth, W. M., & Lee, Y. J. (2007). "Vygotsky's Neglected Legacy": Cultural-Historical Activity Theory. *Review of Educational Research*, *77* (2), 186–232.

Sha, L., & van Aalst, J. (2003). An application of social network analysis to knowledge building. *Paper presented at the Annual Meeting of the American Educational Research Association*, April 21-25, Chicago, IL.

Shaffer, D. W. (2005). Epistemic games. *Journal of Online Education*, *1* (6), Retrieved September 18, 2007, from http://www.innovateonline.info/index.php?view=article&id=79.

Sorama, K., Katajamäki, A., & Varamäki, E. (2004). Cooperation between SMES: Social capital and learning perspective. *Proceedings of the 13th Nordic conference on small business research*, Retrieved May 1, 2009, from http://web.bi.no/forskning/ncsb2004.nsf/23e5e39594c064ee852564ae004fa010/a6cb7066ea59eda6c12567f30056ef4d/$FILE/Sorama&al.pdf.

Sorenson, O., Rivkin, J. W., & Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy*, *35* (7), 994–1017.

de Souza e Silva, A., & Delacruz, G. C. (2006). Hybrid reality games reframed. Potential uses in educational contexts. *Games and Culture*, *1* (3), 231–251.

Steinkuehler, C. (2004). Learning in massively multiplayer online games. *Proceedings of the 6th international conference on learning science*, International Society of the Learning Sciences, 521–528.

Stetsenko, A. (2004). Tool and sign in the development of the child. In R. W. Rieber (Ed.), *The essential Vygotsky*, New York: Kluwer Academic/Plenum, 501–512.

Stetsenko, A. (2005). Activity as object-related: Resolving the dichotomy of individual and collective planes of activity. *Mind, Culture, and Activity*, *12* (1), 70–88.

Thomas, A. (2005). Children online: Learning in a virtual community of practice. *E–Learning*, *2* (1), 27–38.

Thomas, S. (2006). Pervasive learning games: Explorations of hybrid educational gamescapes. *Simulation and Gaming*, *37* (1), 41–55.

Uddin, L. Q., Iacoboni, M., Lange, C., & Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences*, *11* (4), 153–157.

Vianna, E., & Stetsenko, A. (2006). Embracing history through transforming it: Contrasting Piagetian versus Vygotskian (Activity) theories of learning and development to expand constructivism within a dialectical view of history. *Theory & Psychology*, *16* (1), 81–108.

Vygotsky, L. (1933/1978). *Mind in society. The development of higher psychological processes*, Cambridge, MA: Harvard University Press.

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *Knowledge Engineering Review*, *10* (2), 115–152.

Young, M., Schrader, P., & Zheng, D. (2006). MMOGs as learning environments: An ecological journey into Quest Atlantis and The Sims Online. *Innovate*, *2* (4), Retrieved May 1, 2009, from http://students.ou.edu/M/John.S.Madden-1/assets/pdf/mmogs.pdf.

Yuan, Y. C., & Gay, G. (2006). Homophily of network ties and bonding and bridging social capital in computer-mediated distributed teams. *Journal of Computer-Mediated Communication*, *11* (4), 1062–1084.

Shih, K.-P., Chen, H.-C., Chang, C.-Y., & Kao, T.-C. (2010). The Development and Implementation of Scaffolding-Based Self-Regulated Learning System for e/m-Learning. *Educational Technology & Society*, *13* (1), 80–93.

The Development and Implementation of Scaffolding-Based Self-Regulated Learning System for e/m-Learning

Kuei-Ping Shih[1], Hung-Chang Chen[2], Chih-Yung Chang[1*] and Tai-Chien Kao[3]
[1]Department of Computer Science and Information Engineering, Tamkang University, Tamsui, Taipei County, Taiwan // Tel: +886-2-26215656 Ext. 2748 // kpshih@mail.tku.edu.tw // [*]cychang@mail.tku.edu.tw
[2]Department of Information Technology, Ching Kuo Institute of Management and Health, Keelung, Taiwan // Tel: +886-2-24372093 Ext. 273 // gileschen@ems.cku.edu.tw
[3]Institute of Education, National Dong Hwa University, Hualien, Taiwan // Tel: +886-3-8635571 Ext. 5571 // mkao@mail.ndhu.edu.tw

ABSTRACT

This paper proposes a self-regulated learning (SRL) system with scaffolding support in order to develop independent learning skills among students. The SRL system uses self-regulated learning and scaffolding theories to appeal to both instructors and learners. On the part of the instructors, a Content Accessibility Subsystem is provided to easily organize learning materials and to dynamically provide different levels of support for their learners. As for the learners, many subsystems are proposed that provide a conducive mobile learning environment for them. With the application of the scaffolding theory, the system can easily adjust to provide help to the learners, facilitating SRL processes anytime and anywhere, and establishing the learners' SRL patterns gradually. The learners in the experiment deemed that that the proposed system could provide them self-regulatory attributes. The experiment results show that the average SRL score of learners increases, though the improvement is not significant. However, the result also showed that the SR skills of students in the group of Low SR significantly improved.

Keywords

Self-Regulated Learning, Self-Regulatory Learning Cycle, Scaffolding, Mobile Leaning, E-Learning, CAL systems

Introduction

The main goal of education is to develop the character of students and foster in them a spontaneous desire to learn. To achieve this aim, self-regulated learning (SRL) is essential. However, while modern technologies have made learning possible at any time and place, there still is the challenge to provide a conducive environment so that learners can easily schedule their study plans and avail of learning materials outdoors. To address this, this paper proposes an SRL system with wireless technologies.

Lately, people from both academic and government sectors have keenly promoted SRL because they recognize the need to help learners take charge of their own education. However, SRL is not an easy task. Four factors are essential in carrying out SRL: learning schedules, materials, scenarios, and quality (Zimmerman, Bonner, & Kovach, 1996). Along with these factors, other difficulties in performing SRL may be pointed out:

- Learning schedules and materials: A suitable learning schedule makes a person's own learning methodical. However, if there is insufficient experience in the design of a learning schedule, a person may end up having poor SRL performance. In addition, because the study materials are limited and varied, a learner may find it difficult to organize the materials he needs.
- Learning scenarios and quality: The rapid development of modern technologies, such as broadband and wireless communication engineering, makes learning materials easily available. However, because there is no tailor-made learning environment for outdoor scenarios, learners may just give up learning due to the difficulty or the complexity of accessing the learning materials. Moreover, because of many distractions, learners may be unable to focus well.

Therefore, an SRL system that adopts the concept of a self-regulatory learning cycle (Zimmerman et al., 1996) is proposed. Because having an ambitious and unrealistic aim may disappoint learners during the process of learning, the proposed system firstly helps students set a reasonable goal in initiating their motive. Moreover, the system adopts the scaffolding theory (Bruner, 1983), which gradually builds their learning patterns. Through this theory, the system can provide students with information and materials they need. The success of the scaffolding depends on the precise evaluation of the learning outcomes such that the learning scaffolding can be removed properly. Therefore, a reliable evaluation system is provided so that learners can determine their progress. As a result, they can set reasonable goals in the system and even develop their SRL skills as they go through each self-regulated learning cycle.

Another aim of the system is to support a mobile learning environment for learners through modern wireless technologies via mobile devices. The system enables learners to access learning materials easily and conveniently. Moreover, the system automatically observes learners' behaviors to help them terminate unhelpful habits, e.g., using instant messengers or surfing while learning. The mobile learning environment also enables them to share learning materials, allowing them to adjust their strategies based on the data from their companions. Because of this, the system would have the characteristics of a mobile learning environment: *Urgency of learning need*, *Initiative of knowledge acquisition*, *Mobility of learning setting*, *Situating of instructional activity*, and *Integration of instructional content* (Chen, Kao, & Sheu, 2003).

The rest of the paper is organized as follows: Section 2 introduces the design rationale of the proposed system as well as previous works on SRL. Section 3 introduces the proposed SRL system. Section 4 describes the experimental results, and Section 5 concludes the study.

Review of Related Literature and Design Rational

Self-Regulated Learning Theory

While there are various explanations and studies that focus on the definition of SRL (Butler & Winne, 1995; Pintrich, 2000), it can be simply described as a learning process with four attributes (Schunk & Zimmerman, 1994):

- Intrinsically or self-motivated: Self-regulated learners tend to maintain learning behavior with a very strong motivation. Learners can raise this motivation through some practices, such as setting learning goals.
- Planned or automatized: Self-regulated learners are apt to use some strategies along with their learning processes, including both cognitive and self-regulated strategies. Generally, learners improve their learning performance when using self-regulated strategies rather than cognitive strategies. Self-regulated strategies contain goal-setting, goal-planning, organization, transition, exercise, and so on. A self-regulated learner needs to effectively use self-regulated strategies for his learning.
- Self-aware of performance outcomes: Throughout the learning process, self-regulated learners sharpen their self-awareness toward their learning behavior. To approach an ideal outcome, self-regulated learners should be aware of their own learning qualities, and change the behavior or strategies correspondingly.
- Environmentally/socially sensitive and resourceful: The learning environment and resources can affect one's learning pattern. Self-regulated learners have better skills in seeking learning resources or support. With such ability, they should arrange the environmental conditions and search for other resources effectively.

Knowing how to possess the above attributes should be considered when designing an adequate system for self-regulated learners. Once such attributes are possessed, learners can then skillfully self-regulate their learning.

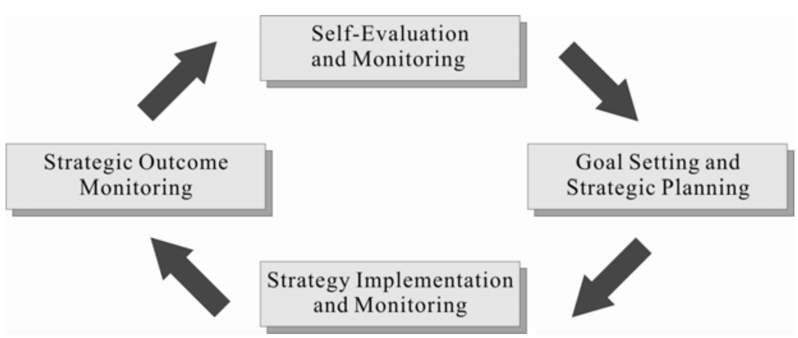

Figure 1. A cyclic model of self-regulatory learning (Zimmerman et al., 1996)

Zimmerman et al. (1996) proposed a self-regulatory learning cycle in order for learners to gain SR skills, as shown in Figure 1. The cycle involves four interrelated processes which assist learners in evaluating their performance. Generally, learners carry out their plans by themselves in these processes. Thus, such a model enables students to arrange their own learning and voluntarily fulfill it at the same time.

SRL with the Support of Computer and Wireless Technologies

With the aid of modern technologies, students can learn efficiently and achieve remarkable performance. Unlike in the traditional face-to-face set-up, today's students can individually determine when and where to learn. Therefore, many computer-based systems have been proposed to enhance a person's performance when he or she learns individually (Hadwin & Winne, 2001; Dabbagh & Kitsantas, 2004). However, Azevedo, Cromley, Thomas, Deibert, & Tron (2003) indicated that, when receiving no assistance, students are less effective at regulating their learning in their hypermedia environment. Because of this, some computer-assisted SRL tools have been proposed (Hadwin & Winne, 2001; Dabbagh & Kitsantas, 2004).

Hadwin & Winne (2001) proposed a prototype electronic notebook, CoNets2, to support self-regulation through explicit scaffolding. Its system can support monitoring and controlling engagement in the phases of SRL, but unskillful self-regulated learners may shun the tool because CoNets2 lacks enough functions to motivate their learning. Moreover, the tool is often limited to taking down notes, and does not develop important SRL skills such as goal-setting, scheduling, and self-evaluation.

Dabbagh & Kitsantas (2004) classified Web-based pedagogical tools (WBPT) into four classes: (1) collaborative and communication tools; (2) content-creation and delivery tools; (3) administrative tools; and (4) assessment tools. These tools are also examined on their support of self-regulatory attributes. However, an integration of these tools should be promoted to gradually make learners become skillful in SRL.

Zurita & Nussbaum (2004) proposed a constructivist learning environment, which allows students to build up their own knowledge. Wireless interconnected handhelds are used in such an environment to achieve the creation of new knowledge. By using wireless interconnected handhelds, students can able to modify their current knowledge schemes which could integrate new information and acquire new knowledge. The process of knowledge construction partially matches the processes of a self-regulatory learning cycle. However, the environment can only be used for information sharing, and is unable to help students completely monitor their learning strategies and outcomes.

Looi et al., (2009) designed and implemented a software on a mobile device called the 3Rs (reduce, reuse, and recycle) software. The mobile device used in the activity can lead students to carry out learning tasks in challenge-experiential cycles, including *Challenge*, *Experience*, *Reflecting*, *Planning*, and *Applying*. The activity also partially matches the self-regulatory learning cycle. However, during the process of *Experience*, learners may not receive any assistance, and their learning experience is not recorded and used for future leaning. The comparisons of the proposed system and the aforementioned tools are summarized in Table 1.

Table 1. Comparisons of the proposed system to the related works

	CoNets2 (Hadwin & Winne, 2001)	WBPT (Dabbagh & Kitsantas, 2004)	Constructivist Learning Environment (Zurita & Nussbaum, 2004)	3Rs software (Looi et al., 2009)	The proposed system
Support self-regulatory attributes?	Yes (Partial)	Yes	Yes (Partial)	Yes (Partial)	Yes
Support scaffolding?	Yes	No	No	No	Yes
E-learning?	Yes	Yes	Yes	Yes	Yes
M-learning?	No	No	Yes	Yes	Yes

The Proposed Learning System

Basic Concept

Shih, Chang, Chen, & Wang (2005) proposed the prototype of the system. According to learner feedback and opinions from educational theorists, we will continually enhance the proposed system to efficiently improve the SRL performance of learners. This section introduces the mapping of the proposed system and the cyclic model. Moreover, the section also tackles the functions and the design considerations of the proposed system.

This paper proposes a state transition diagram that indicates the behavior of a learner by using the cyclic model of SRL to become self-regulatory (as shown in Figure 2). The diagram consists of seven states, indicating the actions of a learner in SRL. Among these seven, Activity Scheduling, Learning and Monitoring, Learning Evaluation, and Analysis are the major states that map the four processes in the cyclic model of Zimmerman et al. (1996).

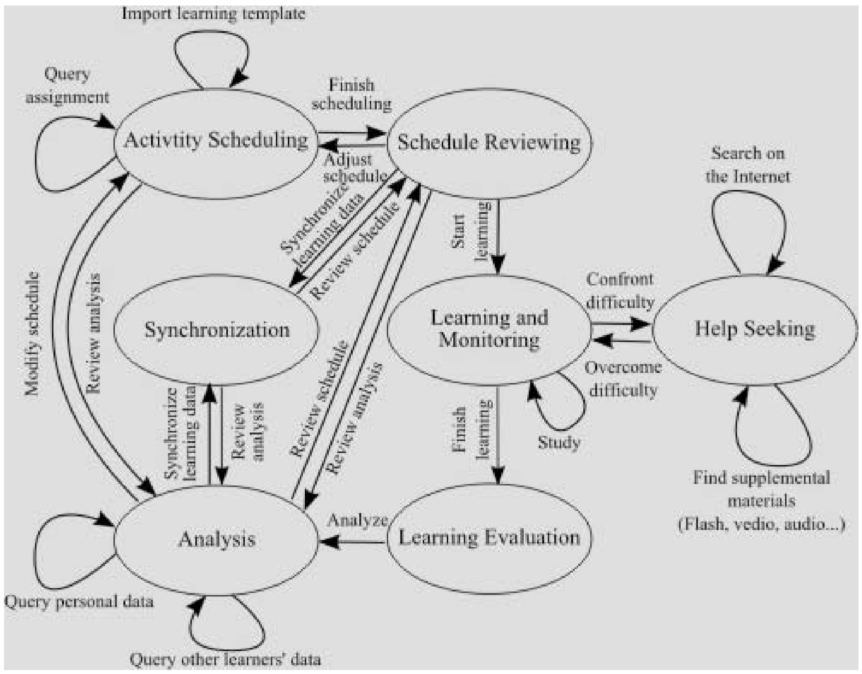

Figure 2. State transition diagram of the system

Through the proposed system, learners in the Activity Scheduling state can obtain information on what to learn and can arrange suitable times according to the provided information, all of which leads to the Goal Setting and Strategic Planning process. After setting their schedules, learners enter the Learning and Monitoring states. Because learners can undertake scheduled activities by using various strategies while they are being observed, the state maps the Strategic Implementation and Monitoring process. In the Learning Evaluation state, learners can assess their progress through tests. Accordingly, the state also maps the process of Strategic Outcome Monitoring. In addition, learners can evaluate their development by means of varied statistical charts, and can discern their SRL patterns in the Analysis state before it leads to the Self-Evaluation and Monitoring processes. In obtaining a detailed understanding of their learning characteristics, students can then go into the Activity Scheduling state again, and arrange more rigid schedules for future learning.

In addition to the aforementioned states, the Synchronization, Help Seeking, and Schedule Reviewing states are also involved in the system to help learners gradually develop their SRL skills.

Figure 3 illustrates the architecture of the proposed SRL System, which supports the state transition diagram in Figure 2. The black arrows in Figure 2 are the data flows between subsystems. This means the learning template is used to help learners in their schedule. The schedule is the learning schedule arranged by the learners. According to the learning schedule, the proposed system can know when learners plan to learn and observe their behavior. The result represents the information generated when learners use our system.

Generally, beginners or unskilled learners cannot arrange their learning well because of their lack of experience in self-regulation. As such, instructors can use the instructor side system to help them in provide directions and scope, or suitable scaffolds. On the other hand, the learner side system aims to form a pleasant SRL environment wherein they would be able to practice their SRL skills. The system is planned to be installed in portable learning devices, where they can schedule, perform, and evaluate their progress anytime and anywhere. In the following, the design of the functions provided by the proposed system is presented according to the state transition in Figure 2 and the data flows in Figure 3.

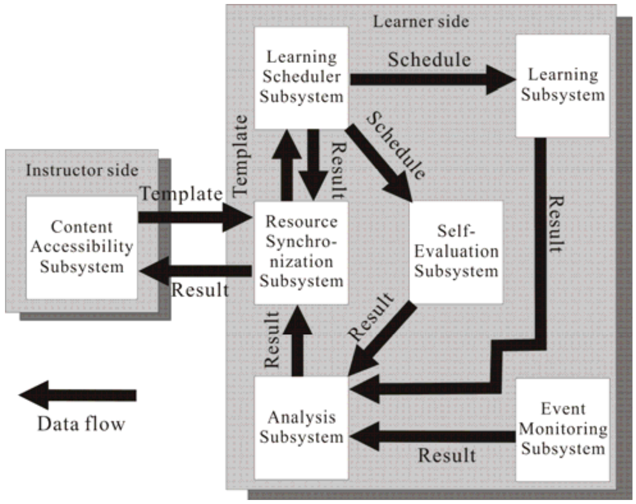

Figure 3. Architecture of the Self-Regulated Learning System

The Data Flow of the SRL System

However, in order to help learners efficiently arrange their learning activities, the information is controlled by a Scaffold Support Module based on their SRL performance. The purpose of scaffolding is to provide novice learners with limited complexities of learning context and to remove limits gradually until they become more skillful (Young, 1993). Therefore, the interface initially shows much information until they become more skillful at SRL, so the learners can then control their learning gradually. Notice that learners are not forced to use given information, but can decide to refer to the given information.

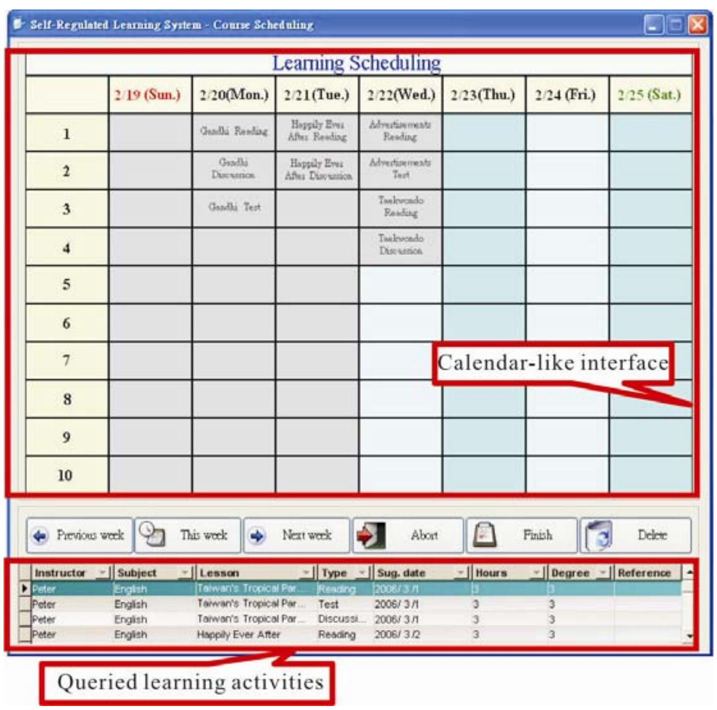

Figure 4. Learning schedule planning

Previous research have pointed out that intervention increases cognition and motivation, and leads to the development of self-regulated capabilities (Hofer, Yu, & Pintrich, 1998). In our design, instructors are involved in helping learners become skillful in SRL through the Content Accessibility Subsystem. Instructors can conveniently design and give assignments and activities to learners. The subsystem also generates a learning schedule template based on the assignments given by the instructors. Interfaces in Figures 5 and 6 enable instructors to arrange these assignments and set up the access to learning materials. Through the interface in Figure 5, instructors can set up detailed information, including degree, session, semester, and credit, in the template. The information can also be shared by instructors to other instructors, making it easy for them to design templates.

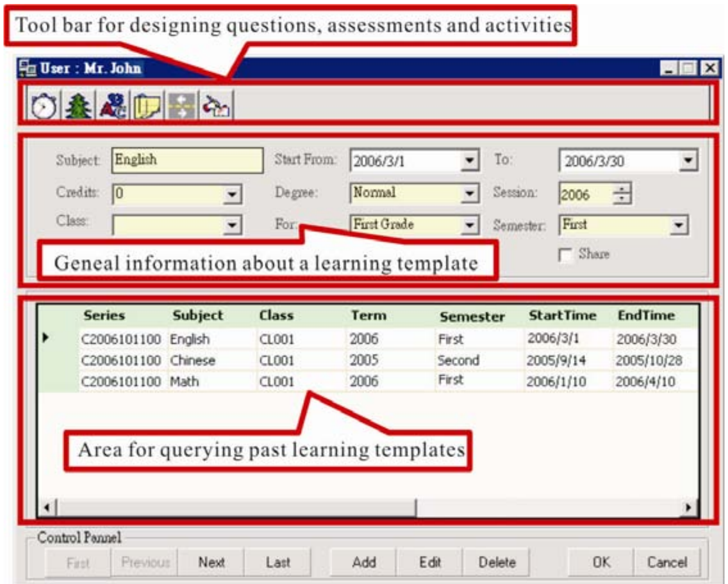

Figure 5. Learning schedule template design

Once learning materials are available, instructors can arrange the details of an assignment through the interface in Figure 6. Instructors can set up information on a learning unit, including its type (activity, self-examination, or discussion), suggestion time, and materials. This way, the information can help learners in their schedules and can also be used by the Scaffold Support Module to generate supplemental information and materials. Similarly, instructors can query designed learning units, which can be directly imported to a schedule template. After the instructor's arrangement, the subsystem then generates a schedule template for learners. Through the aid of wireless technology, the system will automatically download the learning templates from a central database whenever learners move into Wi-Fi hotspots.

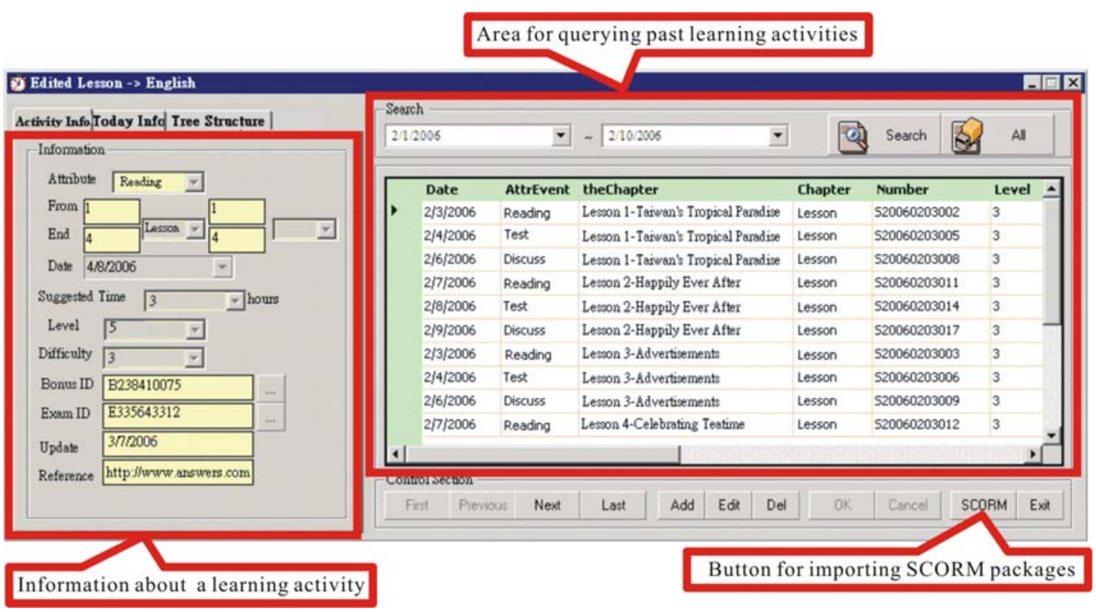

Figure 6 Arrangement for the details of a learning template

After scheduling, the system automatically synchronizes to a central database and downloads learning materials when available. Once learners enter the Schedule Reviewing state via the learning review tool of the Learning Subsystem, a calendar-like interface, where the scheduled and learned activities are marked, is provided, as shown in Figure 7. Each selected item on the interface represents a learning activity, and the interface can show basic statistics (e.g., learning time and the number of interruptions) of a selected item. This information helps learners find their preferred schedules (e.g., in the afternoon or in the morning).

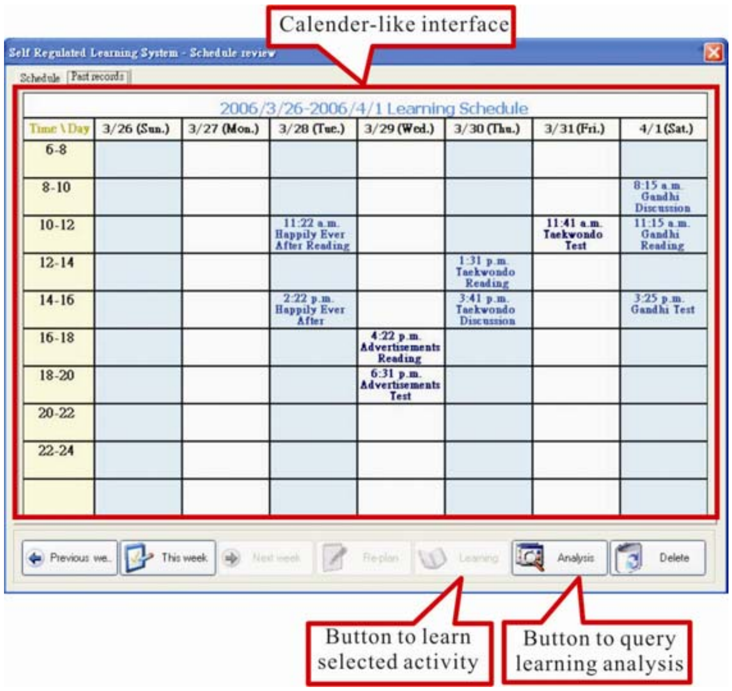

Figure 7. Learning review tool

When in the Learning and Monitoring states, learners start by clicking the “Learning” button on the learning review tool. The Learning Subsystem uses two tools, Hyperbook and Hyperpen, to enrich the learning experience. Hyperbook is a hardcopy book with reference tags. In principle, reading a hardcopy book is comfortable for learners. However, the content of a book is limited and fixed. In learning devices, staring at the monitor for a long time tire students easily. Furthermore, they may shun the complicated operation of learning devices because they may need to alternately use varied input equipment to obtain supplemental materials. Thus, to facilitate operations, students can use a scanning device, termed Hyperpen, to scan reference tags for more supplemental materials, such as Flash (Macromedia, 2005), audio, and video. Supplementary materials will then be shown in the learners’ mobile devices.

Hyperpen is embedded with a Bluetooth solution (Bluetooth, 2003) to avoid cables that may distract students as they scan. Scanned keywords are then submitted to an Internet dictionary, such as Yahoo! Dictionary (Yahoo! Taiwan Inc., 2007), or the database, such as Answers.com (Answers.com, 2005). Because wireless technologies are heavily promoted, hotspots can be found all around, and almost all learning devices have WLAN capabilities. Therefore, the function of searching supplemental materials on the Internet can be carried out everywhere. Figure 8 shows the system interface where learners use Hyperbook and Hyperpen to avail of supplemental materials on the Internet. Hyperbook is manufactured by HardSCORM Editor (Wang & Shih, 2006), an authoring tool that conforms to SCORM 2.0 (ADL Technical Team, 2006). The Content Accessibility Subsystem is able to recognize the edited courses and can split these into several learning activities based on their metadata.

Because the most important performance control process that distinguishes skillful from naive self-regulated learners is self-monitoring (Zimmerman & Paulsen, 1995), an Event Monitoring Subsystem is needed to observe their behavior. If learners can monitor their own progress, their academic performance, achievement, time on task, classroom behavior, and problem-solving abilities can be improved (Lan, 1998). We therefore give much attention to recording learner behavior.

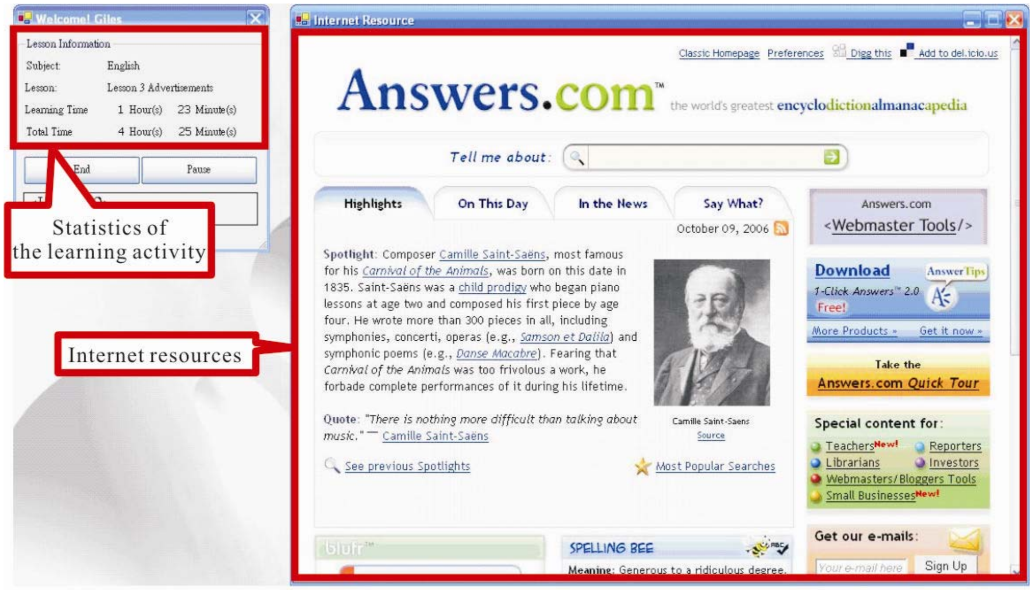

Figure 8. Searching the database in Internet (Answers.com, 2005)

Monitoring items include the learner's schedule and learning behavior. Schedule includes data on whether a learner studies on time and how much time a learner spends on an activity. Learning behavior involves what students do while learning, such as the time and reasons of interruptions, and the frequency, quantity, and their ways of seeking help. Monitoring items can be also easily recorded when learners engage in learning activities. Some learning interruptions may be caused by various situations, such as the presence of a TV or domestic errands. It is impossible for a computer-based system to record events automatically. Because the proposed system is designed for students who want to become self-regulated learners, our system provides an interface for them to manually input interruptions easily. Monitoring items will be demonstrated in the form of charts after the behavior analysis so learners can quickly understand their study habits.

Psychologists argue that regularity, referring to the need to observe behavior frequently instead of sporadically, and proximity, referring to behavior that should be recorded close in time to its occurrence, are important characteristics of effective self-monitoring (Bandura, 1986). The proposed system provides an easy self-monitoring environment that has the characteristics of regularity and proximity. For regularity, the Event Monitor Subsystem continuously observes learners' behavior anywhere and anytime they experience SRL. As for proximity, learning behavior can be monitored immediately. Therefore, it is expected that the system can help them understand their learning habits.

The learners need to evaluate their progress either during or after an activity because both the difficulty of the examination and lack of preparation may incur disappointing performance (Ghatala, Levin, Foorman, & Pressley, 1989). The evaluation should be based on both objective and subjective criteria. The system has a Self-Evaluation Subsystem that includes an Assessment Module and a Self-checking Module to obtain objective and subjective cognitions, respectively. Learners are able to identify the gap between what they think about their learning and what learning outlook they actually have by comparing the objective and subjective cognitions.

After finishing an activity, learners then proceed to the Analysis state, where their behavior is examined. Through the aid of data mining, meaningful information is dug up from the data in the learners' profile. The information includes the differences between learning and the scheduling of an activity and the expected and real scores, among others. This information is used to determine barriers and bottlenecks in the learning processes, and to find solutions to certain problems. Students can also read others' learning analysis, which may help them to either adopt strategies to other learners or help them find their fulfillment (e.g., rank high among ones classmates.) According to the monitoring items, many kinds of analysis charts for learning time, interruptions, used materials, and so on are provided.

When in the Analysis state, learners are able to better understand their learning patterns when provided with the aforementioned information. Learners therefore try to use various strategies to improve their performance. The instructors can also receive the feedback information for the preparation of following learning templates through information exchange. In addition, the system automatically downloads other learners' profiles so that those who learn the same activity can refer to the learning strategies of others. By entering the Activity Scheduling state again, they better arrange their schedules based on the experience acquired from prior learning activities, allowing them to perform more effectively. In the process of doing so, they can be self-regulated learners.

Evaluation of the proposed SRL system

An experiment was conducted in a high school to demonstrate the effectiveness of the proposed SRL system. The experiment addresses the issues: "Can the system help learners possess the four self-regulatory attributes?" and "How much does the system help learners to improve their SRL efficiency?"

Method

In the experiment, the target learners were secondary students because we assumed that, though they were eager to learn, they still lacked SRL skills. The learning topic was English. The experiment consisted of two steps to evaluate the effectiveness of the proposed system. The first step was to classify the students who learned English in a self-regulatory way. Students performed SRL without the help of the system in this step. Since the number of students in this step is large enough, the classification indicates the SRL types of secondary school students in Taiwan. This classification was used in Step 2 to identify the SRL types of students. Through this step, the following step discusses the differences between SRL types of students before and after using our SRL system. A pre-test and a post-test were conducted in the second step to evaluate the improvement of the students' SRL skills after using the proposed system. The SRL types of learners in Step 2 were classified into three types based on the classification built in Step 1. The analysis of the experiment mainly focused on the learners in Step 2 because the proposed system was involved in their learning. The impacts of the use of the proposed system on the learning performance of the different types of SRL learners are observed in this step.

In Step 1, four grade ten classes (42 students in one section, and 43 students in the others) were chosen to self-regulate their English learning for 10 weeks in 2004. Students had the same English instructor and were asked to study 10 lessons from the Studio Classroom magazines (StudioClassroom, 1962), a popular English instructional magazine in Taiwan. Their instructor taught them some SRL skills and gave them hardcopy forms to record their learning behavior and their reflections. After the ten weeks, every student was asked to fill out a MSLQ (Motivated Strategies for Learning Questionnaire). The original MSLQ was designed by Pintrich, Garcia, & Mckeachie (1993) to assess college students' motivational orientations and their use of learning strategies. Wu and Chan translated the MSLQ into Chinese and modified the questionnaire for elementary school students (Wu, 1998). In the experiment, the MSLQ was used for high school students taking up English. There are 91 items in the MSLQ, and some are listed in Table 2.

The questionnaire used a seven-point Likert-type scale. The students would receive one to seven marks for each item in the MSLQ. They were divided into three groups (high SR, medium SR, and low SR) according to normal distribution. Students whose SRL scores were higher than 485 marks (25% students) formed the group with high SR, students whose SRL scores were between 484 marks and 423 marks (50% students) formed the group with medium SR, while the others formed the group with low SR (25% students).

Table 2. Partial items in the MSLQ used in the experiment

I believe I will receive an excellent grade in class.
I'm certain that I can understand the most difficult material in English Learning.
I am very interested in the content area of English learning.

In Step 2, 17 volunteers from one of the grade 11 classes (apart from the four classes in Step 1) were involved in a three-week SRL. Each student was given a Hyperbook, a Hyperpen, and a tablet PC. The Hyperbook contained six English lessons from the IVY magazines (Ivy League, 2006), which are also popular English learning materials in Taiwan.

Before the experiment, a pre-test was employed for the 17 students. The MSLQ used in Step 1 was also used in the pre-test, with the goal of determining the SRL patterns of the students before using our system. Three weeks later, the students took a post-test. They were asked to fill out two questionnaires. One was an MSLQ, which was used to discern the SRL patterns of the students after the experiment. The other was the Self-Regulated System Indication Questionnaire (SRSIQ), which was used to evaluate the support of the self-regulatory attributes.

Generally, an SRL can be surveyed in different psychological dimensions of research by using following scientific questions: *why*, *how*, *what*, and *where* (Schunk & Zimmerman, 1994). SRSIQ asked these scientific questions in a questionnaire filled out by students, as listed in Appendix A. The questionnaire applied a five-point Likert-type scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The results retrieved from this questionnaire were used to determine the assistance of the system during the students' SRL.

Results

The experiment results consisted of two parts. The first focuses on the analysis of the four self-regulatory attributes based on the SRSIQ results of the students. The second part looks into the students' progress in acquiring the SRL skills.

The Support of Self-regulatory Attributes

First, the students were asked if the system helped them acquire the four self-regulatory attributes. The reliability of whether our system motivated students is 0.81, implying that the questionnaire items have a high reliability. The result shown in Table 3 indicates that the system supports the attribute (Mean=3.329, SD=0.897). This subscale is greatly influenced by the amount of motivation inspired by our system. The SRL system encourages students to use supplementary multimedia materials (Q13, Q14, Q15, and Q16) and learning analysis (Q27). Most students thought that supplementary multimedia materials were useful for their learning. According to the results of the behavior observation, each student was willing to access the supplementary materials (74.14 materials per student) Therefore, it can be shown that the supplementary materials and the methods used to access the Internet enriched their learning experiences and made them eager to learn more. Therefore, the students had spontaneous pleasure in learning English.

Table 3. Item statistics of attribute: Intrinsically or self-motivated

	Mean	Std. Deviation	N
Q13	3.24	.970	17
Q14	3.47	.943	17
Q15	3.47	1.007	17
Q16	3.47	1.068	17
Q27	3.00	.707	17

The reliability of whether our system helps students become systematic is 0.758. This indicates that the questionnaire items also have a high reliability. The results also show that all the functions enabled learners to use proper learning strategies (Mean=3.34, SD=0.749). Among the questionnaire items listed in Table 4, prior scheduling experience gave significant assistance to students in planning their schedules (Q9). Nevertheless, item Q10 indicates that the interface of the learning review tool should be more user-friendly. Because the interface of the tool is divided into two parts, students had to frequently switch from one part to the other in order to view their scheduled and learned activities. Aside from this, other functions (e.g., synchronizing learning records and starting a learning activity) were executed through the tool as well, so students thought that the operation of these functions was complicated. Therefore, students deemed that the interface was not user-friendly enough, and that it should be simplified and intuitive.

The response on whether the system helped students become self-aware of their performance outcomes also has a good reliability (Cronbach's Alpha=0.91). Thirteen items tapped on subjective learning performance. The item statistics, shown in Table 5, show that the information brought to learners gave them a sense of fulfillment (Mean=3.357, SD=0.766). Through learning analysis, students thought that our system could precisely monitor their behavior and easily record interruptions. Most students also agreed that the Self-Evaluation Subsystem helped them to conveniently record the cognition to a learning activity, which was used to determine the gap between subjective and objective learning achievements (Q18). Students also agreed that the learning analysis was helpful (Q25). However, though learning monitoring was generally regarded as a useful function to understand one's own learning status (Q19; Q20), the students felt that the function was a little bit hard to use (Q21). This may be attributed to the fact that the behavior was observed when students used the functions of the proposed system. In the experiment, the proposed system was installed in the tablet PC, which ordinarily had no input devices such as a keyboard and a mouse. Unlike desktop PCs, the operation in the tablet PC was more difficult and unfamiliar to the students. Additionally, the recognition ratio of the Hyperpen was about 89%. Students were not satisfied with the ratio. "It is hard to scan the tags," one student said. Therefore, our future work should address this because higher recognition rates can encourage the use of our system.

Table 4. Item statistics of attribute: Planned or automatized

Mean	Std. Deviation	N	Mean
Q1	3.12	.781	17
Q2	3.12	1.166	17
Q3	3.29	.849	17
Q4	3.29	.686	17
Q5	3.47	1.068	17
Q6	3.47	.717	17
Q7	3.47	.874	17
Q8	3.53	.874	17
Q9	3.59	.870	17
Q10	3.06	.966	17
Q24	3.41	.795	17
Q25	3.29	.686	17
Q30	3.29	.772	17

Table 5. Item statistics of attribute: Self-aware of performance outcomes

	Mean	Std. Deviation	N
Q11	3.41	.939	17
Q12	3.35	.702	17
Q18	3.47	.800	17
Q19	3.29	.920	17
Q20	3.35	.702	17
Q21	3.00	.866	17
Q22	3.41	.795	17
Q23	3.41	1.004	17
Q25	3.47	.717	17
Q28	3.24	1.200	17
Q29	3.41	.870	17
Q32	3.47	.800	17
Q33	3.35	.931	17

On average, students were positive on the attribute Environmentally/socially sensitive and resourceful (Cronbach's Alpha=0.852). The result in Table 6 indicates that the function of the learning record synchronization (Q31), the learning materials (videos (Q13), pronunciations (Q14), translations (Q15), and phrases (Q16)), as well as Internet searching (Q17) could inspire students to obtain and seek useful learning resources, as shown in Table 6. Among them, Searching on the Internet (Q17) was the least useful function. Because the topic was English, the students thought that the resources provided by the system were enough and so they had less will to search for extra resources from the Internet. In the future, the system should be modified to support different types of help seeking functions for different kinds of learners. For example, the system should enable skillful learners to search for extra materials and allow unskilled learners to become accustomed to using additional resources.

Genreally, the target students deemed that the proposed system could help them possess the four self-regulatory attributes, albeit some functions have to be improved.

SRL Effectiveness

In this section, the SRL scores of the students are studied. These students are also classified into three groups, according to the results in Step 1. Of the total number of students, 71% increased their SRL scores after using our system. On average, the SRL scores of these students increased by 11.06 points.

For a precise analysis, a T-test was used to determine if the proposed system could efficiently improve the learners' SRL scores. We find that the difference between the means is not statistically significant ($t = -1.606$, $df = 16$, $p > 0.05$, one tailed), so the system cannot help students to significantly improve their SRL performance. This may be because SRL skills should be developed over a long period of time, and three weeks of SRL may not be long enough for learners to improve their SRL skills.

Table 6. Item statistics of attribute: Environmentally/ socially sensitive and resourceful

	Mean	Std. Deviation	N
Q31	3.29	.772	17
Q13	3.24	.970	17
Q14	3.47	.943	17
Q15	3.47	1.007	17
Q16	3.47	1.068	17
Q17	3.18	1.131	17

The SRL scores of all students in the group of low SR increased, except for one student. Those who lack learning experience can easily follow different learning styles, such as computer-assisted learning or distance learning. On the other hand, skillful learners have more difficulty in changing the learning patterns they have developed (Morgan, Dingsdag, & Saenger, 1998). Therefore, whether or not the proposed system significantly improves the SRL scores of the students in the group of low SR is analyzed. The difference between the means is significant at the 0.05 level ($t =-3.136$, $df = 9$, one tailed). The result shows that the SR skills of students in the group of Low SR significantly improved, which corresponds with the findings of Morgan et al.

Conclusions

This paper proposes an SRL system that involves the self-regulatory learning cyclic and scaffolding theories to cultivate self-regulated learners. The system aims to construct a mobile, portable, and personalized learning environment for SRL that can be used anywhere and anytime. To help learners gradually develop their SRL skills, instructors are involved in the system. The learner side system enables students to start SRL anytime and anywhere, obtain learning materials and assistance instantly, realize their learning patterns, cultivate their SRL behavior, and sustain their interest in self-learning.

Generally, the experiment results show that the system has improved SRL skills, though the improvement is not significant. However, the result showed that the SR skills of students in the Low SR group improved significantly. Moreover, most of students deemed that the overall interfaces of the proposed system were user-friendly and could give them valuable progress in SRL. The students also agreed that the system enabled them to acquire the four SRL attributes.

Based on the suggestions of the students, we will make the interfaces of the SRL system more user-friendly and improve the recognition ratio of the Hyperpen to facilitate its operations. We also aim to use other devices such as PDAs or mobile phones as learning instruments because they are easy to carry and can facilitate quick learning. As a result, learners would more be likely to use the proposed system. Experiments on a larger number of students over a longer period of time shall also be conducted to improve the system.

Acknowledgement

The work was partially supported by the National Science Council of the Republic of China under Grants NSC 97-3114-E-119-001, NSC 97-2221-E-032-021, and NSC 98-2218-E-254 -001.

References

ADL Technical Team (2006). *SCORM 2004 3rd Edition*, retrieved Jun. 18, 2009 from http://www.adlnet.gov/scorm/index.cfm.

Answers.com (2005) *Answers.com- Online Encyclopedia, Thesaurus, Dictionary definitions and more*, retrieved Aug. 23, 2009 from http://www.answers.com/.

Azevedo, R., Cromley, J. G., Thomas, L., Deibert, D., & Tron, M. (2003). Online process scaffolding and students' self-regulated learning with hypermedia. *Annual Meeting of the American Educational Research Association*, 31.

Bandura, A. (1986). *Social Foundations of Thought and Action*, Englewood Cliffs, NJ: Prentice-Hall.

Bluetooth® SIG, Inc. (2003). *Bluetooth Core Specification Version 2.1 + EDR*, retrieved Jun. 18, 2009 from https://www.bluetooth.org/spec/.

Bruner, J. (1983). *Child's talk. Learning to Use Language*, New York: W. W. Norton.

Butler, D. L. & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, 65(3), 245-281.

Chen, Y.-S, Kao, T.-C., & Sheu, J.-P. (2003). A Mobile Learning System for Scaffolding Bird Watching Learning. *Journal of Computer Assisted Learning, 19*(3), 347-359.

Dabbagh, N., & Kitsantas, A. (2004). Supporting self-regulation in student-centered web-based learning environments. *International Journal on E-Learning*, *3*(1), 40-47.

Ghatala, E. S., Levin, J. R., Foorman, B. R., & Pressley, M. (1989). Improving children's regulation of their reading prep time. *Contemporary Educational Psychology, 14*, 49-66.

Hadwin, A. F., & Winne, P. H. (2001). CoNoteS2: A software tool for promoting self-regulation, *Educational Research & Evaluation*, 7, 313-334.

Hofer, B. K., Yu, S. L., & Pintrich, P. R. (1998). Teaching College Students to Be Self-Regulated Learners. In Schunk, D. H. & Zimmerman B. J. (Eds.) *Self-regulated learning- from teaching to self-reflective practice* (pp. 57–85), New York: Guilford.

Ivy League (2006). *Ivy League Analytical English*, retrieved May 3, 2009 from http://www.ivy.com.tw/.

Lan, W. Y. (1998). Teaching self-monitoring skills in statistics. In Schunk, D. H. & Zimmerman B. J. (Eds.) *Self-regulated learning- from teaching to self-reflective practice* (pp. 86–105), New York: Guilford.

Looi, C.-K., Seow, P., Zhang, B. H., So, H.-J., Chen, W., Wong, L.-H. (2009). Leveraging mobile technology for sustainable seamless learning: a research agenda. *British Journal of Educational Technology*, 41(2), 154-169.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In Boekaerts, M. & Pintrich, P. R. (Eds.) *Handbook of Self-Regulated* (pp. 13-39), San Diego: Academic.

Pintrich, P. R., Garcia, T. & Mckeachie, W. J. (1993). Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53(3), 801-813.

Macromedia Inc. (2005). *Macromedia Flash MX 2004*, retrieved Sep. 25, 2009 from http://www.macromedia.com.

Morgan, C. J., Dingsdag, D. & Saenger, H. (1998). Learning strategies for distance learners: Do they help. *International Journal Distance Education, 19*, 142-156.

Schunk, D. H. & Zimmerman B. J. (1994). *Self-Regulation of learning and performance: Issues and educational applications*, Hillsdale: New Jersey: Lawrence Erlbaum.

Shih, K.-P., Chang C.-Y., Chen, H.-C., & Wang, S.-S. (2005). A Self-Regulated Learning System with Scaffolding Support for Self-Regulated e/m-Learning. *Proceedings of the 3rd IEEE International Conference on Information Technology: Research and Education* (*ITRE* 2005), 30-34.

StudioClassroom (1962). *StudioClassroom.com-Your Friend for Life*, retrieved Sep. 25, 2009 from http://studioclassroom.com/.

Wang, T.-H. & Shih, T. K. (2006). Integration of multimodal multimedia devices and hardcopy textbooks for supporting pervasive e-learning. *Paper presented at the First International Symposium on Pervasive Computing and Applications* (SPCA06), August 3-5, 2006, Urumchi, Xinjiang, China.

Wu, J.-J. (1998). *Consortium of Research on Creativity and Innovation*, retrieved Oct. 25, 2009 from http://tim.nccu.edu.tw/croci/group/wujingji.htm.

Yahoo! Taiwan Inc. (2007). *Yahoo! Dictionary*, retrieved Sep. 12, 2009 from http://tw.dictionary.yahoo.com/.

Young, M.F. (1993). Instructional design for situated learning. *Educational Technology Research & Development 41*(1): 43-58.

Zimmerman, B. J., Bonner, S. & Kovach, R. (1996). *Developing Self-Regulated Learners: Beyond Achievement to Self-Efficacy*, Washington, DC: APA.

Zimmerman, B. J. & Paulsen, A. S. (1995). Self-monitoring during collegiate studying: An invaluable tool for academic self-regulation. *New directions for teaching and learning*, 1995(63), 13-27.

Zurita, G. & Nussbaum, M. (2004). A constructivist mobile learning environment supported by a wireless handheld network. *Journal of Computer Assisted Learning, 20*(4), 235-243.

Appendix A: Self-Regulated System Indication Questionnaire (SRSIQ)

Q1	Learning Scheduler Subsystem provides clear-cut lesson information to schedule learning.
Q2	With drag-and-drop, I plan my schedule smoothly.
Q3	Learning scheduler subsystem helps me determine the study plan.
Q4	Functions of Learning Scheduler Subsystem are varied and handy.
Q5	Learning Scheduler Subsystem is helpful in setting appropriate learning goals and plan setting.
Q6	Learning Review Tool helps me manage the learning activities.
Q7	Learning Review Tool makes me understand the time spent on prior learning activities.
Q8	I can see every learning activity by examining the information in the Learning Review Tool.
Q9	Tracking the previous learning schedules, I set better plan in the next step.
Q10	Learning Review Tool is designed with user-friendly interface.
Q11	The "pause" function helps me unhurriedly note down every interruption while learning.
Q12	Recording learning time helps me know my own learning progress.
Q13	Introductory videos provided by the system are useful for the beginning of my learning.
Q14	Recordings provided by the system are useful for my learning.
Q15	The function of text translation is useful for my learning.
Q16	The provided vocabularies and phrasal verbs in the Hyperbook are useful for my learning.
Q17	Hyperpen is convenient for searching resources on the Internet.
Q18	Self-evaluation helps me immediately note down my experience and feeling along the learning.
Q19	Items of learning recording and learning monitoring are listed comprehensively.
Q20	Learning and monitoring tools enable me to precisely manage my learning.
Q21	Learning and monitoring tools are designed with user-friendly interface.
Q22	Learning analysis outcome in diagrams and illustrations offers clear ideas.
Q23	Interruption analysis helps me to know where distractions come from.
Q24	In order to keep interruptions out of the process, I amend my learning strategies according to the reasons of interruptions.
Q25	The statistics coming from the use of provided multimedia helps me know my own learning habits.
Q26	The analysis of learning outcome is a good reference for following a schedule design.
Q27	The analysis of all learners improves my learning desire.
Q28	Online analytical dictionary shows me unfamiliar words.
Q29	The analysis of online resource searching record shows me unfamiliar territory.
Q30	The learning analysis of all learners helps me amend my learning strategies.
Q31	The synchronization of learning records is designed with user-friendly interface.
Q32	The analytical tool assists me in understanding my own learning.
Q33	The analytical tool is designed with user-friendly interface.

Chen, H.-P., Lien, C.-J., Annetta, L., & Lu, Y.-L. (2010). The Influence of an Educational Computer Game on Children's Cultural Identities. *Educational Technology & Society*, *13* (1), 94–105.

The Influence of an Educational Computer Game on Children's Cultural Identities

Hsiang-Ping Chen[1], Chi-Jui Lien[2], Len Annetta[3] and Yu-Ling Lu[4]

[1,2,4]Department of Science Education, National Taipei University of Education, Taiwan // [3]College of Education, North Carolina State University, USA // yllu@tea.ntue.edu.tw

ABSTRACT

This study develops an educational computer game, FORmosaHope (FH), to explore the influences that an educational computer game might have on children's cultural identities. FH is a role-playing game, in which children can actively explore a mini-world to learn about science, technology, and society. One hundred and thirty sixth-graders, about 11-12 years old, from four classes in a middle-sized elementary school in Taiwan, participated in the study. A quasi-experimental design was used. The experimental group was two classes that explored FH for a period of six weeks. The other two classes that served as the control group did not receive any experimental treatment. Descriptive statistics, T-test and ANCOVA showed that the experimental group significantly strengthened their cultural identities compared to the control group. This implies that educational games can have an impact on children's cultural identities through their educational contexts.

Keywords

Educational game, Computer game, Cultural identity, Effectiveness evaluation, Individual identity

Introduction

Culture recognition starts from a very early stage of life. Some studies have confirmed that different children hold perceptions about culture differently (Nixon & Comber, 2006). These cultural perceptions not only shape and determine a person's way of perceiving and reasoning (Hofmann, 2006), but also influence one's learning (Nixon & Comber, 2006). Moreover, these cultural identities contribute to group dynamics and the growth of institutes or communities (Ledoux, 2005). Thus, many efforts from a local level, such as a community (Brooklyn Historical Society, 1990), or from national level (Laitin, 1997; Van Gorp & Renes, 2007), have been made to increase the knowledge and understanding of the culture around individuals. In many educational systems, e.g., California State in US and Taiwan, explicitly or implicitly use cultural recognition and respecting different culture as the core of cultural learning (California State Department of Education, 2000; Ministry of Education, 2006).

However, national or state formal curriculums have been doubted for their function on cultural learning. Many studies have maintained that the incorporation of informal approaches into formal learning is a necessity (Kopong, 1995; Ninnes, 1995). It is reasonable to assume that teaching cultural identity may not be successful, if the formal curriculum is the only approach. A value system needs to be built up holistically. Value learning and/or cultural learning need some additional and perhaps more innovative approaches. We believe that giving students opportunities with active involvement and spontaneously emotional attachment, such as those in video games, have more potential to make cultural learning more meaningful than traditional schooling.

Purpose of the study

This study adopted an innovative approach by using a culturally enriched educational game, FORmosaHope (FH), to ascertain the relative effectiveness of gameplay as a form of cultural learning. The research questions of this study are the followings:

1. How effective is the educational game, FH, on Taiwanese student learning of cultural identity?
2. Do students' gender or their family's societal status influence the growth of cultural identities in an educational game environment?
3. How do students feel about the educational game, FH, after they have experienced it?

Theoretical Framework

An important declaration on cultural policies has been made in an UNESCO-based (United Nations Educational, Scientific and Cultural Organization) conference held in Mexico City (World Conference on Cultural Policies, 1982). It asserted that “the cultural identity is a treasure that vitalizes mankind's possibilities of self-fulfillment by moving every people and every group to seek nurture in its past, to welcome contributions from outside that are compatible with its own characteristics, and so to continue the process of its own creation”. This declaration also appealed that “the equality and dignity of all cultures must be recognized” and “the international community considers it its duty to ensure that the cultural identity of each people is preserved and protected (World Conference on Cultural Policies, 1982). Nowadays, developing one’s cultural identity has well recognized as human nature right and protecting this right has also become universal value.

To preserve and to protect this important dimension of culture, an instrument to measure individual’s culture identities was needed. However, after reviewing related literatures, this study found that there is no instrument has been developed. Tracing back to the definition of culture, the Mexico declaration stated culture as, “the whole complex of distinctive spiritual, material, intellectual and emotional features, that characterize a society or social group (World Conference on Cultural Policies, 1982).” This highlighted rough components of “culture.” Another framework that was helpful to establish the component of “cultural identity” was depicted in a booklet of guidelines which was developed to help indigenous peoples to receive equitable culture resources (First Nations and Métis Education Branch, 1995). This booklet examines three portrayals for maintaining qualities of transmitting “culture” and “culture identity.” The three dimensions are “Portrayal of Cultural Interactions”, “Portrayal of Identity”, and “Portrayal of Traditions and Institutions.” After reviewing the design intention and items in each dimension, some characteristics were formulated. First, the “Cultural Interactions” in this booklet stressed contributions of my culture to “other cultures.” Secondly, the “Portrayal of Identity” stresses “what we are”, but not “exchange and interact with other cultures.” Third, the “Portrayal of Traditions and Institutions” was a mix. It examine both “what we are” and “what interactions that we have with others.” Thus, two dimensions could be extracted from above three portrayals: one focused on “my culture” and the other focused on “my culture vs. other cultures.” These viewpoints as well as many studies in Taiwan, those defined cultural identified as spiritual, material, and societal identify (Lee, 1993; Yang, 2001), have provided useful information for further development of instrument for measuring cultural identities.

Cultural identity, in nature, is a matter of perspective-taking. Perspective-taking provides the opportunity to consider others’ viewpoints and induce cognitive conflict. This type of growth and recognition of self does not happen in isolation rather it occurs through the cognitive development of social interactions and/or moral experiences challenging conflict between thought and behavior resulting in more sophisticated, consistent and comprehensive perspective-taking behavior (Hall & Bishop, 2001; Selman, 1977). Through social games and social and moral dilemmas, Selman and Byrne (1974) identified four developmental levels of social perspective-taking that are age-related in a form similar to Piaget’s cognitive operations.

Perspective-taking levels are basic structures of social reasoning and are used in content areas such as interpersonal relations, moral reasoning, social problem solving, communication skills, etc. This is in alignment with the five-step sequence outlined in the practice-oriented Icelandic Project by Adalbjarnardottir (Selman, 2003) which lead class discussions about social conflicts. This approach is designed along the lines of a three-step approach first suggested by John Dewey in the early 1930s. Both approaches were employed by Selman (2003) as a pedagogical practice for social conflicts. In brief social conflicts are approached by: defining the problem; considering the feelings of other’s involved; brainstorming alternative ways to solve the dilemma; choosing a course of action; and evaluating the probable outcome. This implies that intervention research should aim to stimulate perspective-taking through content areas of social reasoning (Selman, 1977).

Gaming is a social construction that varies heavily according to culture, gender, social strata, and the various representations (Brougere, 1999). Video games matter because they present players with simulated worlds; worlds that, if well constructed, are not just about facts or isolated skills, but embody particular social practices. Games bring together ways of knowing, ways of doing, ways of being, and ways of caring, making the player experts in the situated environment. The skills and knowledge are distributed between the virtual actor and the player; hence the values, skills and practices are distributed (Shaffer, Squire, Halverson, & Gee, 2005). In the book, *Got Game*, Beck (2004) stated that Gamers are naturally global: Members of the game generation have been exposed to a multitude of cultures through the gaming experience.

Methodology

This study incorporated a quasi-experimental design consisting of a diverse population and two questionnaires designed to collect information on student understanding of cultural identity and their opinions toward the educational game. This section will explain 1. Participants, 2. The Culture-Enriched Educational Game (FH), 3. Questionnaires, including: The Questionnaire of Cultural Identities (QCI), and The Questionnaire of Students Opinions toward the Educational Game (QSO), 4. Procedures, and 5. Data analysis. These are described in the followings.

Participants

Four classes of sixth-grade students from a mid-size elementary school in Taipei City, Taiwan, participated in this study. The experimental group consisted of 64 students (female, 29 and male, 35) in two classes using FH as a supplemental activity; the control group consisted of 66 students (female, 30 and male, 36) in another two classes proceeding with their normal instruction. Because the school's policy is to try to maintain the equivalence of classes, the class assigning process was based on students' previous academic achievements. From the perspective of students' academic achievements, the four classes maintained equality to some degree, however, there is no evidence to support that equality existed between the experimental and control group, in terms of cultural identities. Thus, a statistical control of variability will be used. The community is located in a residentially and commercially mixed area. Students, in both control and experimental groups, represented a variety of socioeconomic statuses and backgrounds; which made this population viable for this study.

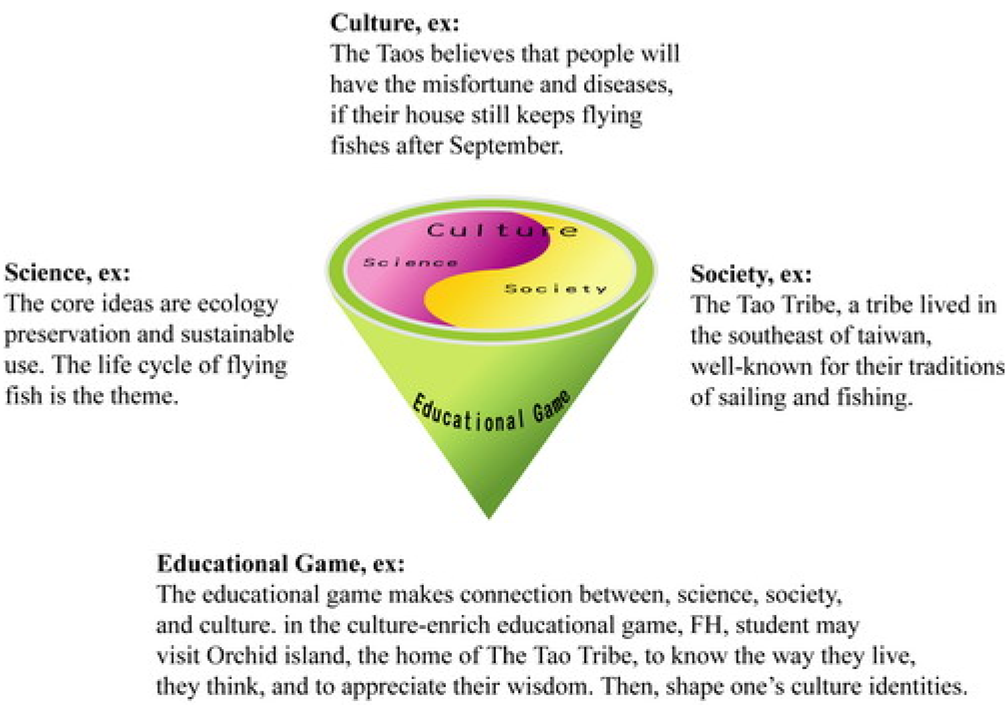

Figure 1. One of the design models used in developing this culture-enriched educational game, FH.

For the purpose of answering the second research question, students were stratified into three groups, the upper, middle, and lower socioeconomic family groups. The criterion for grouping was adopted from a study (Lin, 2001), which is a revised version from a wide-used Hollingshead Index of Social Position (HISP) (Hollingshead, 1965). The Hollingshead grouping system used the education and occupation of the parent and guardian, as indicators. The equation was: HISP score = (Education score X 4) + (Occupation score X 7). Lin's system continued using the equation, however, the seven levels of educations and seven levels of occupations were both degenerated to five levels for simplifying the scoring process. Once, these two factors were assigned numbers from 1-5 (1 for lowest, and 5 for highest), according to level of education and occupation. Then, the Lin's Index of Social Position score was calculated and used as indicator to classifying into group. Scores ranging between 41-55 was defined as upper; 30-40 as middle; and 11-29 as lower socioeconomic status group. After this grouping process, there were 20, 28, and 16 students in the upper, middle, and lower socioeconomic groups, respectively.

The Culture-Enriched Educational Game (FH)

This research group has developed an educational game, FORmosaHope (FH) (Lien, Lu, & Cheng, 2006), targeting 4th through 7th graders. The rationale used for organizing learning in the software is to integrate science, technology, and society in the educational game. A design model used to organize the learning contents and activities in the educational game is shown as Figure 1. In this figure, an example was used to show how the colorful culture, traditional wisdom, and unique society of the Tao Tribe being integrated into a meaningful learning activity in the education game, FH.

The FH consists mainly of two parts. The first one is a role exploration. A player enters the “Village” for a free exploration and may trigger a series of learning events. Figure 2 is one of snapshots of “The Tao Tribe” event which has illustrated as Figure 1. During the exploration, the player may interact with various kinds of people in a virtual society. They also play and gain knowledge about science and technology around them as well as about the society and culture.

Figure 2. Explore and learn in the “Village”

The second part of FH is “Touring Taiwan Island.” Once the player accomplishes some requirement in the village, he or she will be able to visit more than one-hundred towns (stands) in Taiwan (Figure 3). Each will provide opportunities allowing students to learn distinguishable features of that town. The educational content in the game is organized with an inter-disciplinary approach. Students may experience different communities, cities, types of jobs, folklore, and cultures. In the process, students will be able to know, to learn, to appreciate the society and culture, and to grow.

Figure 3. Screenshots of the Touring Taiwan

In addition, there are 11 sub-games to provide many types of learning through game play. These sub-games are arranged randomly and will be triggered when students move to some place or stand. One of the sub-games is shown in Figure 4.

Figure 4. Screenshot from one of the eleven sub-games

Questionnaires

Two types of questionnaire were developed to assess the effects of the FH: The Questionnaire of Cultural Identities (QCI), and The Questionnaire of Students Opinions toward the Educational Game (QSO). The information collected with QCI was used to answer the first and second research questions; the information collected with QSO was used to answer the third research question.

The QCI stated with demographic questions related to student's gender, parent/guardian's level of education and occupation for further analyzing purposes. The construction of the QCI mainly borrowed ideas from the Mexico Declaration (World Conference on Cultural Policies, 1982), Lee (1993), Yang (2001) and The First Nations and Métis Education Branch, Saskatchewan (1995) and modified based on our research emphases. Lee's and Yang's studies used spiritual, material and societal compartments to describe cultural identities. The Mexico Declaration used spiritual, material, intellectual and emotional features to conceptualize the "culture". In QCI, spiritual, and material were adopted and namely, "Spiritual Cultural Identities (SpiCI)" and "Material Culture Identities (MtlCI)". However, considering that the educational game, FH, was heavily science-learning-oriented, the original intellectual and emotional features were replaced to "Scientific Culture Identities (SciCI)." Furthermore, document of The First Nations and Métis Education Branch, Saskatchewan highlighted the importance of societal tradition and institutions, thus, the QCI followed Lee's and Yang's studies and use the name, "Societal Culture Identities (SocCI)." However, there were ambiguities between the SocCI and SpiCI; this was resolved with the ideas derived from the same manuscript. The distinction is that, SocCI in QCI focused on "my culture", while the SpiCI focused on "my culture vs. other cultures."

After the four facets were decided, a total of 26 items were developed and tested. To confirm that the presentation of all items would be suitable to assess the expected cultural identities, eight students represented different learning achievements were assigned to the pilot test and to make sure the items were clear and easy to understand. Students' comments were collected and used to revise the draft. Then, the revised draft was reviewed by experiencing science teachers, social studies teachers and two professors in education to check whether all items matched the goals and to help to revise the questionnaire to enhance the quality of the draft. Once the expert validity of the questionnaire was established, 101 students answered the questionnaire for item analyses and reliability analyses. Three items of twenty-six items did not pass the item analysis, thus were deleted. The final QCI has 23 items in total. Numbers of items in each facet were 5, 7, 5, and 6 for SpiCI, MtlCI, SocCI, and SciCI, respectively. The internal consistency reliability, Cronbach α, of this questionnaire was .84. The result showed that the QCI can be considered as a valid instrument. A description for the four facets, with a sample item, was shown as follows:

- In the SpiCI facet, measuring perception of the extent to which student's confidence in his/her culture when interacting with others cultures, for example, "If I can, I would like to let more people in different parts of the world know more about Taiwan."
- In the MtlCI facet, measuring perception of the extent to which student prefers his/her environment and material life in his/her society, for example, "I feel that imported clothes are better than locally made."
- In the SocCI facet, measuring perception of the extent to which student values his/her society and culture, for example, "I believe that cultures of different ethnic groups in Taiwan are equally fine."
- In the SciCI facet, measuring perception of the extent to which student's confidence in science and technology of his/her society and his/her interest in science, for example, "Science was invented in the western world, and then, passed to us."

The items of the questionnaire were adapted to the Likert format. The numbers from 1 to 5 stands for strongly disagree, disagree, neutral, agree, and strongly agree, respectively. For negative items, scores were converted, so that higher scores always represent student's higher cultural identity on his/her society. The scores of items in each facet are used to indicate the strength of identities in each respective facet; a total score represents the strength of student's cultural identities on his/her society.

The other questionnaire is the QSO. It was designed to understand students' opinions about the educational game, FH. The QSO adapted a two-stage format. In each item, a student needs to make his/her decision either he/she likes or dislikes the educational game, FH. Secondly, the student needs to choose reasons for like or dislike from what researchers have provided in the questionnaire or write down the reasons those are not in the list.

Procedures

This study followed a pre-post test, quasi-experimental design with an additional post-test measure. In terms of cultural identifies, the QCI was used as the pre-, post-test measure, and QSO was used as an additional post-test measure for the experimental group to collect students' overall opinions about the experience of using FH.

Before the intervention, the QCI was administered to all the participants in both groups as a pre-test. During the 6-week treatment period, the experimental and control groups had the same regular curriculum which was based on National Curriculum Guideline (Ministry of Education, 2006). The curriculum included seven major subjects and 33 study hours per week. However, for the experimental group, in addition to their routine classes, students played FH one hour per week for 6 weeks. In the culture-enriched educational game, students traveled in the virtual world to experience the local culture of the real world which was incorporated with science, technology, and society. In the presence of each local cultural and related science event, students learned and responded to specially designed questions or situations. With these, student acted as a tourist, or participant, who was capable of enjoying, learning, appreciating, and even improving the virtual society. After the experimental treatment, the QCI was administered to all the participants in both groups as a post-test to measure participants' cultural identities. Finally, the QSO was used to evaluate this culturally-enriched educational game learning for the experimental group.

Data analysis

The data collected with the QCI and QSO were analyzed with several sets of statistical analyses. For answering the first research question, in addition to descriptive statistics by four facets, one-way analysis of covariance (ANCOVA) was used to compare scores of whole and four facets in terms of cultural identities of the experimental and control groups. Paired t-tests were also used to observe difference of each questionnaire item in QCI, before and after FH intervention.

The second analysis examined the improvement of cultural identities of participants in the experimental group, in terms of different genders and different socioeconomic classes. For this purpose, Two-way analysis of covariance was used. Post-hoc comparisons were used to distinguish differences among socioeconomic family groups. A dependent pre, post t-test was also used to observe the improvement in cultural identity, if needed.

For answering the third research question, descriptive statistics were used. The ratios of different opinions were calculated to be used as the indicator of students' likeness toward the educational game, FH.

Results and Discussion

The means and standard deviations in pre-tests and post-tests of the cultural identity evaluation for both groups are presented in Table 1. The results revealed that mean scores indicated by cultural identities of the experimental group increased in every facet and the total, while for the control group, the total score and scores in each facet did not show significant gains from pre-test to post-test. The pre-post independent t-tests for both control and experimental groups have also demonstrated that the treatment in the experimental group significantly improved students' level of cultural identities in QCI as a whole ($t = 6.956$, $df = 63$, $p = .000$), while a significant level of improvement was not found in the control group. Based on results from previous studies, these findings not only demonstrate the stability of cultural identities and their resistance to change in a traditional educational system based on evidence reported in Table 1, but these results also reveal that a culture-enriched educational game, such as FH, is capable of improving students' understanding of cultural identities.

Table 1. Students' pre- and post-test scores in terms of cultural identities in the experimental and control groups

Measures	Group	**Pre-**: Mean.	S.D	**Post-**: Mean	S.D.
Material (MtlCI)	Experimental	26.36	4.00	28.38	3.81
	Control	24.68	3.88	24.76	4.34
Societal (SocCI)	Experimental	19.53	3.32	21.77	2.70
	Control	19.50	3.16	19.50	3.30
Spiritual (SpiCI)	Experimental	20.08	3.02	22.14	2.64
	Control	20.27	2.90	19.64	3.23
Scientific (SciCI)	Experimental	18.73	4.04	21.56	4.28
	Control	18.97	3.29	19.23	3.55
Cultural Identities (total scores, QCI)	Experimental	84.70	11.06	93.84	10.53
	Control	83.42	9.29	83.12	10.28

Note. Students in the experimental group = 64, the control group = 66.

After confirming that there was no significant difference between slopes of the experimental group and the control group, this study proceeded with ANCOVA analyses. The results, reported in Table 2, shows that there were significant differences between the control and experimental group in all the facets: MtlCI, F = 18.684, p = .000; SocCI, F = 22.641, p = .000; SpiCI, F = 28.096, p = .000; and SciCI, F = 14.625, p = .000. The significant difference of the total score of QCI further demonstrated the difference of improvement between the experimental group and control group, in terms of the growth in cultural identities, F = 43.647, p = .000. In this table, regression effects were all significant (Pre-test, p = .000). The adjusted mean scores of the experimental and control groups were, 93.47 and 83.39, respectively, which revealed that when students were given opportunities of experiencing this cultural-enriched educational game, a significant growth in cultural identities were observed when compared with that of students not exposed to FH.

Table 2. One-way ANCOVA on cultural identities between the experimental and control groups

Measures	Source	df	Sum of Squares	F	p
Material (MtlCI)	Pre-test（covariate）	1	625.489	52.551	.000*
	Between groups	1	222.389	18.684	.000*
	Error	127	11.903		
	Total	130			
Societal (SocCI)	Pre-test（covariate）	1	243.398	33.433	.000*
	Between groups	1	164.831	22.641	.000*
	Error	127	7.280		
	Total	130			
Spiritual (SpiCI)	Pre-test（covariate）	1	146.576	19.143	.000*
	Between groups	1	215.133	28.096	.000*
	Error	127	7.657		
	Total	130			
Scientific (SciCI)	Pre-test（covariate）	1	302.096	22.984	.000*
	Between groups	1	192.225	14.625	.000*
	Error	127	13.144		
	Total	130			
Cultural Identities (total scores, QCI)	Pre-test（covariate）	1	4479.117	60.642	.000*
	Between groups	1	3223.825	43.647	.000*
	Error	127	73.861		
	Total	130			

Note. * for p < .05

Table 3 shows results that students changed before and after the intervention of the educational game, FH, paired t-tests for each item were used.

Table 3. Paired t-tests of mean scores before and after educational game intervention

Items	Means Pre-	Means Post-	T	p
(MtlCI) 1. I prefer Taiwanese traditional food to western fast food.	3.5	4.0	4.42	.00*
2. Comparing with imported computer games, I lack confidence in games that are designed in Taiwan. (-n)	3.5	3.9	2.20	.03*
3. I feel that imported clothes are better than locally made clothes. (-n)	3.0	3.5	2.81	.01*
4. I think, there is no need to keep those historical relics located in valuable city land. (-n)	4.4	4.5	0.82	.42
5. I feel that Taiwan is not as beautiful as many other countries. (-n)	4.4	4.4	0.61	.54
6. We may find and import cheaper goods from abroad, thus, there is no need to develop our traditional industries. (-n)	4.6	4.6	0.36	.72
7. I feel that Taiwan has convenient transportation.	3.1	3.4	2.07	.04*
(SocCI) 1. I feel that we need to preserve and protect different cultures of different people in Taiwan.	4.3	4.6	2.30	.03*
2. I am interested in reading and knowing information concerning local customs and practices.	3.1	4.1	6.63	.00*
3. I hope I can taste local specialties all over Taiwan.	4.3	4.5	1.93	.06
4. I feel there is no need to preserve traditional cultures in my society for they are outdated. (-n)	4.4	4.4	0.34	.74
5. I believe that cultures of different ethnic groups in Taiwan are equally fine.	3.5	4.1	4.47	.00*
(SpiCI) 1. If I could, I would like to let more people in different parts of the world know more about Taiwan.	4.1	4.5	3.47	.00*
2. Respect cultures of different ethical groups in my country will make us more united.	4.3	4.7	3.81	.00*
3. To worship our ancestor is a superior traditional virtue. (-n)	4.0	4.4	3.19	.00*
4. All cultures of different societies need to be preserved.	4.4	4.6	2.47	.02*
5. Invading of foreign cultures through media will eventually extinguish our local culture.	3.3	4.0	3.17	.00*
(SciCI) 1. I am interested in doing science activities.	3.1	3.7	3.58	.00*
2. Learning science can help me to solve my future problems.	3.6	4.2	5.50	.00*
3. I consider that science in my country is not well developed as other countries. (-n)	2.6	3.3	3.37	.00*
4. Technology products' quality of my country is no lower than those of other countries.	3.7	4.1	2.58	.01*
5. I hope to be a scientist when I grow up.	2.6	2.8	1.12	.27
6. Science was invented in the western world, and then, passed to us. (-n)	3.2	3.4	1.15	.26

Note. 1. (-n) represented this item is a negative item. Score was then converted, so that the higher score represented student's higher cultural identity on his/her society; 2. N of experimental group = 64; 3. * for $p < .05$

The above t-tests results showed that means of post-tests for all items were enhanced. Moreover, there were 16 out of 23 showed significant differences. With the findings that cultural identities in the experimental group were superior to that in the control group and the significant improvement demonstrated in each item after the educational game intervention, both strongly supported the position of this study: an informal approach (i.e., through a computer game), other than traditional education is a potential way of improving students' cultural learning.

For answering the second research question, "Do students' gender or their family's societal status influence the growth of cultural identities in an educational game environment", the descriptive statistics of different genders and socioeconomic groups were calculated, and then a two-way ANCOVA tests were implemented. Table 4 presents the means and standard deviations of the cultural identities pre-tests and post-tests for different genders and socioeconomic statuses in experimental group. Table 5 presents the two-way analysis of covariance results for experimental group by gender and socioeconomic status. In the two way ANCOVA, tested for possible differences in the cultural identities by gender and socioeconomic status, the post-test scores were entered as the independent variable while the pre-test score served as a covariate. The result of the analysis revealed significant main effects for gender (F = 4.245, p = .044) and socioeconomic status (F = 5.158, p = .009).

Table 4. Student's pre- and post-test scores of cultural identities by different gender and socioeconomic status in the experimental groups

Source	Group		Pre-test		Post-test	
			M	SD	M	SD
Gender	Male	35	85.71	12.21	96.09	11.14
	Female	29	83.48	9.56	91.14	9.21
Socioeconomic Status	Upper	20	83.85	12.34	97.55	11.13
	Middle	28	85.39	11.25	90.93	10.79
	Lower	16	84.56	9.55	94.31	8.07

Table 5. Two-way ANCOVA on cultural identities in the experimental group: Gender and socioeconomic status

Source	Type III Sum of Squares	df	Mean of Sum of Squares	F	p
Pre-test（covariate）	1877.895	1	1877.895	26.419	.000*
Gender	301.760	1	301.760	4.245	.044*
Socioeconomic Status	733.308	2	366.654	5.158	.009*
Gender * Socioeconomic Status	35.718	2	17.859	0.251	.779
Error	4051.584	57	71.080		
Total	570614.000	64			

Note. * for p < .05

In the gender aspect, the adjusted post-test scores revealed that male students grew significantly greater cultural identities than female students. This result of finding differences between different genders is consistent with many studies. It is believed that this could be attributed to three factors. The first, would this be caused by their attitude of using computer (Cherney & London, 2006; Li & Kirkup, 2007; Ogletree & Drake, 2007)? A large-scale survey study of 10,000 public school students in Texas, U.S. have stated, that, by Grades 4 and 5, girls are more positive in their enjoyment, however, starting about Grade 6, girls' likeness of computers begins to become less positive than boys, and by Grade 8 becomes significantly lower than boys (Christensen, Knezek, & Overall, 2005). Are Taiwanese students in grade 6 in our study already demonstrating the same gender effect? Second, or, this could be caused by the format and design of the game, as previous studies stated that boys prefer role playing games and they play mainly because of curiosity, challenge, competition, and popularity in or of the game. On the other side, girls prefer puzzle games, and they play mainly because of imagination, and relaxation in the game (Chuang, 2002; Hartmann & Klimmt, 2006). We further postulated that to adopt more imaginary and relaxation factors in future educational games as well as to give more attention and encouragement to female students may have more positive effects on their learning in the educational game. Third, there are still possibilities that the gender difference in cultural learning, which has been found, is substantially due to the true effect that female students are more resistant in changing cultural perception, in terms of cultural identities. However, identifying which factors are important and how do they influence learning of different gender students need further studies. It is worth noting that the dependent pre-, post t-test of female students shows significant improvement (t = 3.869, df = 28, p = .001). Based on this finding and the observed improvement in the female student group, when compare with the control group, it is contented that the educational game is still reasonably effective in the female student group.

In the socioeconomic status aspect, the significant differences of means across the upper, middle, and lower groups, requested a post hoc comparison. Although, the differences of means of student from the upper, middle, and lower class families could be easily observed, no significant differences among groups was found except between the upper and middle class groups (p = .002). One notable aspect of these results is that the students from lower societal economics families, who were usually left behind in the learning or in most of innovations or transitions, did as well as the other students from high and middle classes families (p = .133; p = .190). This reveals the potential of an educational game in improving student learning from disadvantaged families or disadvantaged communities. Sociologists affirmed that "stereotype threat effects" occur when members of a stigmatized group perform poorly on a task for they fear confirming a negative stereotype that is associated with their ingroup (Spencer & Castano, 2007). Students from lower societal economics families would more frequently feel threat from traditional learning and evaluation. In the learning context of the educational game, FH, students feels they were playing while they were learning. This may have reduced the "stereotype threat effects" and helped their learning.

For answering the third research question, "how do students feel about the educational game, after they have experienced it", the data were collected with the QSO. Students' opinions about FH are shown in Table 6. This analysis revealed that the majority of participants in the experimental group liked this educational game (88%). They felt this cultural-enriched educational game was novel and interesting (88%). By playing the game, they felt they may learn science (86%), while experiencing the wonder of culture (80%), in this cultural-enriched educational game, FH.

In addition, students also provided some other opinions about the effectiveness of FH. This included such statements regarding: characters in the game are cute; it is easy to learn while playing; the game is challenging; it gives me a sense of accomplishment; the game makes me enjoy my playing, the game is creative, etc. However, some others opinions include: too much talking, need to have more characters, some parts of the game are difficult.

Table 6. Students' opinions about their computer game learning experiences in FH

	Preference and Reasons	N	Percentage
	Like	56	88%
Reasons	I feel the FH educational game is novel and interesting.	49	88%
	By playing the FH, I recognize more about the wonder of our cultures.	48	86%
	By playing the FH, I learn more about the science.	45	80%
	Dislike	8	12%
Reasons	I do not think that the FH computer game is novel and interesting.	5	63%
	By playing the FH, I do not think it help me to recognize the wonder of our cultures.	3	38%
	By playing the FH, I do not think it help me to learn more about science.	4	50%

From the above quantitative and qualitative data, the computer game is a potential way to learn which is different from tradition education and the culture-enriched educational game, FH, is effective in promoting students' learning in terms of cultural identities.

Limitations of Study

Culture is truly a complex whole (Tylor, 1920). The complexities have made the definition nearly impossible. Moreover, identities involve sophisticated psychological factors. It's influenced by one's previous experiences, personal characteristics, and societal interactions, etc. Thus, to design an instrument for measuring this quantity with high variability is difficult in nature. To some extent, we expected that some uncertainty is inherent in the measurement of cultural identities in this study. Authors of this study believe that the measurement of cultural identities can be very sensitive to different culture and different ways of seeing what cultural identities are, thus, suggest that QCI needs to be further modified for their special context and purposes.

Conclusions

Ladson-Billings (1995) defines culturally relevant pedagogy from the student perspective. Three propositions are stated: students experience academic success; students develop/maintain cultural awareness; and students develop a critical consciousness to question the status quo. The culture-rich educational game, FH, attempted impact the three aforementioned propositions.

A simulation “approximates a real world setting” and provides a complex virtual learning environment (Reigeluth & Schwartz, 1989; Winn, 2002). Simulations can be utilized to enhance student perspective-taking experiences by providing the opportunity to consider others’ viewpoints and induce cognitive conflict. This type of growth and recognition of self does not happen in isolation rather it occurs through the cognitive development of social interactions and/or moral experiences challenging conflict between thought and behavior resulting in more sophisticated, consistent and comprehensive perspective-taking behavior (Hall & Bishop, 2001; Selman, 1977).

Through social games and social and moral dilemmas Selman and Byrne (1974) identified four developmental levels of social perspective-taking. Perspective-taking levels are basic structures of social reasoning and are used in content areas such as interpersonal relations, moral reasoning, social problem solving, communication skills, etc. Selman (1977) implies that intervention research should aim to stimulate perspective-taking through content areas of social reasoning. Consistent with Piaget and Kohlberg, Selman (2003) believed individuals form their ways of thinking through social experiences which help to influence our thinking about morality, justice, and fairness. It is a move from how to understand oneself to how one actually relates to others. As Norman (1998) suggested, individuals encounter cognitive conflict when experiencing different cultural. By their nature, video games provide conflict which we posit creates even greater cognitive dissonance when learning about different cultures.

In this study, we demonstrated the feasibility of adopting cultural learning in an educational game platform. Major results of this study also provides evidence that FH has significant effect on students’ cultural learning and is helpful in some degree to students from lower socioeconomic families as compared to students from middle and upper socioeconomic families. Thus, this study concludes that learning in affection domains, such as attitudes, value judgment, or culture, etc. should not be counted on the formal educational system only, the informal education in various formats as well as those may happen outside class needs to be given more weight. A well-developed educational game, in addition to their potential for knowledge learning and entertainment, can promote the growth in the affection domain. Through the equal learning effectiveness in different socioeconomic groups, educational games are also helpful in attaining the principle of equal learning opportunity for all. The full potential of the educational games is out there and just waiting to be discovered

Acknowledgements

The authors wish to thank specialists, students, Professor James A. Shymansky, Professor Chao-Ti Hsiung, and Ms. Chien-Ju Li for their valuable efforts, opinions and help. This study was supported by grants from National Sciences Council, Taiwan (Projects: NSC 93-3111-P-008-001-Y19, 92-2524-S-152-001, 152-003, 260-003, 94-2524-S-152-002, 96-2511-S-152-004-MY3).

References

Beck, J. C., & Wade, M. (2004). *Got game: How the gamer generation is reshaping business forever*, Boston, MA: Harvard Business School Press.

Brooklyn Historical Society. (1990). *Many faces, many ways. Multi-cultural diversity of Brooklyn. A guide for teachers*. New York: Brooklyn Historical Society.

Brougere, G. (1999). Some elements relating to children's play and adult simulation/gaming. *Simulation & Gaming, 30*(2), 134-146.

California State Department of Education. (2000). *Updated history-social science framework,* Sacramento, CA: California Department of Education.

Cherney, I. D., & London, K. (2006). Gender-linked differences in the toys, television shows, computer games, and outdoor activities of 5- to 13-year-old children. *Sex Roles, 54*(9), 717-726.

Christensen, R., Knezek, G., & Overall, T. (2005). Transition points for the gender gap in computer enjoyment. *Journal of Research on Technology in Education, 38*(1), 23-37.

Chuang, Y.-Y. (2002). *The relationship of children's computer game behavior, creativity, and loneliness of Taipei's upper elementary school.* Unpublished Master Thesis, Chinese Culture University, Taipei (in Chinese).

First Nations and Métis Education Branch, Saskatchewan (1995). *Diverse Voices: Selecting Equitable Resources for Indian and Métis Education*. Retrieved May 31, 2009, from http://www.education.gov.sk.ca/adx/aspx/adxGetMedia.aspx?DocID=244,234,140,107,81,1,Documents&MediaID=518&Filename=diversevoices.pdf.

Hall, A., & Bishop, R. (2001). Teacher ethics, professionalism and cultural diversity. *New Zealand Journal of Educational Studies*, *36*(2), 187-202.

Hartmann, T., & Klimmt, C. (2006). Gender and computer games: Exploring females' dislikes. *Journal of Computer-Mediated Communication, 11*(4), 910-931.

Hofmann, S. G. (2006). The importance of culture in cognitive and behavioral practice. *Cognitive and Behavioral Practice, 13*(4), 243-245.

Hollingshead, A. (1965). *Two-factor index of social position*, New Haven, CT: Yale Station.

Kopong, E. (1995). Informal learning: a case study of local curriculum development in Indonesia. In Tedesco, J. C. (Ed.), *Prospects: quarterly review of comparative education, 25*, 639-652.

Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory Into Practice, 34*(3), 159.

Laitin, D. D. (1997). The cultural identities of a European state. *Politics & Society, 25*(3), 277.

Ledoux, M. W. (2005). Institutional mission and identity: How do we carry the culture to the electronic forum? *Educational Technology & Society, 8*, 191-197.

Lee, H.-C. (1993). Nationlism and cultural identities. In N. Polzer (Ed.), *Nationlisium*. Taipei: Li Ming Cultural Enterprise. (in Chinese)

Li, N., & Kirkup, G. (2007). Gender and cultural differences in internet use: A study of China and the UK. *Computers & Education, 48*(2), 301-317.

Lien, C. J., Lu, Y. L., & Cheng, S. H. (2006). An educational software game for learning science and society- "The Formosa Hope". *Elementary Education, 46*(3), 9-15. (in Chinese)

Lin, S. C. (2001). *Educational Sociology*, Kaoshiung: Fu-Wen Publishing (in Chinese).

Ministry of Education, Taiwan. (2006). *General guidelines of grades 1-9 curriculum for elementary and junior high school education in Taiwan*. Retrieved May 31, 2009, from http://english.moe.gov.tw/public/Attachment/66618445071.doc

Ninnes, P. M. (1995). Informal learning contexts in Solomon Islands and their implications for the cross-cultural classroom. *International journal of educational development, 15*(1), 15-26.

Nixon, H., & Comber, B. (2006). Differential recognition of children's cultural practices in middle primary literacy classrooms. *Literacy, 40*(3), 127-136.

Norman, A. J. (1998). Managing conflict building a multicultural collaborative. *Cities, 15*(3), 209-214.

Ogletree, S. M., & Drake, R. (2007). College students' video game participation and perceptions: Gender differences and implications. *Sex Roles, 56*(7), 537-542.

Reigeluth, C., M., & Schwartz, E. (1989). An instructional theory for the design of computer-based simulations. *Journal of Computer-Based Instruction, 16*(1), 1-10.

Selman, R. L. (1977). A structural-developmental model of social cognition: Implications for intervention research. *Counseling Psychologist, 6*(4), 3-6.

Selman, R. L. (2003). *The promotion of social awareness: Powerful lessons from the partnership of developmental theory and classroom practice*, New York: Russell Sage Foundation.

Selman, R. L., & Byrne, D. F. (1974). A structural-developmental analysis of levels of role taking in middle childhood. *Child Development, 45*(3), 803-806.

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan, 87*(2), 105-111.

Spencer, B., & Castano, E. (2007). Social class is dead. Long live social class! Stereotype threat among low socioeconomic status individuals. *Social Justice Research, 20*(4), 418-432.

Tylor, E. (1920). *Primitive culture*, New York: J.P. Putnam's Sons.

Van Gorp, B., & Renes, H. (2007). A European cultural identity? Heritage and shared histories in the European Union. *Tijdschrift voor Economische en Sociale Geografie (Journal of Economic & Social Geography), 98*(3), 407-415.

Winn, W. (2002). Current trends in educational technology research: The study of learning environments. *Educational Psychology Review, 14*(3), 331-351.

World Conference on Cultural Policies, UNESCO. (1982). *Mexico City Declaration on Cultural Policies*. Retrieved May 15, 2008, from http://portal.unesco.org/culture/en/files/35197/11919410061mexico_en.pdf/mexico_en.pdf.

Yang, S. F. (2001). Multi-dimension national identities. *Dr. Sun Yat-Sen Academic Journal, 22*, 175-191 (in Chinese).

Hannon, J., & Bretag, T. (2010). Negotiating Contested Discourses of Learning Technologies in Higher Education. *Educational Technology & Society*, *13* (1), 106–120.

Negotiating Contested Discourses of Learning Technologies in Higher Education

John Hannon and Tracey Bretag*
Curriculum, Teaching and Learning Centre, La Trobe University, , Melbourne, Victoria 3086 // Tel: 61 3 9479 1533 // J.Hannon@latrobe.edu.au
*School of Management, University of South Australia, Adelaide, South Australia 5001 // Tel: 61 8 8302 0224 // tracey.bretag@unisa.edu.au

ABSTRACT

This paper explores the way that learning technologies frame teaching practice in higher education using both autoethnography and discourse analysis (interpretative repertoires). The analysis juxtaposes our own experience in the form of data from two interviews, with teaching and learning policy documents from the group of five Australian Technology Network universities, as a means of investigating the centrality of these technologies in the reconfiguring of teaching practice in higher education for the networked university. The data yielded three distinct discourses: technology as a bridge to globalised opportunity; technology as delivery of learning; and technology as communication and building relationships for learning. The first repertoire provides a utopian vision which glosses over the complex practice of implementation. The second repertoire also omits details of implementation, presenting learning technology unproblematically. The third repertoire, not present in the policy documents, but central to the autoethnographic accounts, focusses on both the possibilities and challenges of learning technologies in practice, and points to the potential for a complementary approach which foregrounds the student-teacher relationship. How these discourses can be reconciled is a central issue for academic teaching practice in higher education.

Keywords

Learning technologies, higher education, policy, practice, discourse

Introduction

The current climate in higher education

Soucek (1994) suggests that the function of tertiary institutions has changed since the 1980s “from guardianship of knowledge and wisdom to ancillary production of knowledge for corporate capital” (p.54). During this time there has been a redefinition of the role of the teacher, “from progressive educator and participant in educational politics to one of competent performer of relatively neutral tasks related to efficient and profitable delivery of pre-specified curriculum, and of being a responsible manager of learning contexts” (Seddon, 1998, p.5). There is widespread concern at the effect that “supermarket” policies have had on teachers’ professional lives (Moloney, 2000, p.73). In the current climate few academics have been openly critical of policies within their own institutions, although according to Gaita (2000), far from betraying the institution, such open criticism is the mark of “the true champions of a university” (p. 41). The rhetoric that accompanies online learning at the policy level does not necessarily match actual practice (Conole, de Laat, Dillon, & Darby, 2008, p. 511). With networked technologies central to the changing nature of universities (Lewis, Marginson & Snyder, 2005; Cornford & Pollock, 2003, p. 6), competing agendas concerning online learning have emerged within institutions, with consequences for practice (Lewis et al., 2005, p. 72). It is our contention that when educators are confronted with policies that potentially disregard both the learning needs of diverse students and the recent research on teaching and learning, we have a responsibility to engage with those policies, interrogate them, and make a space for constructive debate.

The contest over online learning

Teaching and learning in higher education across the globe is said to be undergoing profound change and transformation by the technologies of learning (Castells et al., 1999; 1996; Kellner, 2003), and organisations have invested in “big” solutions with mixed results. Others argue that online learning, or “e-learning” has heralded “successive false dawns” (McMullin, 2005, p. 67) or “stalled” (Zemsky & Massey, 2004). Pollock & Cornford (2002) discussed three failed online learning projects, and identified the issues contributing to this as not the technology itself, nor any negative attitude of staff, rather “the underlying problem is the sheer volume and complexity of the work required to configure people, machines, objects, texts, and money” (p.371).

This complexity does not seem to be acknowledged in the implementation of institution-wide implementations of online learning (Barnett, 2000; Hannon, 2008), particularly in its demands on people, time and resources. While online learning may be still in a developing phase, its technologies are integral to the reconfiguration of both pedagogy and the organisation of learning. Learning technologies bring with them interests, internal and external stakeholders and a politicised "covert curricula" (Roberts, 2004).

Teaching staff in universities are caught between two trajectories: on the one hand, the promise of the economies of scale with "big" solutions in online learning, with heavy investment in proprietorial online learning, huge commitment of resources and support, and on the other, teaching and learning practices opened up by innovative uses of learning technologies. Lewis et al. (2005) described these two perspectives as inherent in the "networked organisation" (p. 72), from which emerged contested discourses within institutions (p. 72). They noted how networked technologies in organisations did not follow a single logic, but were "socially embedded and therefore highly variable" (p. 71).

With the online technology administered and managed by an information technology unit, there is the risk that "learning experience takes a back seat to the management functions" (Siemens, 2006). The focus of learning management systems (LMS) on "content" and the management of students places pedagogy and engagement in the background. Hotrum (2005) observes that LMS are "progressively being regarded as a hindrance to effective online learning." With the emergence of easy to use collaborative Web services, Hotrum suggests that "a new generation of Web-based tools and approaches is evolving that are better suited to meet the need for dynamic online learning content, interaction, collaboration, and networking." The transformative potential of "social software" marks a clear shift away from pedagogies based on managed "content" and repositories of learning objects towards student-centred learning focussed on knowledge production (Gibbs & Gosper, 2006; McLoughlin & Lee, 2008).

The production of instructionally designed online content is one of the trajectories of learning technologies, in which curriculum is relatively fixed, and produced for delivery by many instructors to learners en masse. The other trajectory is the opportunities provided by networked communication to treat curriculum as always unfinished, as "content for meddling with" (McWilliam, 2005). Each trajectory points to different structural arrangements for teaching and learning in institutions, with different notions of learning, of expertise and of the division of pedagogical labour. As the shape of online teaching and learning is still in flux, its contested nature is being played out in a struggle of discourses, where online learning is being shaped by the terminology and the framing notions of institutional, technological and pedagogical interests.

Two examples of the contested discourses of the "networked university" are described by Lewis et al., (2005), where a concern with centralised, standardised control of teaching with technology, co-existed with an interest in the collaborative and democratic use of technologies (pp.72-3). They noted these discourses, though "highly contested", were not necessarily opposed. A contentious recurring issue for academics concerns how academics negotiate governance of their day to day work, caught between the opposing poles of collegiality and managerialism (Marginson & Considine, 2001). This inquiry takes up an instance of this discussion.

Our research agenda

In this paper, our intention was to follow the mismatch suggested by Conole et al. (2007) between the rhetoric of online learning and its practice, by selecting and contrasting discourses from both these realms. Our questions were: how do rhetoric and practice connect or fail to connect? And what are the underlying issues or dilemmas in each discourse? Both discourses concerned online learning in the changing environment of higher education: the realm of rhetoric is represented by the institutional policies from the five Australian Technology Network (ATN) Universities that reflect a similar orientation towards learning technologies; the realm of practice is represented by the interview accounts of the two authors about their own experience with learning technologies. Our intention was to select a case of authentic, situated practice that represents the typical enactment of those teaching and learning policies in higher education practice, that of teaching in a mass learning environment enabled by networked technologies. Our decision to deploy autoethnographic interview accounts by the authors, arose, in part, from experiences in our day-to-day practice of being confronted with issues of large class teaching, approaches to tutoring online groups, reports of intensification of work, all of which raised the ambiguous role of networked technologies and the way they occurred in institutional policy texts and in practice. Our intention was to provide an "emic" or insider perspective on how we negotiated our own pedagogical practice with institutional imperatives in this networked environment. We did not intend to make universalising claims, but to enter this contested space, and track our personal judgements about online pedagogies and their intersection with institutional policy, on the assumption of relevance: that other teaching academics also encounter pedagogical dilemmas at the intersection of policy and practice in the new higher education environment. While the setting and participants in this study were local, the learning technologies, the policy context, and the teaching and learning practices are replicated globally in higher education. In this paper we bring a innovative approach to analysing practice by deploying a type of discourse analysis which bridges local and global contexts. This strategy problematises discourses of online learning, aims to expand the agency of teaching academics, and thereby reclaim our practice.

Local context

Our current interest in online learning technologies occurs in a particular policy context. In 2004 the Vice Chancellor at The University circulated a discussion paper to all staff entitled *Towards an online strategy 2005-2015*. We refer to the institution in which we worked during the period of research as "The University" because we do not wish to imply that the situation there was unique in terms of a general push towards online learning in Australia, or indeed across the higher education sector internationally. The discussion paper in question documented a vision of how The University would use new technologies in its service to staff, students, alumni, prospective students and partners. At the heart of this vision was the desire to develop a sustainable, customer-focused, competitive, cost-effective, standardised and disaggregated teaching and learning framework that made use of e-learning and online technology. In particular, the online strategy advocated the use of "learning objects" – decontextualised online content that could be shared within and outside The University – as a means of increasing efficiency and reducing costs. While the paper reiterated a commitment to "high levels of customer service and well-developed customer relationship management strategies" (*Towards an online strategy* 2004, p.3) there was no mention of the difficult-to-measure, but nonetheless valuable, student-teacher relationship.

Data and methodology

A mixed methodology was designed to match the discourses of the public, institutional policies of teaching and learning on the one hand with those of academic practice on the other. Following the researcher-practitioner tradition, an "inside-out" approach (Mann, 2004, p. 206) was used, in which personal enquiry into practice combined with critical reflection becomes "a valid source of knowing" (p. 207). Our approach juxtaposed this with an "outside-in" (p. 207) perspective which applied empirical approaches of coding and categorising to all the data, using a discourse and content analytic approach to analyse the intersection of policy and practice, that is, draw analytical comparisons between the public and personal discourses. This approach located the researchers in the enquiry setting with a focus on practice settings, "both in terms of the *discourse* in which practices are described and understood and in terms of socially and historically constructed *actions and their consequences*" (Kemmis & McTaggart, 2008, p. 292).

Two sources provided the data for analysis of these discourses: public documents on teaching and learning policy from five ATN universities, and interview transcripts from two academics who had been working at one of the ATN universities: Bretag (interviewed on 5 August 2005), a lecturer teaching communication courses in a business faculty, and Hannon (interviewed on 26 April 2007), an academic developer based in the university teaching and learning unit. This work is part of a larger project which began in 2005 when Hannon interviewed Bretag for his doctoral research. In early 2007, after conducting research together on the use of computer mediated communication to develop a community of learners (Bretag & Hannon 2008), we decided to revisit the initial interview with the idea of reversing our roles. Using the same interview questions, Bretag interviewed Hannon (see Appendix 1). We then contextualised our analysis of the interview data with a content analysis of five ATN university policies relating to learning technologies. From both these sources, patterns and themes of discourses were analysed.

Teaching and learning strategy documents were from the five ATN universities (Curtin University, University of South Australia, Royal Melbourne Institute of Technology, University of Technology Sydney and Queensland University of Technology). These universities are grouped together under the ATN banner because of their shared development history from colleges to universities in the 1990s, and because of their common focus on "undertaking solution based research", underpinned by a commitment to equity and access (Australian Technology Network of Universities). The five ATN documents were:

- RMIT 2010: Designing the Future, (RMIT University, 2007)
- Curtin University of Technology, Teaching and Learning Enabling Plan, (Curtin University of Technology, 2006)
- Queensland University of Technology, Approach to and context of teaching and learning, (Queensland University of Technology, 2007)
- University of Technology, Sydney, Setting the Pace 2006-2009: Strategic Directions for the Current Decade (of Technology, Sydney, 2007)
- University of South Australia, Teaching and Learning Strategy 2006–8, (University of South Australia, 2007).

There were three stages of analysis. Both ATN documents and interview data were analysed using the qualitative analysis software program N6. The first stage of analysis used a simple content analysis by weighting to indicate the keywords most used relating to issues of higher education teaching and learning with technologies. The second stage involved coding sentences and paragraphs to identify the concerns and preoccupations of authors and speakers, then the coded units were organised into themes or categories. In the third stage, these categories of most concern were examined for the regularities in the accounts as they reflected attempts to resolve particular dilemmas, or achieve certain ends. These were identified as interpretative repertoires, for which an associated vocabulary and rhetorical techniques were deployed.

Autoethnography

Rather than seeking to generate a "theory", the aim in this research was to juxtapose policies relating to learning technologies with our own experience, as a means of exploring and participating in this contested field of inquiry. As practitioner researchers we wanted to situate our own and each other's experience within the broader policy environment, using autoethnography. Patton (2002) describes autoethnography as "self-awareness about and reporting of one's own experiences and introspections as a primary data source" (p. 86), and Ellis and Bochner (2000) define it as "writing and research that displays multiple layers of consciousness, connecting the personal to the cultural" (p. 739). Research which utilises autoethnography may include personal narratives, first-person accounts, personal essays, opportunistic research, self-ethnography, and autobiography to mention just a few approaches (as cited in Ellis & Bochner, 2000, p. 739). In keeping with the tenets of qualitative research (see Cresswell, 1998), at times we strategically change to the first person in a conscious attempt to foreground our own positions and experiences. While the personal revelation used in autoethnographic research has been critiqued as a potential risk to the integrity of research, Bruni (2000) counters this view by offering a revisioning of research ethics protocols based on the uniqueness of each research enterprise (p. 30). We would argue further that in foregrounding our perspectives and context we make visible what is rarely apparent to the reader, our own agendas. We present ourselves as insider researchers, and claim the benefit of prolonged and not disinterested, but critical engagement in the research context. This engagement acts as an aid to validity and has also provided us with access to knowledge which may be undiscoverable to outsiders (see Edwards, 2002).

We follow the example of Adams (2007), an ex-nurse, who interviewed a number of nurses to explore their everyday work experience, and then used the same questionnaire to "provide a similar and comparable record of [her own] experience" (p. 10). Adams then coded the autoethnographic interview using the same process used in the other interviews. Like Adams, having foregrounded our own insider status, we interviewed each other using the same interview structure, and then coded the interviews to reveal "interpretative repertoires", as described in the following section. In this research, autoethnographical material (as exemplifed in the coded interviews) was contrasted with the coded policy documents of the five Australian ATN universities. By insisting on the legitimacy of our own experience, and applying the same rigorous analysis to it, we have responded to Smyth and Hattam's (1998) call for academics to be "reflexively engaged" in policy issues which directly impact on our practice.

Discourse analysis using interpretative repertoires

A discourse analysis approach to spoken or textual accounts focuses on language use: rather than analysing social representations or cognitive processes, it is concerned with how utterances and accounts are constructed, and the social effects of what they may represent. Discourse can be viewed as a “social practice” (Wetherell & Potter, 1988, p. 168), and as having an “action-orientation”, with “particular consequences” (p. 171). These effects range from actions in local contexts, the “speech acts” of Austin (1962), where talk may be described as accusing, excusing, justifying, and so on, to accounts which may be described as having broader effects such as arguing a particular institutional position.

Much discourse analysis occurs in local settings, focused on the specifics of an encounter, and one of the criticisms and challenges of such analysis is how to make connections beyond the situated context (Alvesson & Karreman, 2000, p. 1127). In our methodology we wished to bring an “outside-in” perspective to both sets of data, the author interview transcripts and policy documents, and make the link between the local, conversational, “transient” discourse to the extended, durable, “meanings ‘existing’ beyond” (p. 1130). We selected the particular discourse analytical technique of “interpretative repertoires”, developed by Potter and Wetherell, (1987), to make this link between the local setting and the social world of practice: in our study, to connect the discourses of the localised interview transcripts and public policy documents. An interpretative repertoire was defined by Wetherell as: “a culturally familiar and habitual line of argument comprised of recognisable themes, commonplaces and tropes” (Wetherell, 1998, p. 400). Hence speakers and writers deploy rhetorical and lexical patterns to accomplish particular tasks. An analysis of interpretative repertoires can explore the discourses individuals draw upon to construct their social world through talk and texts in order to achieve particular ends. An example of an interpretative repertoire may be the use of the term “flexibility”, and the effects this produces when used in powerful policy documents.

Interpretative repertoires tend to be deployed to resolve a dilemma in the social context for the speaker/writer (Potter & Wetherell, 1987, p. 152). In the current study, repertoires were analysed to identify in the authors’ own talk or text how they attempted to resolve contextual dilemmas in their practice and accomplish particular tasks, and their use of specific rhetorical techniques to do this. We analysed the discourse of each of our interview transcripts, identified repertoires in these accounts, and compared these with the repertoires expressed in the ATN university teaching and learning strategy documents.

Findings

The two sets of data (ATN policy documents and authors’ interview transcripts) were coded using the qualitative software N6. The findings below consist of content analysis for keywords, coding into categories, and interpretative repertoires identified from both sets of data.

Keywords

The data was examined for recurring keywords relating to how institutions and academics engage with learning technologies. Total occurrences of keywords, truncated for variations using an asterisk, are shown in Table 1: Occurrences of key terms.

Table 1: Occurrences of key terms

Key terms	ATN documents (Total 25,696 words)	Authors’ interview transcripts (Total 19,510 words)
improv* (improve, improving, improvement)	135	6
communit* (community, communities)	57	10
engag* (engage, engagement)	43	7
respons* (response, responsible, responsibility)	41	8
techn* (technology, technological, technical)	37	50
flexib* (flexible, flexibility)	28	0
online	27	75
value	23	6
innovate* (innovate, innovation)	12	3

The key terms mentioned most frequently were:

- For the five ATN institutional policy documents: improve/improving/improvement (135 occurrences), followed by community/communities (57), then engage/engagement (43 occurrences), technology/technological/technical (37), flexible/flexibility (28), online (27).
- For the authors' interview transcripts, two keywords were mentioned with much greater frequency than all others: online (75), and technology/technological/technical (50).

Categories

The process of coding the texts from both data sources resulted in a total of 29 codes, with the coding unit being the paragraph. The codes were grouped into four categories in response to the question: what issues of concern are expressed around use of technologies for learning? These are shown with their coding frequencies in Table 2: Categories and codes for each data source.

Table 2: Categories and codes for each data source

Coding	ATN documents	Authors' interview transcripts
Academics		
Attitude to new technologies	0	15
Changing role	0	13
Critique	0	15
Engagement with technology	0	11
Innovation	0	13
Interculturality	0	1
Leading	0	9
Multiple roles	0	15
Research	0	7
Sharing and mentoring	0	10
Teaching	0	25
Time constraints	0	15
Institutional context		
Flexibility (of learning arrangements)	35	6
Institutional requirements	0	19
Institutional support / Workload	16	4
Policy vs practice	0	27
Principles & Values	50	0
Strategy	111	0
Strategic Actions	25	0
Online technologies		
Communication issues	0	17
Complement to face-to-face	0	24
Embodiment	0	17
Limitations	0	16
Marginalisation	0	17
Technophobia	0	4
Types of e-learning	6	14
Students		
Engagement with technology	0	11
Student needs	0	21
Student-teacher relationships	0	39

The four categories were ordered to reflect the density of coding, shown in Table 3: Categories from data sources.

Table 3: Categories from data sources

ATN documents	Coded units	Authors' interview transcripts	Coded units
Institutional context	237	Academics	149
Online technologies	6	Online technologies	109
Academics	0	Students	71
Students	0	Institutional context	56

The order of categories in Table 3 provides an indication of the relative emphasis each category or issue was given by the authors of the ATN documents, compared to our interview transcripts: the institutional context was the predominant concern of the ATN documents, comprising 97% of codes for that data source; and for the interview transcripts, the main concerns were academics (39%) and online technologies (28%). A striking feature of the comparison of codes in Table 2 was that the concerns expressed by us in our interview transcripts were shared with only three codes with the ATN policy documents: flexibility, types of e-learning, and institutional support.

A strong disparity between policy documents and interview transcripts emerges from the coding. The two primary codes for the ATN documents, strategy, and principles and values, comprised 66% of codes for that data source, and were not coded at all in the Hannon and Bretag interviews. The coding for the two interviews encompassed a total of 26 codes, the main concerns indicated by five codes with over 20 coded units, which comprised 36% of codes for the transcripts. These five codes were: *teacher-student relationships* (39 units), *teaching* (35), *policy vs practice* (concerns with institutional pressure on teaching and learning practice) (27), *complement to face-to-face* (how technologies complement face-to-face learning settings (24), and *student needs* (21).

Interpretative repertoires

Interpretative repertoires were identified first by focusing on those categories which reflected the greatest concern expressed in the interviews and policy documents, then by locating specific uses of text or talk that attempted to resolve uncertainties or dilemmas of practice. Three interpretative repertoires were identified, or ways of writing or talking about the technologies of teaching and learning in higher education in the data sources. These were:

- technology as a bridge to globalised opportunity;
- technologies as delivery of learning; and
- technologies as communication and building relationships for learning.

The first repertoire was used exclusively by the ATN teaching and learning policy documents, and presents an institutional perspective on technology in higher education, where the reader is offered, via the institution, access or a bridge to a networked world. This repertoire reflects the concerns of the ATN documents coded as *strategy, principle and values, flexibility*. The remaining two repertoires were both drawn on by Hannon and Bretag in their interviews, but in contrasting ways: technologies as delivery of learning, and technologies as connecting people to build relationships for learning. These correspond to the areas most frequently coded for the interview transcripts: teacher-student issues, technologies and learning, and concerns with policy and practice. Each repertoire is discussed below.

Analysis and discussion

Technologies as a bridge to globalised opportunity

The ATN teaching and learning strategy documents reflected the imperatives for universities to position themselves in a global market, the need to respond to the demands of learners, and the fluid nature of the times and spaces of the learning environment. Keywords used frequently were variations on "improve", "community", "engage", "response", "technology", "flexible" (see Table 1). These were associated with terms such as global passport, global engagement, connected, strengthen, incorporate, provide a real-world education, best practice, appropriate technology, define our global network. These keywords and terms located the university as an active agent in a disaggregated, global field of endeavour.

Technology was offered to prospective students as a key which offers access to the world:

> technological advances open up opportunities (QUT)
> provision of resource-rich, technologically-mediated forms of delivery that enable access (UniSA)
> Our flexible learning environment and effective use of technology in teaching and learning will underpin the University's reputation for excellence in the facilitation of learning (UTS)

In the teaching and learning strategy documents, technologies both underpinned and enabled institutional goals. The authors of the ATN documents consistently framed a field of operations which reflected a volatile, globalised and unpredictable world of shifting markets and demands, "major challenges" in global education, and competition with other educational institutions. Having constructed such a challenging environment, a response was proffered by institutions in the form of self-descriptions: they were adaptable, internationalised and technologically cutting edge, and were able to offer their audience global reach, and access to opportunities in this unpredictable world. The rhetoric of "community" was linked with work and business:

> engage students and staff with the professions, industry, business and the community to maximise opportunities (UTS)
> a global university grounded in Melbourne and connected to communities, enterprises and industry across the world (RMIT)

The use of metaphors of connection and community was associated with descriptions of a global, competitive, dispersed and internationalised field. The response of the institutions was to embrace this dynamic world:

> shaping the educational foundations of existing and emerging fields of practice (UTS)
> reflect our global engagement with industries and communities. (RMIT)
> research and teaching will be conducted in many countries through innovative use of e-learning and e-business (UniSA)

Institutions were positioned as locations of access to a complex global field, with access offered through technologies and international networks. Potential dislocation of the times and spaces of learning was resolved with a virtuous rhetoric associated with technology:

> Technological advances open up opportunities for adding newer and more innovative methods to the spoken lecture and the face-to-face seminar. (QUT)
> The University will be a leader in global access to learning that is enabled by emerging technologies. (UniSA)

The uncertainty of the "complex workplace and community of the 21st century" (UTS) is made palatable by the keywords "community", "engage", "response", "innovative". Technologies, then, are a key to a dual network, technical and social: one offers the reader (presumably a student) a network which provides "technologically-mediated forms of delivery", the other offers social access and connection to this high-tech, utopian community.

The policy documents used the metaphor of access or the "bridging" repertoire to orient the reader to an imagined high-tech world beyond the institution, and was not deployed at all in the interview transcripts. While the interviews did, in fact, reflect institution concerns, in the codes *policy and practice*, and *institutional requirements*, these concerns were oriented towards adapting teaching and learning practice to the institutional context; that is, they were inward looking. The contrasting outward looking orientation of the policy documents constitutes one disjunction between discourses of technology that impacts on practice.

Technologies as delivery of learning

In Bretag's account of her teaching practice (Interview August 2005), she expressed a dilemma concerning her position in the middle between institutional strategies of teaching and learning and her teaching practice as an academic. Bretag described two contrasting learning settings in her practice: one involved small groups where face-to-face communication and email dialogue were combined and resulted in successful deep learning outcomes; the other involved coordinating large international student cohorts and resulted in a heavy workload:

> We've got, over the year, about 700 in our course now, you know, because we teach in Malaysia and Hong Kong and here, and just maintaining the standard service to those students is as tough as it gets. (Bretag)

This dilemma was expressed in terms of the difficulty of negotiating with an online strategy that "frames what we do", that is, determines the context of teaching practice:

> It seems to me we've gone too much that way, economic, economic, economic, and I'm worried about the online strategy being that way. I'm not saying there aren't elements of it that would actually make a lot of sense, but if we're constantly thinking about hey, we'll be more efficient… (Bretag)

Bretag located her practice at two points, with the "we" that thinks about efficiency, as well as the "I" that attempts to adapt "elements" of the economically driven strategy. In her discussion of this issue, the two points are too far apart for Bretag to find a balance or negotiate her practice satisfactorily with either.

The rhetoric invoking the economic determinants of the contexts of teaching was associated with the institutional application of online learning technologies. Hannon (Interview, April 2007) described an encounter with an academic manager for whom he was conducting staff development in using online discussion, who equated learning technologies with a future institutional shift to wholly online courses, stating that, "this is the model we're going towards, where the university has no option." In Hannon's account he was concerned that the online strategy of the institution was interpreted unproblematically with a technology-led shift to online courses.

Hannon then identified an institutional effect in which learning technology became first separated from teaching practice, then implemented by organisational units removed from teaching and learning priorities:

> In the organisational sense you've got IT departments and they have their own interests and they put up technologies to use, and in the sense the technologies become the end themselves. (Hannon)

A consequence of the separate interests of those who implement learning technologies ("IT departments") from those who teach with them, then, created the conditions for a technology-led approach to emerge to a teaching setting, which functions "like a default pedagogy". Hannon acknowledged Goodyear and Jones (2003, p. 40) for this term:

> I think it's wrong to promote people to use a particular technology, it's got to be the use to which it's put, and almost always in an academic issue there's no simple one solution. … so the technology always tries to take centre stage, I've realised, and if you leave it alone it will do that, and if it does that it starts corroding or corrupting pedagogy. (Hannon)

In expressing this strong view, Hannon indicates a conflict in his practice of academic development in online learning, between "opening people's eyes" to the possibilities of learning technologies and institutional constraints on teaching practice.

Bretag identified a value conflict in her practice, where online strategy was "largely based on an economic rationale", and doesn't fit "with an educational rationale". Bretag identified efficiency as a value in the underpinning of the online strategy, with implied time saving. In fact, the use of learning technologies came at a cost to her and her colleagues:

> It's very time intensive to really use it to its full capacity, so then that takes me back to that question about our online strategy. I don't believe it saves you time; this didn't save me time, this just changed my life. I'm sorry, I've got two different priorities going there; this changed my student's life, it didn't save anybody time, it took a lot of time. (Bretag)

Bretag was concerned that her style of teaching practice, which had produced profound effects on student learning, had no place in the institutional online strategy exemplified in the "delivery" repertoire.

Technologies as communication and building relationships for learning

Hannon contrasted the open nature of the Internet with learning management systems, which he described as emphasising "managing content", despite some interactive software:

> there's an assumption that doing a course is accessing content... As well they have a discussion facility but in a sense it's an add on, it's really about accessing content and behind that there's this vision of distance education which is the postal one. (Hannon)

Hannon equates the learning management systems with the “default pedagogy” of the model of content access, and draws a contrasts to the “new social software” which places communication as the centre of its model. This clash of models – learning as communication versus learning as access – has limited the extent to which innovative and interactive software can be brought into institutional teaching and learning practice.

Bretag described as central to her teaching approach the process of building a relationship with a student in an intercultural context, and she had developed a research informed practice which was founded on achieving successful intercultural or “third space” communication. This approach is closely tied to a perception of student needs in the learning environment:

> They’re actually here for a much broader and deeper cultural experience, and you only get that cultural experience with interacting with people, and sharing culture and sharing of yourself. (Bretag)

For Bretag, a dilemma arose when strategy documents promoted the use of online technologies to bridge dispersed locations and conduct totally online education, yet she perceived the needs of students as primarily met by classroom contact:

> Why would students be wanting to come to university and actually be in a classroom environment? People could just go and do all their learning sitting in front of a computer, but they don’t. Why don’t they? Because they actually want human contact. (Bretag)

However, Hannon frames this as a need for contact, rather than a need for being in a particular physical location:

> When it comes down to contact, ways of contacting people, whether you do this in a group form or by email I think, well, it comes down to that contact, that’s what students tend to like. (Hannon)

Bretag attempted to accommodate her practice to the institutional online strategy, and despite professing scepticism and ambivalence, suggested that online materials and online learning, “can complement the face-to-face, but I just don’t think it should replace it.” Moreover, “something about the medium”, in her pedagogical use of email, afforded greater sharing and disclosure than the face-to-face class setting. For Bretag, modelling “complementarity” was part of her role of a teaching academic, and in the intercultural context, this involved sharing experiences mutually. Such modelling gave students new to Australia an “insider view”, which she described as part of the reason they are studying in Australia. The student need for contact, particularly contact of an intercultural nature, underpinned her discourse of relationship building as central to her teaching approach, and student-teacher relationships was the code that attracted the highest level of coding in both Bretag’s and Hannon’s interview transcripts at 39 units (see Table 2).

The keywords most frequently used in the Bretag and Hannon interview transcripts were “flexible/flexibility”, and “technology/technological/technical” (Table 1). These tended to be used in a descriptive manner, reflecting the pervasiveness of these words in descriptions of teaching and learning contexts and practice. By contrast, the ATN policy documents reflected a strategic, goal-oriented purpose, and technology related words were softened by positive associations, and closely positioned with terms such as, “opportunities”, “resource-rich”, “excellence”, “enable access”, “effective use”.

Bretag invoked two distinct but opposing ways in which online technologies impacted and shaped her practice. In the first, online modes were described as offering greater opportunity for blended learning: they were able to liberate students’ expression when informal language was used online, and enable the formation of deeper relationships via one-to one online communication, particularly in the intercultural context. The second way her practice was shaped was by the pressures of teaching large cohorts afforded by online learning systems, with the result that teaching became “maintaining the standard service”, and the economic rationale underpinning the online teaching and learning strategy which marginalised the “educational rationale”.

For Hannon, central to his academic development practice was innovative online learning approaches that were situated and contextualised in their use, enabled reflective and deep learning, and were responsive to workload issues for teaching academics. A dilemma arose for Hannon when online teaching approaches were derived from an ideal of all online, off-campus, mass learning contexts, the technology use overshadowed the pedagogy, and technology became a platform rather than a learning space.

Summary of the three repertoires

The first repertoire, deployed exclusively by the ATN policy documents, was *technology as a bridge to globalised opportunity*. We have shown that these documents metaphorically connect the reader, via the institution, to the global community, with learning technologies taking the role as that bridge. In contrast to this strategist's perspective, the second repertoire, *technology as delivery of learning* was drawn from Hannon and Bretag's perspective on using learning technology systems scaled to large, distributed teaching contexts, in a similar globalised higher education scenario. Both repertoires related to technology enabled education with the capacity for global reach, but with a crucial distinction: the former reflected a strategic and visionary perspective which was unifying and community oriented, whereas the latter used a rhetoric which was economically oriented, institutionally focussed and deployed the vocabulary of the IT help-desk.

The third repertoire described the deployment of *technologies as communication and building relationships for learning*, where technology augmented the learning context, extended pedagogical scope by offering more modes of interaction and means of "sharing culture". This repertoire was not used in the ATN policy documents. The perspectives and goals of each repertoire are schematised in Table 4: Contrasting repertoires of learning technologies.

Table 4: Contrasting repertoires of learning technologies

Interpretative repertoire	**bridge to global opportunity**	**delivery of learning**	**building learning relationships**
Source	ATN policy documents	Author interview transcripts	Author interview transcripts
Perspective	Strategic: visionary, a global, technologised world. Outward looking.	Implementation: access to learning, reach, scalable. Inward looking.	Practice: situated contexts, interactive, communicative style
Goal or accomplishment of repertoire	positions institution as global player	provides access to large-scale, distributed cost-effective education	augments face-to-face settings and enables deep and effective learning
Rhetoric	improve, engage, response, flexible, community, global, appropriate, innovate	economic inevitability, normative positioning of large-scale online learning	relationships, interaction, sharing cultural experiences, learning as change

Both the bridging and delivery metaphors of the first two repertoires have the effect of diverting attention from the actual conditions of teaching and learning practice, from the concrete considerations of organising and coordinating time, place, people, interactions and technologies, and especially maintaining quality of teaching and learning in mass or spatially dislocated settings. In the second repertoire, of technology as delivery, learning technologies are presented as neutral, a platform or system which is inherently separate from pedagogy. This repertoire deflected attention from the work involved with learning technology systems, in making a delivery platform work in the peopled settings of teacher-learner engagement. The work of organising the teaching and learning environment and engagement with colleagues and students is left to the third repertoire of communication and building relationships. The two repertoires of technology as "bridging" and as "delivery" are institutionally oriented, and project an idealised world where the work of building relationships is invisible.

Conclusion

This research aimed to explore the intersections and disjunctures between our own situated practice and the rhetoric expressed in policy documents relating to the use of learning technologies. An autoethnographic approach provided the starting point from which to reflexively engage in this process, as it allowed us as both practitioners and researchers to draw on our personal experience. Using autoethnography in combination with discourse analysis, interviews by and of the authors were contrasted with the institutional policies from the five Australian Technology Network (ATN) Universities to identify "interpretative repertoires" or recognisable themes, commonplaces and tropes.

It was apparent from the findings that there was a separation of discourses around teaching and learning with online technologies that was not easily reconcilable. The author interview transcripts revealed teaching practice that embodied a strong commitment to using online technologies to develop relationships for deep learning. This repertoire, which emerged from the context of practice, stood in stark contrast to the strategic discourse exemplified in the university strategy documents, and the functional discourse of implementation. The question raised was whether any shared ground exists between the "social worlds" (Wetherell & Potter, 1988, p. 171) depicted by these discourses.

Our findings confirmed research that found contested discourses around learning technologies in institutions, where technologies intensified the tension between technological management functions and teaching practice (Lewis et al., 2005; Siemens, 2006). Our analysis also explored the resources which we personally drew upon to respond to the resulting dilemmas in our own practice. While we clearly recognised that learning technologies provide access to large-scale, distributed cost-effective education, the autoethnographic interview data also demonstrated that our commitment to learning technologies went well beyond this functional stance, despite the challenges in our day to day practice. As noted previously (Pollock & Cornford, 2002; Roberts, 2004; Hannon, 2008), the actual work of organising people, technology, and resources, is complex and does not match the abstracted model of the delivery platform and the futuristic e-learning promise. The dilemmas of practice experienced in online learning settings were glossed over by the repertoire of *technology as bridging*, and occluded by the repertoire of *technology as delivery*. However, the repertoire of *technology as building relationships for learning*, identified in the autoethnographic element of this research, offers an alternative discourse to academics who wish to define, maintain and restore the centrality of their practice, while simultaneously working within a mandated policy framework.

References

Adams, V. (2007). Laughing It Off: Uncovering the Everyday Work Experience of Nurses. *International Journal of Qualitative Methods, 6*(1), 1–15.

Australian Technology Network of Universities. Retrieved February 29, 2008, from http://www.atn.edu.au/.

Austin, J. (1962). *How to do things with words*, Oxford: Clarendon Press.

Barnett, R. (2000). *Realizing the University in an age of supercomplexity*, Buckingham: Open University Press.

Bretag, T. & Hannon, J. (2008). Online Close and Personal: Developing a Community of Inquiry Using Computer Mediated Communciation. In Hellsten, M. & Reid, A. (Eds.) *Researching international pedagogies: Sustainable practice for teaching and learning in higher education* (pp. 221–239). Netherlands: Springer.

Bruni, N. (2000). The Crisis of Vulnerability: Ethical Dilemmas of Autoethnographic Research. *Qualitative Research Journal, 2*(1), 24–33.

Castells, M. Giroux, H., Freire, P.,Willis, P. & Macedo, D. (1999). *Critical Education in the New Information Age*, London: Rowman & Littlefield.

Castells, M. (1996). *The Rise of the Network Society*, (1). Oxford: Blackwell.

Churchman, D. (2006). Institutional commitments, individual compromises: Identity-related responses to compromise in an Australian university. *Journal of Higher Education Policy and Management*, 28(1), 3–15.

Conole, G., de Laat, M., Dillon, T., & Darby, J. (2008). 'Disruptive technologies', 'pedagogical innovation': What's new? Findings from an in-depth study of students' use and perception of technology. *Computers & Education*, 50, 511–524.

Cresswell, J.W. (1998). *Qualitative inquiry and research design: Choosing among the five traditions*, Thousand Oaks: Sage Publications.

Curtin University of Technology (2006). *Teaching and learning enabling plan.* Retrieved July 10, 2007, from http://lsn.curtin.edu.au/ssu/tlplan.html.

Edley, N. (2001). Analysing masculinity: Interpretative repertoires, ideological dilemmas and subject positions. In Wetherell, M., Taylor, S., & Yates, S. (Eds.) *Discourse and data: A guide for analysis* (pp. 189-228). Milton Keynes: The Open University.

Edwards, B. (2002). Deep insider research. *Qualitative Research Journal*, 2(1), 71–84.

Ellis, C. & Bochner, A.P. (2000). Autoethnography, personal narrative, reflexivity. Chapter 28 in N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 733–768). 2nd Ed. Thousand Oaks: Sage Publications.

Gaita, R. (2000). Truth and the university. In Coady, T. (Ed.) *Why universities matter* (pp. 26–48). Sydney: Allen & Unwin,.

Gibbs, D., Gosper, M. (2006). The upside–down-world of e-learning. J*ournal of Learning Design*, 1(2), 46–54. Retrieved September 12, 2008 from http://www.jld.qut.edu.au/.

Goodyear, P. & Jones, C. (2003). Implicit theories of learning and change: Their role in the development of e-learning environments in higher education. In Naidu, S. (Ed.) *Learning and teaching with technology: Principles and practice* (pp. 29–41). London & New York: Routledge Farmer.

Hannon, J. (2008). Doing Staff development: Dilemmas, practices and technologies, *Australasian Journal of Educational Technology,* 24(1). Retrieved February 29, 2008, from http://www.ascilite.org.au/ajet/ajet24/ajet24.html

Hotrum, M. (2005). Breaking down the LMS walls. *The International Review of Research in Open and Distance Learning*, 6(1). Retrieved February 21, 2008, from http://www.irrodl.org/index.php/irrodl/article/view/212/295.

Kalaja, P. (2006). Research on students' beliefs about SLA within a discursive approach. In Kalaja, P. & Ferreira Barcelos, A.M. (Eds.) *Beliefs about SLA: New research approaches.* Netherlands: Springer.

Kellner , D. (2003). *Technological transformation, multiple literacies, and the re-visioning of education*. Graduate School of Education & Information Studies, UCLA. Retrieved February 21, 2008, from from http://www.gseis.ucla.edu/faculty/kellner/essays.html.

Kemmis, S. & McTaggart, R. (2008). Participatory action research: Communicative action in the public sphere. In N. Denzin & Y. Lincoln (eds.). *Strategies for qualitative inquiry* (pp. 271–330). Thousand Oaks, CA: Sage.

Lewis, T. Marginson, S. & Snyder, I. (2005). The network university? Technology, culture and organisational complexity in contemporary higher education. *Higher Education Quarterly*, 59(1), 56–75.

Mann, S. (2004). A personal inquiry into an experience of adult learning online. In S. B. P. Goodyear, V. Hodgson, D. McConnell (Eds.), *Advances in research in networked learning* (pp. 205–220). Boston: Kluwer Academic Publishers.

McLoughlin, C., & Lee, M. (2008). Future learning landscapes: Transforming pedagogy through social software. *Innovate,* 4(5).

McMullin, B. (2005). Putting the learning back into learning technology. In O'Neill, G., Moore, S., & McMullin, B. (Eds.) *Emerging issues in the practice of university learning and teaching*, Dublin: AISHE. Retrieved February 21, 2008, from http://www.aishe.org/readings/2005–1/

McWilliam, E. (2005). Unlearning pedagogy. *Journal of Learning Design*, 1 (1). Retrieved February 21, 2008, from http://www.jld.qut.edu.au/Vol 1 No 1.

Marginson, S. & Considine, M. (2001). *Enterprise university in Australia. Governance, strategy and reinvention,* Cambridge: Cambridge University Press.

Miller, R. (2001). The social life of information. Briefing paper from the *Greater Expectations National Panel of the Association of American Colleges and Universities.* Retrieved May 15, 2007, from http://www.greaterexpectations.org/ briefing_papers/SocialLifeOfInformation.html.

Moloney, J. (2000). Australian universities today. In Coady, T. (Ed.) *Why universities matter* (pp.72–84). Sydney: Allen & Unwin,.

Patton, M.Q. (2002). *Qualitative research and evaluation methods*, Thousand Oaks, CA: Sage.

Pollock, N. & Cornford, J. (2002). The theory and practice of the virtual university: Working through the work of making work mobile. *Minerva* 40, 359–373.

Potter, J. & Wetherell, M. (1987). *Discourse and social psychology*, London: Sage.

Queensland University of Technology. (2007). *Approach to and context of teaching and learning.* Retrieved August 1, 2007, from http://www.mopp.qut.edu.au/C/C_01_01.jsp.

RMIT University. (2007). *RMIT 2010: Designing the Future.* Retrieved July 3, 2007, from http://www.rmit.edu.au/ browse;ID=if1c83z42eua1.

Roberts, G. (2004). The new covert curriculum: A critical actor-network approach to learning technology policy, *Proceedings of the fourth international conference on networked learning*, Lancaster University and the University of Sheffield, pp. 637–644.

Seddon, T. (1998). *Steering futures: Practices and possibilities of institutional redesign in Australian education and training.* Retrieved July 13, 2005, from http://www.aare.edu.au/98pap/sed98354.htm.

Siemens, G. (2006). Learning or management system? A review of learning. *Management System Reviews,* University of Manitoba, e*learnspace.* Retrieved February 21, 2008, from http://www.elearnspace.org/Articles/index.htm.

Smyth, J. & Hattam, R. (1998). Intellectual as hustler: Researching against the grain of marketisation. In I. Hunt & J. Smyth (Eds.). *The ethos of the university: West and beyond*. Adelaide: Flinders University Press, pp. 145–175.

Soucek, V. (1994). Flexible education and new standards of communicative competence. In Soucek. V. (Ed.) *Economising education: The post-Fordist debates* (pp. 43–103). Geelong: Deakin University,.

University of Technology, Sydney. (2007). S*etting the pace 2006–2009: Strategic directions for the current decade.* Retrieved August 1, 2007, from http://www.planning.uts.edu.au/pdfs/settingthepace.pdf.

University of South Australia. (2007). *Teaching and Learning Strategy 2006–8.* Retrieved July 3, 2007, from http://www.unisa.edu.au/staff/teachlearn/default.asp.

University of South Australia. (2004). *Towards an online strategy*. Internal policy document no longer available online as at 22 February 2008.

Wetherell, M. (1998). Positioning and interpretative repertoires: Conversation analysis and post-structuralism in dialogue. *Discourse & Society*, 9, 387–413.

Wetherell, M. & Potter, J. (1988). Discourse analysis and the identification of interpretative repertoires. In Antaki, C. (Ed.) *Analysing everyday explanation: A casebook of methods* (pp. 168–183). London: Sage.

Zemsky, R., & Massy, W. (2004). Thwarted innovation: What happened to e-learning and why. *A final report for the Weatherstation Project of the Learning Alliance at the University of Pennsylvania in cooperation with the Thomson Corporation.* Retrieved March 14, 2005, from http://www.irhe.upenn.edu/Docs/Jun2004/ThwartedInnovation.pdf.

APPENDIX 1: Interview Questions

1. Your role as educator:
   Can you describe your role – it may be multiple - your area of expertise, your teaching and research areas
2. First use:
   Describe your first use(s) of e-learning technologies. What was it, when, was it successful and so on. How did you discover it?
3. Changes in use:
   Have you discarded or shifted away from any uses of elearning approaches or technologies? Why?
4. A specific current use:
   Can your describe a current use of e-learning technologies or computer mediated communication that is significant for you? Briefly, how did this project arise, and what do you hope to achieve?
5. How does this project fit in with the organisation framework and IT system.
6. What has worked well in this project or related areas? What has been opened up by this engagement for you or others?
7. What hasn't or doesn't work well in this project or related areas? Has anything or anyone been constrained, excluded or foreclosed?
8. Can you describe any unexpected consequences of in your use of networked communication technologies in this project?
9. Can you describe any innovative uses, adaptations, or workarounds involving technologies for this project that you have discovered or used with some success?
   – Innovative uses include "official" uses, which are supported and presented by your organisation, and "unofficial" uses, which are those discovered through your own research and contacts.
10. What concerns you about where e-learning is heading? Can you comment on the direction of your work with e-learning.

Hoon, T. S., Chong, T. S., & Binti Ngah, N. A. (2010). Effect of an Interactive Courseware in the Learning of Matrices. *Educational Technology & Society*, *13* (1), 121–132.

Effect of an Interactive Courseware in the Learning of Matrices

Teoh Sian Hoon[1], Toh Seong Chong[2] and Nor Azilah Binti Ngah[2]
[1]Information Technology & Quantitative Sciences, Universiti Teknologi MARA, Penang, Malaysia // [2]Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia, Penang, Malaysia // sianhoon02@yahoo.com // tohsc@usm.my // azilah@usm.my

ABSTRACT

The main aim of this study is to integrate cooperative learning strategies, mastery learning and interactive multimedia to improve students' performance in Mathematics, specifically in the topic of matrices. It involved a quasi-experimental design with gain scores and time-on-task as dependent variables. The independent variables were three instructional strategies (CCL, CML and CCML) with academic abilities as the moderator variable. The sample for the study was 262 Form Four Malaysian students. A courseware entitled "Matrices" was developed using Macromedia Authorware as the authoring tool. In this study, the collected data was used to investigate the effects of the three learning strategies on the gain scores and time-on-task. Based on the gain scores and time-on-task, the effectiveness of the three learning strategies was discussed. This study showed that the CCML and CML strategies were superior compared to the CCL strategy; CCML strategy produced the highest gain score. For students with low academic ability, the CML strategy was found to be the most effective strategy. The findings of this study also suggested that high academic ability students would obtain high gain scores regardless of the instructional strategies. In terms of time-on-task, students in CCL and CML strategies demonstrated significant lower time-on-task than CCML strategy.

Keywords

Computer-assisted cooperative learning (CCL), Computer-assisted mastery learning (CML) and Computer-assisted cooperative mastery learning (CCML)

Background of the study

One of the major problems among mainstream secondary school students is the performance difference between the low achievers and the high achievers. To overcome this problem, various interventions had been offered including curriculum-based assessment (Fuchs, Fuchs and Tindal, 1986), direct instruction curriculum design (Engelman and Camine, 1982), mastery learning (Bloom, 1984), tutoring (Sleeman and Brown, 1982), learning strategies (Mason, Burton and Stacey, 1982), and so forth. Unfortunately, most of these interventions required additional resources such as teachers' efforts and time needed to use them. However, the advent of Information and Communication Technology (ICT) in the last few years has eased the burden on the resources needed for the teaching and learning process. The use of computer as an ubiquitous teaching tool has become prevalent in Malaysian schools. As a result, the use of computer in conjunction with effective teaching strategies has tremendous potential in the teaching and learning process.

With the use of computer, mastery learning has a high potential to become an effective and extensive teaching and learning tool (Guskey, 1997; Guskey and Gates, 1986; Kulik, Kulik, and Bangert-Downs, 1990). The mastery learning method divides subject matter into units and each unit has predetermined objectives. Students should achieve mastery on unit tests, typically 80%, before moving on to the following units. Students who do not achieve mastery receive remediation and students who achieve mastery have the opportunity to participate in enrichment activities. Mastery learning fits well with other strategies and complements cooperative learning (Guskey, 1997). Researchers, such as Dansereau (1988), Gunderson and Johnson (1980), Hooper, Temiyakarn, and Williams (1993), had strongly recommended that cooperative learning should be used in the teaching and learning process. As suggested by Guskey (1997), it needed a comprehensive framework to incorporate other instructional strategies into mastery learning. Guskey (1997) had suggested cooperative learning as one of the instructional strategy. Cooperative learning as part of collaborative learning has gained educators' attention to include it into process of learning (Wells and Brook, 2004). A meta-analysis based on 39 rigorous studies on a common basis in science, mathematics, engineering and technology showed that generally, cooperative learning significantly increased academic performance, decreased dropout rates and increased student self-confidence (Springer, Stanne & Donovan, 1999).

Cooperative learning was preferable to be incorporated into mastery learning since the goal of using cooperative learning was to accomplish a specific learning task through people working together in groups (Panitz, 1997). The learners were more concerned with mastery of a pre-determined body of knowledge. In other words, mastery learning complemented and fitted well with cooperative learning (Guskey, 1997).

Over the years, studies by Guskey and other researchers (Atkinsola, 1996; Mevarech, 1985) had found that cooperative learning could be incorporated into mastery learning to present a conducive learning environment for students. Students who focused on specific instructional goals were actively engaged in cooperative learning activities, thus, effective study time was increased. Students were properly guided with mastery learning materials in a cooperative learning environment in order to strengthen their skills of self-awareness and personal controls in learning. Hence, cooperative learning was considered an efficient way of increasing effective study in mastery learning. Results from the above studies had shown that the combination of mastery learning and cooperative learning were found to be superior to the traditional lecture teaching format. Specifically, these studies indicated that mastery learning and cooperative learning had an impact on affective and academic outcomes of students.

The combination of systematic design and integration of cooperative learning strategies, mastery learning and interactive multimedia might have a great impact on the teaching and learning of subjects, such as Mathematics, where hierarchical knowledge is the requirement of the field. Mathematics learning skills could easily be learned in a cooperative setting. Cooperative learning provided opportunities to students with low academic abilities to model their study skills and work habits as compared with high academic abilities. With the help given by students with high academic abilities in explaining in detail the steps in the worked-out examples, weaker students were then convinced to use these skills in order to obtain a mathematical solution. Besides, students with high abilities often developed greater mastery during discussions by obtaining a deeper understanding of the task (Becker, Silver, Kantowski, Travers, and Wilson, 1990; Stigler, Lee, Lucker, and Stevenson, 1982). With mastery learning, cooperation needs were structured and guided through systematic instruction and feedback.

In learning mathematics, a diagnostic test (Teoh, 2003) had shown that students with difficulties in matrices were also weak in the basic skills of mathematics, such as solving equations. Specifically, students who experienced difficulties in matrices would find doing multiplication of two matrices confusing. It showed that the basic skills in mathematics had become a necessity in solving problems and understanding other concepts in mathematics (Wu, 1999). Furthermore, to avoid omission of important processing skills, students had to be trained to master the basic skills in the early stages of the learning experience with enough time and quality instruction (Bloom, 1968). If the students were not provided with enough time, they might find difficulty to proceed to a higher stage of learning (Harrell, Walker, Hildreth and Tayler-Wood, 2004). Without the awareness, weak students were found with no improvement in their skills in mathematics. In contrast, if teachers tended to focus only on weak students, then students with good performance would not be able to get the teachers' attention in the learning process. Subsequently, mastery learning played an important role in providing an environment for students to be involved in their study, whereby students with high and low abilities were able to learn at their own pace, with the help of feedback corrective and enrichment activities. However, each component of mastery learning involved a great amount of work which made it inapplicable to the manageability and constraints relating to time (Anderson and Jones, 1981; Levine, 1985). As an example, excessive amount of testing, corrective and enrichment activities were needed during 'feedback', an important component in mastery learning. The time allocation for subjects in the normal school curriculum was evidently not sufficient for mastery learning to be applied. Currently, the use of the e-learning platform to teach developmental mathematics in a mastery learning format had been promoted to overcome this obstacle (Boggs, Shore and Shore, 2004).

With the advent of ICT as a teaching tool and the availability of computer hardware in the schools, the problem in applying mastery learning could be improved by using interactive courseware. Feedback activities could also be easily conducted by using computers. In addition, for recording purposes of students' performance, the technology could also reduce time and effort required to implement comprehensive interventions needed in mastery learning materials.

Objectives of the study

This study is mainly aimed to integrate cooperative learning strategies, mastery learning and interactive multimedia to improve the students' performance in Mathematics, specifically in the topic of matrices. The integration of cooperative learning, mastery learning, and interactive multimedia environment would provide a comprehensive framework needed for an effective and efficient teaching and learning of mathematical concepts. A computer-based systematically designed interactive courseware was created to test the hypotheses of this study. The effects on the gain scores and time-on-task would be investigated to determine the effectiveness of using the courseware in three different strategies, namely, Computer-assisted Mastery Learning (CML), Computer-assisted Cooperative Learning (CCL), and Computer-assisted Cooperative Mastery Learning (CCML). Students in the three learning strategies used the same instructional materials. The CML strategy was based on individual learning, while the CCML and CCL strategies were based on cooperative learning. Certain elements of mastery learning were added to the courseware, which were used in the CML and CCML strategies. The CCL strategy was based on cooperative learning and used the version of the courseware without the elements of mastery learning. The effects of the three learning strategies on the gain scores and time-on-task were investigated.

Research framework

This study examined the effects of the three learning strategies, which were measured using gain scores and time-on-task. The moderator variable was the academic ability. The dependent variables were the gain score and time-on-task. The relationship among the variables is depicted in Figure 1, describes the research framework of this study. As evidenced from the past social and cognitive psychology researches, academic achievement outcomes (gain scores) had been a significant variable in learning success within education classroom (Schwarz, 1998). In addition, many researchers (Schremmer, Hortz and Fries, 2001; Toh, 1998) used gain scores to investigate the effectiveness of treatment in instruction.

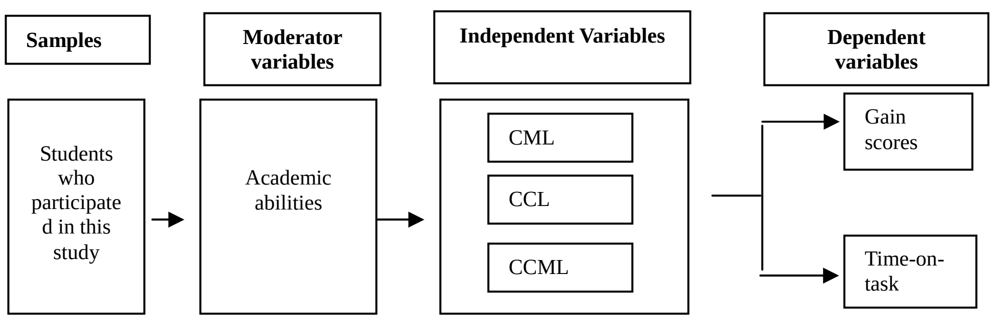

Figure 1. Research framework

Carroll (1989) emphasized that if a student really spent time in learning as needed, then he would achieve competence in learning. Bloom (1984) described that mastery learning would take more time than the normal teaching. With high quality instruction, a variety of methods were included that made learning easier for students to understand and remember. Thus, all students could learn differently with their individual abilities and work at their own pace through the planned sequence of lessons. It helped to motivate them to learn important concepts in order to proceed to the subsequent learning units. In this situation, the different academic abilities among learners would not affect their learning to a great extent. Mastery learning could be easily adapted to reduce achievement differences among students. Besides, time-on-task for a student could be shorter if an opportunity is given to a student to learn a learning unit through a series of quality instruction. With quality instruction, students would be more persistent in learning and increase their ability to understand the learning unit.

In mastery learning, students were grouped into high and low academic abilities. In this study, a standardized Mathematics examination results from the PMR examination (Penilaian Menengah Rendah, a standardized examination in Malaysia) were used to classify the students into high and low academic abilities. The examination gauges students' abilities after nine years of education in Malaysia. Therefore, it would be an accurate representation of students' mathematical abilities vis-à-vis the national norm as students could be classified into high and low academic abilities.

The followings are the hypotheses of this study.

H1 There are no significant differences in the dependent variables among students in the CCL, CML and CCML strategies.

H1.1 There is no significant difference in the gain scores among students in the CCL, CML and CCML strategies.

H1.2 There is no significant difference in the time-on-task among students in the CCL, CML and CCML strategies.

H2 There are no significant differences in the dependent variables among students with high academic abilities in the CCL, CML and CCML strategies.

H2.1 There is no significant difference in the gain scores among students with high academic abilities in the CCL, CML and CCML strategies.

H2.2 There is no significant difference in the time-on-task among students with high academic abilities in the CCL, CML and CCML strategies.

H3 There are no significant differences in the dependent variables among students with low academic abilities in the CCL, CML and CCML strategies.

H3.1 There is no significant difference in the gain scores among students with low academic abilities in the CCL, CML and CCML strategies.

H3.2 There is no significant difference in the time-on-task among students with low academic abilities in the CCL, CML and CCML strategies.

Methodology

The design of this research is a quasi-experimental design. This study involved 262 students from four different secondary schools, namely schools *A*, *B*, *C* and *D*. Randomly, school *A* was assigned to the CCL treatment, school *B* and *D* was assigned to the CML treatment and school *C* was assigned to the CCML treatment. The number of students in CCL, CML and CCML were 77, 81 and 104 respectively.

The researcher had developed the courseware entitled "Matrices". Two sets of courseware were used in this study. The first courseware was designed with mastery learning elements, used in the CML and CCML strategies. The second courseware was designed without mastery learning elements, used in the CCL strategy. Before conducting the experiments, the courseware was field-tested for revision purposes.

The courseware was used as the instruction in those three groups, which were CCL, CML and CCML. Gain scores and time-on-task were used to investigate the effectiveness of the mentioned strategies. Before using the courseware, an entry test was conducted to filter students' basic knowledge in Matrices and determine whether they possessed the requisite prior knowledge on arithmetic, which involved addition, subtraction, multiplication and division of numbers (integers, fraction and decimal), and solution of one linear equation as well as two linear equations. A student's prior knowledge was considered high if he or she scored 80% and above. If a student scored less than 80%, then the student had to complete an interactive courseware program on arithmetic. The interactive courseware was specially created to strengthen those students' prior knowledge in matrices. Students had to obtain the required level of mastery before they could be given the treatment. Thus, before the actual commencement of the experiment, all samples in each group would have achieved the required prior knowledge.

The PMR Mathematics result was used to classify the students into different academic abilities. Students with Grade A and Grade B were grouped in the high academic ability category. Students with Grade C and Grade D were grouped in the low academic ability category.

The pretest and posttest questions were developed to determine students' understanding of important concepts related to Matrices. These tests, that consisting of 51 questions, had the reliability of 0.7051 based on the Kuder-Richardson Formula (KR20).

On the first day of the data collection, students were given a briefing on the learning strategies. Next, students were given a pretest on matrices and followed by a lesson on Matrices and Equal Matrices on the second day. After the lesson, students were given the first formative test using the computer. The subtopics covered in the whole process were:

(1) Matrices and Equal Matrices
(2) Addition and Subtraction on Matrices
(3) Multiplication of a matrix by a number; Multiplication of two matrices
(4) Identity Matrix, Inverse Matrix and solution of simultaneous linear equations by using Matrices.

The whole lesson took four to six hours to finish. Students took a test after each subtopic. The differences of three treatment groups in terms of presentation of the lessons, team function and individual improvement were as depicted in the following discussion. Students in the CML, completed all formative tests or quizzes independently. Students who failed to meet the required performance level received supplementary instruction and corrective activities immediately after each question until the requirement was met. At the end of a test, extra corrective activities were given to those who could not achieve the success level of 80% as evaluated by the computer.

Students in the CCL completed all activities in cooperative groups and they completed all tests independently. Students in the CCML completed all activities in cooperative groups and all tests were carried out independently for this group. Students who failed to meet the required performance level received supplementary instruction and correction activities immediately after each question until the requirement was met. At the end of a test, extra corrective activities were given to those who could not achieve the success level of 80% as evaluated by the computer. Each student needed to wait until all members in the group had achieved the level of 80%. Those who had finished and achieved 80% of the score were able to help other students who had not achieved 80% of the score. All the cooperative work was examined using checklist for the approach named Student Team Achievement Division (STAD).

The design of the courseware was based on a macro and micro design. Alessi and Trollip's instructional design model (Alessi and Trollip, 2001) was used for the macro design. Gagné's nine events of instruction (1985) was used for the micro design of the courseware. Motivational elements were incorporated into the courseware which was created based on Gagné's Motivational elements were incorporated into the courseware which was created based on Gagné's nine events of instruction.

Results

MANOVA was used with gain scores and time-on-task as the two dependent variables and the learning strategies (CCML, CML and CCL) as the group factor. Follow-up analyses would be conducted if the test on MANOVA was significant.

Gain scores and time-on-task for the three learning strategies

The descriptive statistics on the gain scores and time on task for CCL ($Gain_{CCL}$), CML ($Gain_{CML}$) and CCML ($Gain_{CCML}$) are shown in Table 1, where $Gain_{CCML} > Gain_{CML} > Gain_{CCL}$ and $Time_{CML} < Time_{CCL} < Time_{CCML}$.

Table 1. Descriptive statistics on gain score for CCL, CML and CCML

	CCL		CML		CCML		TOTAL	
	Gain scores	Time-on-task	Gain scores	Time-on-task	Gain scores	Time-on-task	Gain scores	Time-on-task
Mean	31.47	3.90	42.79	3.70	49.40	4.71	42.09	4.16
N	77	77	81	81	104	104	262	262
Std Dev	19.206	0.771	19.678	1.030	17.849	0.784	20.164	0.973

The results of the MANOVA test (Table 2) showed that the Wilk's lambda of 0.549 was significant, $F = 45.032$, $p < 0.05$. Thus, Hypothesis One, which stated that the population means on dependent variables (i.e., gain scores and time-on-task) were the same for the three groups, was rejected. The multivariate Eta Squared indicated that 25.9% of multivariate variance of the dependent variables was associated with the group factor.

Table 2. Multivariate tests of the effect of learning strategies on the dependent variables

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
strategies	Wilks' Lambda	.549	45.032	4.000	516.000	.000	.259

Using multiple univariate ANOVAs, a follow-up approach was conducted. Two ANOVAs were conducted, one for each dependent variable (i.e., gain scores and time-on-task). The results of the univariate ANOVAs are shown in Table 3. The univariate ANOVA for gain scores was significant, $F = 20.155$, $p < 0.025$. Likewise, the univariate ANOVA for time-on-task was significant, $F= 36.066$, $p < 0.025$. Both results showed that there were significant differences of gain scores and time-on-task among the groups. Therefore, Hypothesis 1.1 and Hypothesis 1.2 were rejected.

Table 3. Univariate tests of the effect of learning strategies on the dependent variables

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Gain Scores	14291.342	2	7145.671	20.155	.000	.135
	Time-on-task	53.863	2	26.932	36.066	.000	.218
Intercept	Gain Scores	437568.203	1	437568.203	1234.189	.000	.827
	Time-on-task	4336.984	1	4336.984	5807.944	.000	.957
Strategies	Gain Scores	14291.342	2	7145.671	20.155	.000	.135
	Time-on-task	53.863	2	26.932	36.066	.000	.218
Error	Gain Scores	91825.639	259	354.539			
	Time-on-task	193.404	259	.747			
Total	Gain Scores	570219.000	262				
	Time-on-task	4782.000	262				
Corrected Total	Gain Scores	106116.981	261				
	Time-on-task	247.267	261				

The analyses revealed that there were significant differences in gain scores for the two pairs- CCML with CCL and CML with CCL. Also, there were significant differences in time-on-task for the two pairs- CCML with CCL and CCML with CML.

The effect sizes of learning strategies on the gain score

Effect sizes of CML and CCML towards CCL were studied because there were significant differences on gains scores between CML and CCL as well as between CCML and CCL. Calculations of the effect size (ES) of CML and CCML towards CCL are illustrated in Table 4. The results showed that the effect size of CML towards CCL was 0.5603, which was moderate. This indicated that an individual learner in CML had a 0.5603 standard deviation increase. The effect size of CCML towards CCL was 0.8778. Therefore, effect size of CCML towards CCL was stronger if compared to the effect size of CML towards CCL.

Table 4. The effect size of CCML and CML towards CCL

Learning Strategies	Difference of Means	Pooled Standard Deviation	Effect Size, *ES* $= \frac{\text{The Difference of Means}}{\text{Pooled Standard Deviation}}$
CML CCL	11.32	20.202	$ES = \frac{11.32}{20.202} = 0.5603$
CCML CCL	17.93	20.424	$ES = \frac{17.93}{20.424} = 0.8778$

The learning strategies effects on the dependent variables among students with high academic ability

The descriptive statistics on gain scores and time-on-task for CCL, CML and CCML of students with high academic abilities are illustrated in Table 5. For gain scores, the mean of CML ($Gain_{CML}$) and CCML ($Gain_{CCML}$) were close to each other, with the respective values of 53.49 and 52.52. Subsequently, CCL showed a lesser value in the gain scores ($Gain_{CCL}$) of 46.03. Thus, students with high academic ability obtained gain scores in the following sequence, $Gain_{CML} > Gain_{CCML} > Gain_{CCL}$.

Table 5. Descriptive statistics on gain scores and time-on-task among students with high academic ability

	Strategy	Mean	Std. Deviation	N
Gain Scores	Cooperative (CCL)	46.03	18.128	35
	Mastery(CML)	53.49	18.505	45
	Cooperative mastery(CCML)	52.52	16.254	92
	Total	51.45	17.372	172
Time-on-task	Cooperative(CCL)	3.46	.657	35
	Mastery(CML)	3.27	.780	45
	Cooperative mastery(CCML)	4.57	.684	92
	Total	4.00	.930	172

It can be seen in Table 5 that the time-on-task of CML and CCL are close to each other which were 3.27 hours ($Time_{CML}$) and 3.46 hours ($Time_{CCL}$) respectively. Subsequently, students in CCML spent longer time-on-task ($Time_{CCML}$), that was 4.57 hours.

Results of the MANOVA test (Table 6) showed that Wilk's lambda of 0.52 was significant, $F = 32.542$, $p < 0.05$. Thus, Hypothesis Two, which stated population means on the dependent variables among students with high academic ability were the same for the three groups, was rejected. The multivariate Eta Squared showed 27.9% of multivariate variance of the dependent variables was associated with the group factor.

Table 6. Multivariate tests of effect of learning strategies on the dependent variables among students with high academic ability

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Strategy	Wilks' Lambda	.520	32.542	4.000	336.000	.000	.279

A follow-up approach was conducted by using multiple univariate ANOVAs. Two ANOVAs were conducted, one for each dependent variable (i.e., gain scores and time-on-task). Results of the univariate ANOVAs are shown in Table 7. The univariate ANOVA for gain scores was not significant, $F = 2.221$, $p > 0.025$, but the univariate ANOVA for time-on-task was significant, $F = 64.214$, $p < 0.025$ and the associated Eta Squared was 43.2%. The results showed that there were significant differences in time-on-task for the two pairs- CCML with CCL and CCML with CML. Thus, Hypothesis 2.1 was not rejected and Hypothesis 2.2 was rejected.

Table 7. Univariate test of the effect of learning strategies on the dependent variables among students with high academic ability

Source	Dependent Variable	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Gain Scores	1321.456	2	660.728	2.221	.112	.026
	Time-on-task	63.906	2	31.953	64.214	.000	.432
Intercept	Gain Scores	374873.704	1	374873.704	1259.938	.000	.882
	Time-on-task	2066.745	1	2066.745	4153.425	.000	.961
Strategy	Gain Scores	1321.456	2	660.728	2.221	.112	.026
	Time-on-task	63.906	2	31.953	64.214	.000	.432
Error	Gain Scores	50283.172	169	297.534			
	Time-on-task	84.094	169	.498			
Total	Gain Scores	506968.000	172				
	Time-on-task	2900.000	172				
Corrected Total	Gain Scores	51604.628	171				
	Time-on-task	148.000	171				

The learning strategies effects on the dependent variables among students with low academic ability

The descriptive statistics on gain scores and time-on-task for CCL, CML and CCML of students with low academic ability are shown in Table 8. For the gain scores, it can be seen that the mean of CML ($Gain_{CML}$) was the highest, which was 29.42, and followed by CCML ($Gain_{CCML}$) with the mean of 25.50. Mean of gain scores among students in CCL ($Gain_{CCL}$) were the lowest, which was 19.33. Thus, $Gain_{CML} > Gain_{CCML} > Gain_{CCL}$.

Table 8. Descriptive statistics on gain scores and time-on-task among students with low academic ability

	Strategy	Mean	Std. Deviation	N
Gain Scores	Cooperative (CCL)	19.33	8.913	42
	Mastery (CML)	29.42	11.111	36
	cooperative mastery (CCML)	25.50	9.625	12
	Total	24.19	10.909	90
Time-on-task	Cooperative (CCL)	4.26	.665	42
	Mastery (CML)	4.25	1.052	36
	cooperative mastery (CCML)	5.83	.577	12
	Total	4.47	.985	90

As shown in Table 8, time-on-task for CML and CCL are approximately equal which were 4.25 hours ($Time_{CML}$) and 4.26 hours ($Time_{CCL}$) respectively. It was observed that the mean of the time-on-task in CCML ($Time_{CCML}$) was the highest, which was 5.83 hours.

The results of the MANOVA test (Table 9) showed that Wilk's lambda of 0.527 was significant, F= 16.239, $p < 0.05$. Thus, Hypothesis Three, which stated that the population means on the dependent variables (i.e., gain score and time-on-task) among students with low academic ability were the same for the three groups, was rejected. The multivariate Eta Squared indicated that 27.4% of multivariate variance of dependent variables was associated with the group factor.

Table 9. Multivariate tests of the effect of learning strategies on the dependent variable among students with low academic ability

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Strategy	Wilks' Lambda	.527	16.239	4.000	172.000	.000	.274

A follow-up approach was conducted by using multiple univariate ANOVAs. Two ANOVAs were conducted, one for each dependent variable (i.e., gain scores and time-on-task). Results of the univariate ANOVAs are shown in Table 10. The univariate ANOVA for gain scores was significant, F = 10.093, $p < 0.025$ and the associated Eta Squared was 18.8%. Also, the univariate ANOVA for time-on-task was significant, F = 18.586, $p < 0.025$ and the associated Eta Squared was 29.9%. Thus, Hypothesis 3.1 and Hypothesis 3.2 were rejected.

Table 10. Univariate test of the effect of learning strategies on the dependent variable among students with low academic ability

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Gain Scores	1994.706	2	997.353	10.093	.000	.188
	Time-on-task	25.864	2	12.932	18.586	.000	.299
Intercept	Gain Scores	40861.522	1	40861.522	413.507	.000	.826
	Time-on-task	1525.236	1	1525.236	2192.021	.000	.962
Strategy	Gain Scores	1994.706	2	997.353	10.093	.000	.188
	Time-on-task	25.864	2	12.932	18.586	.000	.299
Error	Gain Scores	8597.083	87	98.817			
	Time-on-task	60.536	87	.696			
Total	Gain Scores	63251.000	90				
	Time-on-task	1882.000	90				
Corrected Total	Gain Scores	10591.789	89				
	Time-on-task	86.400	89				

Discussion

Generally, there were significant differences in gain scores between the learning strategies. The effect size in gain scores suggested that the CCML strategy had a more positive effect than the CML strategy. These results support the findings from past research that cooperative mastery learning produced better results (Akinsola, 1996; Krank and Moon, 2001; Laney, Frarich, Farich & Luke, 1996). Furthermore, these results are consistent with Mevarech's (1985) and Okebukola's (1985) findings that discovered positive effects of cooperative learning in the application of STAD approach and even better effects if it was combined with mastery learning.

Although the CCML strategy had better results in the gain scores, it showed no significant difference in the gain scores between the CCML and CML strategies. In this case, the contribution of the CML and CCML strategies are equally important in terms of the gain scores in which both learning strategies had mastery learning. In other words, mastery learning plays an important role in organizing a systematic and more structured instruction to guide students for gaining high scores. In addition, incorporating cooperative learning could strengthen the role of mastery learning. This study also found that students in cooperative mastery learning groups were guided through well-designed instruction. Hence, better effect size was seen when the CCML strategy was used. These were consistent with Okebukola's (1985) findings that cooperative learning could strengthen students' performance. This study shows that the effect size of cooperative mastery learning was the highest in the gain scores and the effect was mainly contributed by mastery learning. The findings also support Mevarech's (1991) view that mastery learning had been successful in producing gains in achievement. There was a tendency to incorporate such programmes with cooperative learning strategies, which was called cooperative mastery learning (Mevarech, 1985).
Although mastery learning (systematic work) was the most important instructional method to make students succeed, it is better if supported by cooperative learning. This finding suggests that the advantages in cooperative learning were not obviously shown in the gain scores without mastery learning. This study shows that mastery learning plays a primary role and when incorporated with cooperative learning, students will learn more and systematically in the cooperative environment. Some students might be weak in the socialization and interaction skills and might need guidelines in mastery learning. Likewise, some students needed peer-guidance during the learning process. Thus, the CCML strategy was found to be the most effective learning strategy in this study.

In terms of time-on-task, the major finding is that students in the CML and CCL strategies spent shorter time-on-task compared to the CCML strategy. The results are consistent with past researches (Mortimore, and Sammons, 1987). According to Zimmerman (1998), there was an apparent correlation between time and achievement. The more time-on-task was made available to the student, the more activities and learning processes were involved. However, past literature suggested that even though time was certainly a critical factor but it had little direct impact on students' performance (Zimmerman, 1998). Simply adding time would not really produce large gains in student achievement. Rather, quality was the key to making time matter (Funkhouser, Humphrey, Panton and Rosenthal, 1995; Levin, 1985). Essentially, students should be provided with activities and instructions that catered to their needs and abilities, engaging them so they would continue to build on what they had learnt. What matters most were the valuable experience when students were absorbed in instructional activities that were adequately challenging, yet allowed them to experience success. This study has shown that the CML and CCML strategies could provide these experiences for the students.

For students with high academic abilities, the analysis showed that there were no significant differences in the gain scores among students with high academic ability in the CCL, CML and CCML strategies. The finding suggests that high academic ability learners could obtain high gain scores regardless of learning strategies. In many ways, students with high academic abilities were more alike in terms of the way they managed themselves in studies (Monaco, 2003). Students with high academic abilities were able to learn under any condition in the school, for example, small groups, alone and in informal settings. Thus, teachers could use any grouping pattern and instructional method as long as the method used could accommodate their teaching objectives.

This study also shows that there were significant differences in time-on-task across the learning strategies among students with high academic abilities. Students with high academic abilities in the CML and CCL strategies significantly spent shorter time-on-task compared to the CCML strategy. Therefore, the CCML students who were involved in using components of mastery learning and cooperative learning spent more time compared to those who were involved in either the mastery or the cooperative learning only. These findings indicated that there were no effects resulting from learning strategies among students with high academic ability. However, students with high academic abilities spent significant higher time-on-task on CCML without showing higher gain scores significantly. This result shows that cooperative learning provided a conducive structure for learning because students with high academic abilities were fully engaged to help other counterparts by clarifying misunderstandings and correcting learning errors to attain a criterion-referenced standard through a well-designed mastery learning instruction as revealed by Bork (1999). Therefore, students with high academic abilities in the CCML strategy showed significant higher time-on-task without showing higher gain scores.

For students with low academic abilities, the analyses reveal that there were significant differences on gain scores for CML and CCL. Also, there were significant differences on time-on-task for the two pairs which were CCML with CCL and CCML with CML.

The above results show that students with low academic abilities did not achieve significantly higher gain scores in the CCL and CCML strategies if compared to the CML strategy. This indicates that cooperative learning which was used in the CCL and CCML strategies did not give assistance to students with low academic ability. Additionally, within cooperative learning groups, students with low academic abilities did not influence each other's learning. Nevertheless, with the well-designed mastery learning instruction within a mastery learning environment, they could attain significantly higher gain scores compared to students in the CCL and CCML strategies. This could be interpreted within the context of the group processing concept in cooperative learning. According to Yager, Johnson, and Johnson (1985), the achievement of low academic ability students in cooperative learning environment also depended on how well their group was functioning.

In terms of time-on-task, the results were consistent with the results of students with high academic abilities, where students in CML and CCL significantly spent shorter time-on-task compared to CCML. Therefore, in terms of getting a higher gain score and shorter time-on-task, CML was the best learning strategies among students with low academic abilities. With the CML strategy, students' result increased within a shorter time compared to the students who used the CCML strategy.

Summary and conclusion

In conclusion, this study has led to the following key findings. Firstly, this study shows that the CCML and CML strategies are superior compared to the CCL strategy. Secondly, the CCML strategy is the best choice among the three learning strategies to obtain a higher gain score. However, for students with low academic ability, the CML strategy is found to be the best choice. This finding also suggests that high academic ability learners could obtain high gain scores regardless of learning strategies. Thirdly, in terms of time-on-task, students in CCL and CML strategies demonstrated significant lower time-on-task than CCML strategy.

The findings of this study propose a simple yet powerful approach to facilitate the learning process of students through the use of a multimedia integrated learning system with a series of high quality instructions in cooperative mastery learning and mastery learning.

References

Alessi, S. M., & Trollip, S. R. (2001). *Multimedia for learning*. NJ: Allyn and Bacon.

Anderson, L. W., & Jones, B. F. (1981). Designing instructional strategies which facilitate learning for mastery. *Educational Psychologist*, 16(3), 121-137.

Atkinsola, M. K. (1996). *Combined mastery learning and cooperative learning strategy was found to be more suitable in facilitating achievement in integrated science in the Junior Secondary School*. Retrieved September 24, 2001, from http://www.ipn.uni-kiel.de/projekte/esera/book/b132-aki.pdf.

Becker, J.P., Silver, E.A., Kantowski, M.G., Travers, K.J., & Wilson, J.W. (1990). Some observations of mathematics teaching in Japanese elementary and junior high schools. *Arithmetic Teacher*, 38(2), 12-21.

Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment*, 1(2), 1-5.

Bloom, B. S. (1984). The search for methods of group instruction as effective as one-to-one tutoring. *Educational Leadership*, 41(8), 4-18.

Boggs, S., Shore, M., & Shore, J.A. (2004). Using e-Learning platforms for mastery learning in developmental mathematics courses. *Mathematics and Computer Education*. 38(2), 213-220.

Bork, A. (1999). The future of learning. *Educom Review*. 34(4).

Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher*, 18(1), 26-31.

Dansereau, D. F. (1988). Cooperative learning strategies. In C. E. Weinstein, E. T. Goetz, & P.A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (p. 103-120). New York: Academic Press.

Engelmann, S., & Camine, D. (1982). *Theory of instruction:* Principles and application. New York: Irvington.

Fuchs, L. S., Fuchs, D., & Tindal, G. (1986). Effects of mastery learning procedures on student achievement. *Journal of Chemical Education*, 79(5), 286-291.

Funkhouser, J. E., Humphrey, D. C., Panton, K. L. M., & Rosenthal, E. D. (1995). *A research review: The educational uses of time*. (Vol. IV). Washington, DC: Policy Studies Associates, Inc.

Gagné, R. M. (1985). *The conditions of learning and theory of instruction*. Fort Worth: Holt, Rinehart and Winston.

Gunderson, B., & Johnson, D. W. (1980). Building positive attitudes by using cooperative learning groups. *Foreign Language Annals*, 13, 39-46.

Guskey, T. R. (1997). *Implementing mastery learning*. New York: Wadsworth.

Guskey, T. R., & Gates, S. L. (1986). Synthesis of research on the effects of mastery learning in elementary and secondary classroom. *Educational Leadership*, 43(8), 73-80.

Harrell, P. E., Walker, M., Hildreth, B., & Tayler-Wood, T. (2004). Mentoring BUGS: An integrated science and technology curriculum. *The Journal of Computers in Mathematics and Science Teaching*. Austin. 23(4). 367-378.

Hooper, S., Temiyakarn, C., & Williams, M. D. (1993). The effects of cooperative learning and learner control on high- and average-ability students. *Educational Technology Research and Development*, 41(2), 5-18.

Krank, H. M., & Moon, C. E. (2001). Can a combined mastery/cooperative learning environment positively impact undergraduate academic and affective outcomes?. *Journal of College Reading and Learning*, 31(2), 195-208.

Kulik, J. A., Kulik, C. C., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, 60(2), 265-299.

Laney, J. D., Frarich, D. K., Frarich, L. P., & Luke, K. P. (1996). The effect of cooperative and mastery learning methods on primary grade students' learning and retention of economic concepts. *Early Education and Development*, 7(3), 253-74.

Levine, D. (1985). *Improving student achievement through mastery learning programs*. San Francisco: Jossey-Bass.

Mason J., Burton L., & Stacey K. (1982). *Thinking mathematically.* London: Addison-Wesley.

Mevarech, Z. R. (1985). The effects of cooperative mastery learning strategies on mathematical achievement. *Journal of Educational Research*, 78(6), 372-377.

Mevarech, Z. R. (1991). Learning mathematics in different mastery environments. *Journal of Educational Research*, 84(4), 225-231.

Monaco, T. (2003). Gifted education research practices recommended in Doctoral Dissertations. *Academic Exchange- EXTRA*. Retrieved June 23, 2009, from http://asstudents.unco.edu/students/AE-Extra/2003/11/Art-3.html.

Mortimore, P., & Sammons, P. (1987). New Evidence on effective elementary schools. *Educational Leadership*, 45 (September): 4-8.

Okebukola, P. A. (1985). The relative effectiveness of cooperative and competitive interaction techniques in strengthening students' performance in science classes. *Science Education*, 69, 501-509.

Panitz, T. (1997). Collaborative versus cooperative learning: comparing the two definitions helps understand the nature of interactive learning. *Cooperative Learning and College Teaching*, 8(2), 1-13.

Schremmer, Hortz, H., & Fries, S. (2001). Testing the knowledge gained in multimedia-enhanced learning. Retrieved September 4, 2009, from http://www.informatik.uni-mannheim.de/pi4/publications/Schremmer2001i.pdf.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, 2(2), 87-99.

Sleeman, D., & Brown, J. S. (1982). Introduction: Intelligent tutoring systems. In D. Sleeman & J. S. Brown (Eds.), *Intelligent* *tutoring systems* (p. 1-11). London: Academic Press.

Springer, L., Stanne, M.E. & Donovan, S. (1999). Effects of small-group learning on undergraduates in Science, Mathematics. Engineering and Technology: A meta-analysis. *Review of Educational Research*, 69(1), 21-51.

Stigler, J.W., Lee, S.Y., Lucker, G.W.. & Stevenson. H.W. (1982). Curriculum and achievement in mathematics: A study of elementary school children in Japan, Taiwan, and the United States. *Journal of Educational Psychology*, 74, 315-322.

Teoh, S. H. (2003). Mistake made in the topic of Matrices among secondary school students. Presented at *Conference on Scientific and Social Research* (CSSR) 2003, the Golden Horses, Seri Kembangan, Malaysia.

Toh, S. C. (1998). *Cognitive and motivational effects of two multimedia simulation presentation modes science learning*. Unpublished Ph.D. thesis, University of Science Malaysia.

Wells, M.A., & Brook, P.W. (2004). Conversational KM - Student Driven Learning. *Sixth Australasian Computing Education Conference (ACE2004)*, Dunedin, New Zealand.

Wu, H. (1999). Basic skills versus conceptual understanding: a bogus dichotomy in mathematics education. *American Educator*, 23(3), 14-19.

Yager, S., Johnson, D. W., & Johnson, R. (1985). Oral discussion groups-to-individual transfer, and achievement in cooperative learning groups. *Journal of Educational Psychology*, 77(1), 60-66.

Zimmerman, J. (1998). *Improving student achievement by extending school: Is it just a matter of time?* The PACE Media/ Education Writers Seminar.

Chen, C.-H., & Howard, B. (2010). Effect of Live Simulation on Middle School Students' Attitudes and Learning toward Science. *Educational Technology & Society, 13* (1), 133–139.

Effect of Live Simulation on Middle School Students' Attitudes and Learning toward Science

Ching-Huei Chen and Bruce Howard*
Graduate Institute of e-Learning, National Changhua University of Education, Changhua City, Taiwan // chhchen@cc.ncue.edu.tw
*Center for Educational Technologies®, NASA-Sponsored Classroom of the Future, Wheeling Jesuit University, Wheeling WV, USA // howard@cet.edu

ABSTRACT
This study examined the effect of live simulation on students' science learning and attitude. A total of 311 middle school students participated in the simulation, which allowed them to access and interpret satellite data and images and to design investigations. A pre/post design was employed to compare students' science learning and attitude before and after the simulation. The findings revealed positive changes in students' attitudes and perceptions toward scientists, while male students had more positive adoption toward scientific attitudes than females. The study also found that the change in student's science learning was significantly influenced by the teacher. Hence, teacher classroom preparation for the simulation experience proved vital to students' attitudes toward science as well as their scientific understanding. Implications for effective use of simulation to increase science-related career awareness and inform effective teaching practice are shared and discussed.

Keywords
Simulation, Science attitude, Videoconferencing, and Teaching practice

Introduction

There has been a prolonged discussion on the use of technology as cognitive tools for teaching and learning. Jonassen and Reeves (1996) characterized cognitive tools as "technologies that enhance the cognitive powers of human beings during thinking, problem solving, and learning" (p.693). Technology holds great potential for students to develop deeper knowledge and execute reflective thoughts by the specific tasks that they otherwise will not have access to. Technology also provides capabilities to complement students' learning styles and multiple intelligences.

Accordingly, simulation has emerged as one of the most popular instructional tools for delivering quality instruction. The use of realistic simulation often requires students to apply newly acquired skills while motivating them toward advanced learning (Hsu & Thomas, 2002; Lewis, Stern, & Linn, 1993; Moreno & Mayer, 2007; Weller, 2004). Frequently, students participating in a simulation perceived the experience as helpful in providing a clear context for the application of learned knowledge and in being a motivating experience (Spinello & Fischbach, 2004). Previously, Jarvis and Pell (2005) investigated the effect of live simulation experience on middle school students in the United Kingdom. They found such experience engaged students in performing expert-like thinking and acting as real scientists would to analyze and assess real-time data. Although the development of technology-rich learning environments has progressed greatly in recent years, researchers have just begun to signify the application of technology for science learning and its impact on students' achievement (Jonassen, 2003; Kim & Reeves, 2007).

In this paper, we sought to review previous research on scientific inquiry-based learning and its impact on students' attitudes toward science, evaluate how simulation would be an ideal way to support inquiry learning and promote positive attitudes toward science, and describe potential underpinning factors to optimize a successful simulation experience for students.

Theoretical Perspectives

Inquiry science learning

Scientific concepts are complex, highly technical, and often considered among the most difficult subjects to teach K-12 students. Recently, many science educators have advocated an inquiry-based approach to learning science in which students are given opportunities to actively build scientific knowledge by asking overarching questions, planning strategies, exploring solutions, constructing new knowledge, and reflecting on their own inquiry process

(Linn, 2000). The learning sciences community agrees that deep and effective learning is best promoted by situating learning in purposeful and engaging activity (Cognition and Technology Group at Vanderbilt, 1993). The goal of this study is to use the technology to mimic the real-world scientific activities and make the inquiry processes become salient to the students.

Technological Support for Science Learning and Attitudes

Several longitudinal investigations into the use of technology in students' science, technology, engineering, and mathematics (STEM) learning are ongoing, but very little attention has been given to discovering the outcomes of such endeavors (Boyle, Lamprianou, & Boyle, 2005). Technology can help in the scientific learning process because of its potential to support activities such as data collection, visualization, meaningful thinking, problem solving, and reflection. In fact, much of our current educational practice grows out of curriculum reform efforts that have emphasized the teaching of process skills involved in the construction of scientific knowledge—diverse skills such as observing, classifying, measuring, conducting controlled experiments, and constructing data tables and graphs of experimental results (Linn & Hsi, 2000). Various strategies to promote better science learning have been explored. For example, the Web-based Inquiry Science Environment (WISE) is one of the curriculum projects that Linn and her colleagues created to help students develop more cohesive, coherent, and thoughtful accounts of scientific phenomena (Linn, Clark, & Slotta, 2003). WISE is guided by an instructional framework called scaffolded knowledge integration (SKI) that requires students to reflect on their deliberately developed repertoire of models for complex phenomena, and to work toward expanding, refining, reconciling, and linking these models (Bell, 2002; Linn, 1995). In WISE, students who engage in various inquiry activities such as compare ideas, distinguish cases, identify the links and connections among ideas, seek evidence to resolve uncertainty and categorize valid relationships show better understandings of complex scientific concepts (Davis & Linn, 2000). Further research into how the effects of using technology-mediated tools to facilitate science practices, such as applying various real data to empower students to understand the scientific enterprise itself, are worth further discussion.

Researchers have long discussed whether students' positive attitudes toward science can influence whether students consider science as a career (Papanastasiou & Zembylas, 2004). Several studies have found that students' attitudes toward science correlated with science achievement and participation in advanced science courses (e.g., Lee & Burkam, 1996; Simpson & Oliver, 1990). It is also well known that students' attitudes toward a subject as well as their learning environment impact school achievement. In this study we sought to understand whether a technology-supported simulation learning environment can improve students' positive attitudes toward science subjects and careers.

Teaching Practice on Science Learning and Attitude

High-quality teachers are essential to improving the teaching and learning (Darling-Hammond, 1997). Teaching practice and instructional decisions influence the quality of students' academic performance and their motivation, effort, and attitudes toward school and academic pursuits (Hidi & Harackiewicz, 2000). They also promote or reduce students' learning and achievement (Hardre & Chen, 2005). Research involving both secondary and older students appears to indicate a relationship between teacher behaviors and students' attitudes toward science (Haladyna, Olsen, & Shaughnessy, 1982; Myers & Fouts, 1992). Children with positive attitudes toward science are more likely to be found in classrooms that have high levels of involvement, teacher support, and use of innovative teaching strategies (Myers & Fouts, 1992). Teachers who lack ability, confidence, and enthusiasm for the subject tend to use less stimulating, more didactic methods and do not respond effectively to students' questions (Osborne & Simon, 1996). Those teachers also are more likely to have students with poor attitudes toward science. Effective teachers adapt learners' needs and evaluate how information should be presented. To meet these demands, teachers need to adjust instruction to student ability levels and background. In fact, one study showed that teachers' teaching style and instructional decisions are the most noticeable factors in students' attitude toward science (Jarvis & Pell, 2005). Therefore, we surveyed teachers about their teaching practices to understand how their approaches might affect students' knowledge and attitude toward science as the result of a simulation learning experience.

The Challenger Mission: Emergency Responsive Learning Experience

The live simulation learning experience conducted for this study originates from the Challenger Learning Center at Wheeling Jesuit University in Wheeling, WV, one of more than 50 Challenger Learning Centers in the United States, Canada, and the United Kingdom. These centers for space science were created in memory of the space shuttle Challenger. More than 25,000 students fly missions each year through the Wheeling facility, and it has been honored nine times for having served the most children of all the centers.

The live simulation, or e-Mission™, is a real-world adventure delivered into the classroom via distance learning technology. With the use of the internet and video conferencing equipment, these live simulations take place in the classroom by a flight director at mission control from the Challenger Learning Center at Wheeling Jesuit University. The learning approach is a student-centered, team-based, interactive educational experience that encourages students to use scientifically accurate data to solve problems. Before the live simulation, teachers conduct a pre-mission preparation for their students, which covers all the mission materials needed for the culminating "live" event. On the mission day, students are assembled into emergency response teams. Via the Internet and videoconferencing equipment, teams connect to a flight director at mission control in Wheeling. The emergency response teams work together and with mission control to handle a "live" problem while the scenario unfolds. Every few minutes new data is delivered to the classroom. Students perform calculations, create graphs, assess the situation, and make decisions based upon their analysis of the data.

The Challenger Learning Center offers a number of e-Missions for all age groups covering mathematics, Earth science, and biology. In this study we researched teachers and students who participated in the Operation Montserrat e-Mission, which uses actual data collected during a 1996 volcanic eruption on the Caribbean island of Montserrat and during a hurricane that hit the island a few years earlier.

Methods

Subjects/Procedures

The participants were 311 (186 males; 125 females) middle school students and 7 teachers from West Virginia, Ohio, Pennsylvania, and New York. Before the e-Mission teachers attended a one-day workshop at the Challenger Learning Center, where they covered Earth science curriculum and learned about the study procedure. Teachers and students completed surveys before any mission-related activity. Teachers spent a week preparing students for the mission. Although all of the participating teachers received one-day professional training on the e-Mission, the actual classroom implementation and time allocation for the mission preparation was left up to each teacher. On the mission day classrooms connected with Challenger Learning Center via videoconferencing to interact and solve problems with a flight director. The scenario unfolds as the Soufriere Hills volcano is ready to erupt while at the same time a Category 3 hurricane is approaching Montserrat from the east. Students worked in teams to take charge of different tasks. The volcano team calculated rock fall and volcanic tectonic data to predict what would happen with the volcano. The hurricane team tracked the approaching hurricane and estimated when it would arrive on the island. The evacuation team used population maps and available transportation options to move residents out of danger zones to shelters on the island. The communications team informed mission control about the situation brewing on the island and relayed recommendations from all teams. All the data were real and sent from the satellite every 5-6 minutes. The length of the live simulation activity was approximately two hours. A week after the mission, students and teachers took the post-surveys.

Surveys/Instruments

Attitudes toward Science

Osborne, Simon, and Collin (2003) suggested that attitudes toward science is not a unitary construct, but "rather of a large number of subconstructs, all of which contribute in varying proportions towards an individual's attitudes towards science" (p. 1054). Our study used the Test of Scientific-related Attitudes (TOSRA), designed by Fraser (1978) from Klopfer's Classification (Klopfer, 1971) because this instrument contains more focused scales to measure the sub-constructs of attitude toward science among middle school students. The instruction for completing the TOSRA survey was given in the beginning: This test contains a number of statements about science. You will be asked what you think about these statements. There are no "right" or "wrong" answers. Your opinion is what is wanted. For each statement, draw a circle around the specific numeric value corresponding to how you feel about each statement. 5 as strongly agree (SA), 4 as agree (A), 3 as uncertain (U), 2 as disagree (D), and 1 as strongly disagree (SD). TOSRA with a total of 70 items includes seven distinct science-related attitudes. The first factor is social implications of science. Here's a sample statement measuring it: "Public money spent on science in the last few years has been used wisely." The second subscale is normality of scientists: "Scientists do not have enough time to spend with their families." Attitude toward scientific inquiry is the third subscale: "I would prefer to do experiments than to read about them." The fourth subscale is adoption of scientific attitudes: "Finding out about new things is unimportant." Enjoyment of science lessons is the fifth subscale: "Science lessons are a waste of time." The sixth subscale is leisure interest in science: "I would like to belong to a science club." The last subscale is career interest in science: "Working in a science laboratory would be an interesting way to earn a living."

Teacher Content Knowledge Covered

A content knowledge covered sheet was given to the teachers asking for the information on the kinds of topics that teachers covered during their classes and the information on the level of coverage for each science vocabulary. The ratings are: 1 as not at all mentioned; 2 as mentioned briefly in class; 3 as discussed in class or observed in homework, and 4 as covered thoroughly in class or through homework.

Student Content Knowledge Pre- and Post-Tests

The pretest and posttest each consisted of 40 items. The items represented the following categories: near transfer, far transfer, and selected items from standardized testing. In this study, the tests include two forms and contain respectable reliabilities (A=0.79; B=0.86) on the previous study (Howard, 2004). Tests were administered by participating teachers. Choice of test form and order of test form administration was left up to each teacher (i.e., A-A; A-B; B-A, B-B).

Data Analysis

A one-way ANOVA (analysis of variance) was performed to examine the changes on students' science attitudes and learning before and after participating in the live simulation. A Pearson coefficient correlation was then conducted to examine whether there is a relationship between students' content knowledge tests, TORSA, and teachers content knowledge covered.

Results

The one-way ANOVA showed a significant difference on the normality of scientists before and after the live simulation experience, F (1, 557) =5.00, p<.05. Table 1 presents their mean scores and standard deviations for the TOSRA and its scales. Students' perception toward normality of scientists increased from 3.14 to 3.22.

The results showed significant gender differences on adoption of scientific attitudes and career interest in science. Male students (Mean=3.08, SD=0.42) showed significantly higher adoption of scientific attitudes than did female students (Mean=3.01, SD=0.41), F (1,550) =4.13, p<.05. Male students (Mean=3.00, SD=0.39) also showed significant higher interest in science careers than did female students (Mean=2.88, SD=0.37), F (1, 541) =5.10, p<.05.

Additionally, the results showed that there was a significant difference on the students content knowledge tests, F (1, 510) =8.13, p=0.005. A further analysis showed that there is a significant difference on students content knowledge among different teachers, F (4, 369) =7.94, p<.001, effect size=0.81, power=.99. Table 2 shows overall students content knowledge test scores by teachers. Teacher B and F students showed significant higher students test results than Teachers A, C, D, and E after the live simulation. Furthermore, the Pearson test revealed significant correlation resided between teachers' content knowledge covered on students' content knowledge tests and attitudes toward science. Specifically, teachers who covered most of science vocabulary thoroughly in class or through homework had major impact on students' content knowledge and positive attitudes toward science.

Table 1. Descriptive Statistics of Students' TOSRA Before and After the Live Simulation

	0-pre/ 1-post	N	Mean	Std. Deviation
Social implications of science	0	279	3.09	0.46
	1	287	3.09	0.43
Normality of scientists*	0	277	3.14	0.40
	1	282	3.22	0.38
Attitude toward scientific inquiry	0	283	3.20	0.46
	1	286	3.18	0.44
Adoption of scientific attitudes	0	279	3.03	0.43
	1	288	3.05	0.41
Enjoyment of science lessons	0	275	2.93	0.48
	1	284	2.92	0.42
Leisure interest in science	0	270	3.05	0.50
	1	281	3.00	0.45
Career interest in science	0	273	2.91	0.38
	1	285	2.92	0.38

*p<.05

Table 2. Descriptive Statistics of Students' Content Knowledge Tests Before and After the Live Simulation by Teachers

Teachers	0-pre/ 1-post	N (of Students)	Mean	Std. Deviation
A	0	21	11.57	4.59
	1	19	11.58	2.67
B*	0	14	11.86	3.98
	1	12	15.75	4.05
C	0	51	10.14	2.69
	1	38	10.61	3.23
D	0	77	10.51	3.05
	1	91	13.99	2.51
E	0	22	11.92	3.44
	1	24	12.05	4.13
F*	0	24	10.98	3.79
	1	21	15.46	2.14
G	0	27	11.49	4.07
	1	25	12.36	2.36

*p<.05

Discussion/Implications

The usefulness of technology simulation as a method for learning has been applied sporadically within education. This study supports the significant role of simulation in transmission of knowledge for educational purposes. We found that participation in the live simulation may have influenced students' attitude toward science over time. Students' normality of scientists was one of the science-related attitudes that showed significant change. This result is confirmed by prior research into realistic simulations showing the relevance of science and change in how students perceive scientists (Jarvis & Pell, 2005). This has significant implication for promoting STEM-related career awareness. As the nation strives to increase students' knowledge of STEM-related careers, the live simulation learning environment has potential for changing students' self-perception and goal orientation. Such high-tech, computer-assisted cooperative simulation of a real-life situation helped to trigger application learning as well as professional identity/awareness/interest/construction.

Additionally, we found there is a difference in gender toward science attitudes and interest in science careers. Boys tended to be more adoptable and receptive to the notion of science attitudes, such as discovering new things, and more interest in careers toward science. Gender has been characterized as the most significant variable towards students' attitude to science (Gardner, 1975). Previous research has shown that boys have a consistently more positive attitude to school science than girls (Becker, 1989; Breakwell & Beardsell, 1992; Hendley, Stables, & Stables, 1996; Weinburgh, 1995). With regards to the career choices, Whitehead (1996) discussed gender stereotype may influence career choices. For example, boys are more likely to choose sex-stereotype careers than girls. Despite of gender stereotype, the findings of this study showed that the live simulation may have increased boys' interests in science or science careers. Therefore, future research should continue to study how the instructional content or technology-enhanced learning environment will lead to a significant increase in the choices of science-related careers by girls.

We also recognized that the potential gain depends on the quality of teacher preparation as well as teacher's instructional strategy. From this study, we learn that teachers who spent more time on science vocabulary had impacted students' content knowledge and attitudes toward science. Such support seems to build up students' knowledge about the subject matter and support self-confidence so that students can participate effectively and make real gains in the complex simulated learning environment. The study reaffirms the value of a synergy between effective teaching practice and use of simulation in optimizing students' learning of science.

Future Research

Our next step is to extend current findings by conducting an experimental study comparing students who experience live simulation to those who do not. We also can further our research on the live simulation by investigating what kinds of teaching strategies (e.g., coaching, inquiry, or mentoring) best align with a simulated learning environment.

Note

An earlier version of this article was presented at the 2008 annual meeting of the American Educational Research Association in New York. We thank the Challenger Learning Center at Wheeling, WV for their support of data collection, and teachers and students who participated in the simulation and the study.

References

Becker, B. J. (1989). Gender and science achievement: a re-analysis of studies from two metaanalyses. *Journal of Research in Science Teaching, 26*, 141-169.

Bell, P. (2002). Using argument map representations to make thinking visible for individuals and groups. In T. Koschmann, R. Hall & N. Miyake (Eds.), *CSCL 2: Carrying forward the conversation.* (Vol. 2, pp. 449-505). Mahwah, NJ: Lawrence Erlbaum Associates.

Boyle, B., Lamprianou, I., & Boyle, T. (2005). A longitudinal study of teacher change: What makes professional development effective. *School Effectiveness & School Improvement, 16*(1), 1-27.

Breakwell, G. M., & Beardsell, S. (1992). Gender, parental and peer influences upon science attitudes and activities. *Public Understanding of Science, 1*, 183-197.

Cognition and Technology Group at Vanderbilt (1993). Anchored instruction and situated cognition revisited. *Educational Technology*, 52-71.

Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching.* NY: National Commission on Teaching and America's Future.

Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: prompts for reflection in KIE. *International Journal of Science Education, 22*(8), 819-837.

Fraser, B., J. (1978). Development of a test of science-related attitudes. *Science Education, 62*(4), 509-515.

Gardner, P. L. (1975). Attitudes to science. *Studies in Science Education, 2*, 1-41.

Haladyna, T., Olsen, R., & Shaughnessy, J. (1982). Relation of student, teacher, and learning environment variables to attitudes toward science. *Science Education, 66*(5), 671-687.

Hardre, P., & Chen, C. H. (2005). A case study analysis of the role of instructional design in the development of teaching expertise. *Performance Improvement Quarterly, 18*(1), 34-58.

Hendley, D., Stables, S., & Stables, A. (1996). Pupils' subject preferences at Key Stage 3 in South Wales. *Educational Studies, 22*, 177-187.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the twenty-first century. *Review of Educational Research, 70*(2), 151-179.

Hsu, Y.-S., & Thomas, R. A. (2002). The impacts of a web-aided instructional simulation on science learning. *International Journal of Science Education, 24*(9), 955-979.

Jarvis, T., & Pell, A. (2005). *Secondary pupils of different abilities response to an e-Mission simulation of the Montserrat volcanic eruption.* Paper presented at the American Education Research Association, Montreal, CA.

Jonassen, D. H. (2003). Using cognitive tools to represent problems. *Journal of Research on Technology in Education, 35*(3), 362-379.

Jonassen, D. H., & Reeves, T. C. (1996). Learning with technology: using computers as cognitive tools. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 693-719). New York: Macmillan.

Kim, B., & Reeves, T. C. (2007). Reframing research on learning with technology: In search of the meaning of cognitive tools. *Instructional Science, 35*(3), 207-256.

Klopfer, B. J. (1971). Evaluation of learning in science. In B. S. Bloom, J. T. Hastings & G. F. Madaus (Eds.), *Handbook on summative and formative evaluation of student learning*. New York: McGraw-Hill.

Lee, V. E., & Burkam, D. T. (1996). Gender differences in middle grade science achievement: Subject domain, ability Level, and course emphasis. . *Science Education, 80*(6), 613-650.

Lewis, E., Stern, J., & Linn, M. (1993). The effect of computer simulations on introductory thermodynamics understanding. *Educational Technology, Jan.* , 45-58.

Linn, M. C. (1995). Designing computer learning environments for engineering and computer science: The scaffolded knowledge integration framework. *Journal of Science Education and Technology, 4*(2), 103-126.

Linn, M. C. (2000). Designing the knowledge integration environment. *International Journal of Science Education, 22*(8), 781-796.

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education, 87*(4), 517-538.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners.* Mahwah, NJ: Lawrence Erlbaum Associates.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(309-326).

Myers, R. E., & Fouts, J. T. (1992). A cluster analysis of high school science classroom environments and attitude toward science. *Journal of Research in Science Teaching, 29*(9), 929-937.

Osborne, J., & Simon, S. (1996). Primary science: past and future directions. *Studies in Science Education, 26*, 99-147.

Osborne, J., Simon, S., & Collin, S. (2003). Attitudes towards science: a review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049-1079.

Papanastasiou, E. C., & Zembylas, M. (2004). Differential effects of science attitudes and science achievement in Australia, Cyprus, and the USA. *International Journal of Science Education, 26*(3), 259-280.

Simpson, R. D., & Oliver, J. S. (1990). A summary of major influences on attitude toward science and achievement in science among adolescent students. *Science Education, 74*(1), 1-18.

Spinello, E. F., & Fischbach, R. (2004). Problem-based learning in public health instruction: A pilot study of an online simulation as a problem-based learning approach. *Education for Health, 17*(3), 365-373.

Weinburgh, M. (1995). Gender differences in student attitudes toward science: a meta-analysis of the literature from 1970 to 1991. *Journal of Research in Science Teaching, 32*, 387-398.

Weller, J. M. (2004). Simulation in undergraduate medical education: Bridging the gap between theory and practice. *Medical Education, 38*, 32-38.

Whitehead, J. M. (1996). Sex stereotypes, gender identity and subject choice at A level. *Educational Research, 38*, 147-160.

Yeh, Y.-C. (2010). Analyzing Online Behaviors, Roles, and Learning Communities via Online Discussions. *Educational Technology & Society, 13* (1), 140–151.

Analyzing Online Behaviors, Roles, and Learning Communities via Online Discussions

Yu-Chu Yeh
Institute of Teacher Education; Research Center for Mind, Brain & Learning; Center for Creativity and Innovation Studies; National Chengchi University, Taipei 116, Taiwan // ycyeh@nccu.edu.tw

ABSTRACT

Online learning communities are an important means of sharing and creating knowledge. Online behaviors and online roles can reveal how online learning communities function. However, no study has elucidated the relationships among online behaviors, online roles, and online learning communities. In this study, 32 preservice teachers participated in an 18-week instruction program. Analyses of online group discussions revealed the following: (a) of thirteen identified online behaviors, the most common were constructing a positive atmosphere, providing opinions for group assignments, and providing reminders of assignment-related work; (b) of eight online roles identified within a group, the most common roles were information providers, opinion providers, and troublemakers; (c) four online learning communities based on "collaboration" and "participation" were identified. The evolution of these online learning communities indicates the interrelationships among online behaviors, roles, and learning communities.

Keyword

Behavior, Learning community, Preservice teacher, Online discussion, Roles

Introduction

The emergence of learning communities is an interesting and recent pedagogical development in higher education. Various strategies have been developed to foster learning communities in an online setting. The objectives of these strategies include communicating effectively, strengthening social ties, collaborating in small teams, establishing social networks, and collaborating in knowledge construction (e.g., Chang, Chen & Li, 2006; Jones & Issroff, 2005; Wang & Poole, 2004; Yang, Wang, Shen & Han, 2007). Online behaviors and roles that are fundamental to the functioning of online learning communities, however, have seldom been compared (Yang et al., 2007). Moreover, although a few studies (e.g., Cho, Gay, Davidson & Ingraffea, 2007; Lin, Lin & Huang, 2007) have attempted to define online learning community styles or types, none has developed clear criteria for defining online learning community types from a holistic viewpoint. For instance, Lin et al. (2007) classified the products and processes of knowledge sharing and creation in a professional virtual community into six types. However, they did not define these groups according to their characteristics. Although another study by Cho et al. (2007) defined the styles of online learning communities, it only focused on willingness to communicate in learning communities.

Online learning communities are a collaborative means of achieving "shared creation" and "shared understanding," in which mutual exchange between community members are encouraged to support individual and collective learning (Ludwig-Hardman & Woolley, 2000). Some studies indicate that online learning communities promote active participation, increase academic achievement, contribute to knowledge creation, and improve learner cognitive abilities (e.g., Lin et al., 2007; Ludwig-Hardman & Woolley, 2000; Moller, 1998; Waltonen-Moore, Stuart, Newton, Oswald & Varonis, 2006). However, the question of how these benefits are obtained remains unanswered. Lin et al. (2007) found that for any group to perform well via an online setting, group members must recognize their functional roles in knowledge-related activities, and each functional role requires a corresponding behavior in the processes of knowledge sharing and creation. Therefore, identifying the important online roles and their corresponding behaviors should elucidate how online learning communities function, what online learning communities can be formed, and which online learning communities best benefit learners. This information can help teachers improve their e-learning instruction methods. Briefly, this study examines online roles and corresponding behaviors exhibited in an online learning community and further, based on these analyses, develops objective criteria for categorizing online learning communities.

Online behaviors, roles, and learning communities

Online behaviors and roles

Chang, Cheng, Deng, and Chan (2007) identified the following ten basic elements of structured network learning societies: participants, shared visions, devices, services, rules, relations, manners, learning domains, learning goals, and learning activities. They argued that because participants are the lifeblood of online learning groups, identifying participants by analyzing "roles" is crucial for identifying the interpersonal behaviors of network community members. Similarly, Yang et al. (2007) determined that learning communities inevitably include learners with similar approaches and different interests and that each learning behavior reflects learner interests as well as their resource and information needs. Accordingly, analyzing behaviors and roles in an online learning community is essential for understanding online learning communities.

According to the *Dictionary of Psychology* (Corsini, 2002), a role is "the set of behaviors expected of a person possessing a certain social status" (p. 850). Accordingly, a role is an upper-level concept of a behavior and can comprise a set of behaviors. To date, few studies have analyzed online roles and behaviors or clarified their corresponding relationships. In a study of online learner roles, Lin et al. (2007) compared widely varying examples of inferior and superior consequences of special-interest groups at several group levels. Their analysis of group roles revealed that inferior group roles comprised information/opinion seekers or givers, encouragers, and followers whereas superior group roles included initiators, orienters, encouragers, recorders, gatekeepers, information/opinion seekers or givers, coordinators, and clowns. Lin et al. also found that, for knowledge-creation roles, the inferior group is primarily comprised of idea providers whereas the superior group consists of task performers followed by idea providers and integrators. Similarly, Agre (1998) noted the critical importance of "thought leader" for building trust within a community; thought leaders are individuals who foresee issues, gather positions and arguments, network with people relevant to the issue, and articulate the issue in a manner that provokes thinking by individual community members.

In their further analysis of online behaviors or strategies, Lin et al. (2007) classified collaborative strategies into two categories: task performance and team maintenance. Task performance strategies are related to coordination tasks for problem solving or goal attainment, such as initiating, seeking information/opinions, providing information/opinions, coordinating, orienting, evaluating and recording. On the other hand, team maintenance strategies such as encouraging, gate-keeping, following and clowning build friendly relationships among group members and maintain team coherence. Lin et al. concluded that each functional role requires a corresponding behavior in knowledge-sharing and -creation processes. However, the relationship between roles and behaviors has not been clarified. The present study therefore attempts to elucidate these relationships.

Indices for online learning communities

Cooperation and motivation to participate are two crucial indices for distinguishing between the achievement of online groups (Guzdial & Turns, 2000; Lin et al., 2007). Lin et al. (2007) found that while over 50% of participants in the superior group habitually cooperated, few participants in the inferior group did so. They also indicated that participants in the superior group were more enthusiastic about sharing knowledge than those in the inferior group. Similarly, Ligoria (2001) proposed that when communities are organized into groups consisting of members with different abilities, the overall purpose of the community must be kept in mind along with a sense of collaboration.

According to Ligorio (2001), the collaborative dimension of knowledge building comprises the community of learners model and the community of practices model. In the learners model, each learner is invited to formulate problems and hypotheses, search for solutions, share knowledge, explore new fields, learn about new topics, and adopt new perspectives. In the practices model, learning is a function of an activity, context, culture and social interaction between people with different competencies. Consequently, peripheral participation is legitimate; that is, even when not directly participating in an activity, learners can still benefit from observation, analysis, and discussion of that activity (Ligorio, 2001). Clearly, both the community of learners and the community of practices models assume the interdependence of participants during cognitive learning (Salomon, 1993; 1998). Therefore, collaboration is needed to build knowledge in online learning communities.

The second index of online learning communities, participation, is considered a general measure of successful online discussions (Guzdial & Turns, 2000). Notably, although online learning communities have considerable potential for encouraging students to construct and share knowledge, in most online discussions, only a few key students actively do so (Chang, Chen & Li, 2006). Moreover, frequent messaging does not constitute a genuine community (Guzdial, & Turns, 2000). However, from a meaning-making perspective, the content and context of messages are critically important. In support of this perspective, Havelock (2004) suggested that, although the number and frequency of connections provide a sense of community activity, they say little about how these interactions impact identity formation, meaning-making, and the professional practices of participants. To determine which messages are meaningful, Baym (1998) proposed that message content in a group should contribute to the development of intra-community trust, and such messages are typically characterized by optimism, excitement, clear task orientation, and shared leadership duties. In the same vein, Agre (1998) advocated the importance of facilitating a sense of group trust and participation. Thus, collaboration and meaningful participation are clearly two important indices for measuring the success of online learning communities. These two indices must also be utilized when categorizing learning communities.

Research questions

Because this study is exploratory, only research questions rather than hypotheses are proposed. The principal research questions are as follows.

1. What online behaviors are exhibited during interactive online discussions?
2. What online roles evolve from online behaviors?
3. What are the relationships among online behaviors, online roles, and online learning communities? Specifically, how do online behaviors and online roles evolve into objective criteria for categorizing online learning communities?

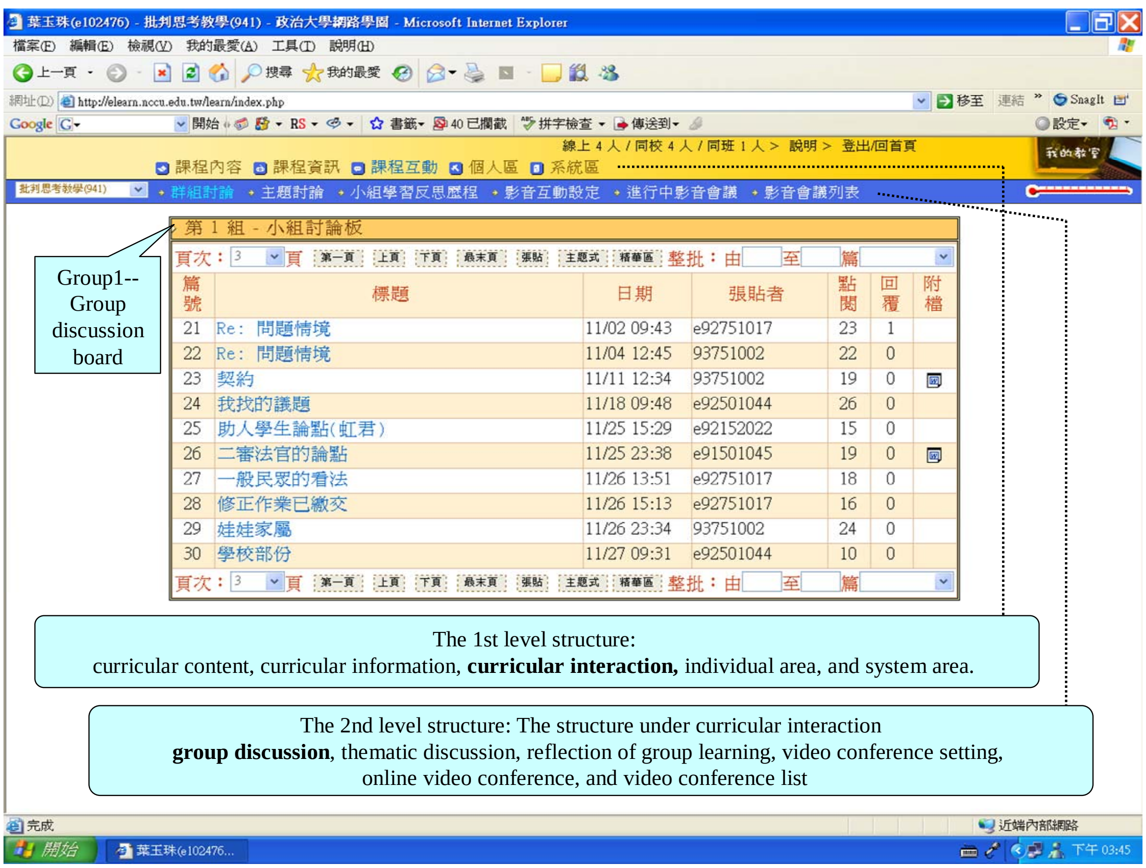

Figure 1. An example screen of group discussion board

Method

Participants

The study participants were 32 preservice teachers (6 males and 26 females) enrolled in the Instruction in Critical Thinking class in a teacher-training program for secondary school teachers. Among the participants, 14 (43.75 %) were undergraduates, and 18 were graduates (56.25 %). Mean subject age was 23.00 years (*SD* = 2.54 years).

Instruments

The research instrument employed in this study was the e-learning website developed by National Chengchi University. The e-learning interface consists of three levels. The first level includes the functions of curricular content, curricular information, curricular interaction, an individual area, and a system area. The instructional design of this study required that participants complete several group assignments and engage in online discussions. Consequently, "curricular interaction," particularly the "Group Discussion Board" under this function, became the most frequently used interface. Figure 1 presents an example screen of this Group Discussion Board. Since the interface is written in Mandarin, the main functions of the menu bars are translated.

Procedures and instructional design

An 18-week experimental instruction program based on teaching critical thinking was developed to encourage the formation and use of online learning communities. To achieve these two goals, collaborative problem-based learning (PBL) was incorporated into the experimental program. According to Lee and Kim (2005), collaborative PBL is a method in which learners share a common goal, perform given tasks at the same level, and interact with each other during problem solving. Accordingly, collaborative PBL, which emphasizes the importance of interactive discussions, is ideal for analyzing online learning communities. Specifically, the instructional design had two phases: formation of online learning communities (weeks 1–7) and use of online learning communities (weeks 8–16).

In the first phase, participants were divided into six groups; each participant was allowed to select the group of his or her own choice. Each group consisted of five to six members. However, one participant in Group 4 dropped out during the semester; therefore, Group 4 was comprised of only four participants. During the second week, they started preparing for their group project, which employed collaborative PBL. During this instruction period, the researcher encouraged the formation of online learning communities by assigning group work on the following topics: (1) develop test items for five critical-thinking skills—assumption identification, induction, deduction, interpretation, and argument evaluation; (2) develop a situation-based problem; and (3) apply strategic thinking to everyday problems.

In the second phase, participants were scaffolded to complete the collaborative PBL assignment via online learning communities. The primary group tasks in this instructional period were as follows: (1) find an authentic case for collaborative PBL; (2) define problems in the case; (3) decide roles in the case; (4) develop arguments for each role; (5) present all arguments and a consensus of solutions via concept maps; and, (6) present arguments for each role and role-play the problem-solving process.

The researcher assumed that the instructional design would prompt participants to take advantage of the online discussion board, especially during group discussions. First, participants were from various departments, and each participant was enrolled in several classes. Thus, face-to-face discussions were difficult to organize. Second, participants were asked to discuss group assignments online and then upload their files to the e-learning website during both phases of the study.

Analyses

The Group Discussion Board content was analyzed. As mentioned, an online role can comprise several online behaviors, and not all online behaviors contribute to the formation and function of an online community (Corsini, 2002; Guzdial & Turns, 2000). Restated, among the many online behaviors, only some can be further combined into online roles that contribute to group trust and participation. Further, identifying online roles is critical for understanding an online learning community (Agre, 1998; Lin et al., 2007). Consequently, to achieve the goals of this study, the online behaviors of participants were determined first. These behaviors were then used to determine a set of online roles likely to influence the formation of an online learning community. When determining the number of online roles, the researcher took the related findings in the applicable studies (Agre, 1998; Lin et al., 2007) as a referenced framework and then tried to propose more elaborate categories based on the data obtained in this study. Finally, based on the analyzed roles, different online learning communities were identified. The online behaviors and online roles in this study were identified via discussions of the researcher and two trained graduate students.

Results

Analyses of online behaviors

Table 1 lists the frequencies of discussions on the Group Discussion Board. The analytical results indicate that most discussions were conducted during weeks 8–16 via asynchronous discussions when participants started working on their PBL projects. The mean number of asynchronous and synchronous discussions for each participant was 36.00 and 4.72, respectively.

Table 1. The frequencies and means of online discussions

		Asynchronous discussions		Synchronous discussions	
Group	*N*	Count	*M*	Count	*M*
G 1	5	110	22.00	0	0
G 2	6	202	33.67	25	4.17
G 3	6	176	29.33	30	5.00
G 4	4	170	42.50	0	0
G 5	5	193	38.60	31	6.20
G 6	6	301	50.17	65	10.83
Total	32	1152	36.00	151	4.72

Analyses of interactions in the asynchronous and synchronous discussions identified the following 13 online discussion behaviors.

1. Providing opinions for group functioning: Such behaviors helped the group function effectively and efficiently. For example, “We should upload personal assignments before Sunday night to ensure efficient discussion on Tuesday.”
2. Providing opinions for group assignments: Such behaviors referred to personal responses to member opinions or ideas related to group assignments. For example, “The suggested story is good, but it’s kind of hard to discuss a gang leader.”
3. Encouraging opinions about/responses to group assignments: Such behaviors were observed when the deadline was approaching, but no one had posted any opinions about the assignments. These behaviors were also observed when personal opinions had been posted, but no one had responded to these opinions. For example, “Everyone posts your opinions on the discussion board; the deadline is Monday.”
4. Sharing information: Such behaviors were related to the sharing of information obtained from the teacher, media, magazines, websites or other sources. For example, “I recently read a magazine article that discussed bullies in schools. Maybe ‘bullies in schools’ can be a topic of our project. What do you think?”
5. Clarifying concepts: Such behaviors were performed to clarify misconceptions about an issue. For example, “The test item you proposed is not about ‘assumption identification’; it is about ‘explanation’.”
6. Constructing a positive atmosphere: Such behaviors included giving encouragement and blessings as well as expressing gratefulness, caring, and forgiveness. For example, “Two groups have decided to take ‘bird flu’ as their project topic. You should continue to exercise to help prevent infection of a bird flu.”
7. Answering questions: Such behaviors occurred when a group member had questions about the assignment or distributed work and asked for help on the discussion board. For example, “Julie: ‘I don’t know what I should do.’ Albert: ‘You need to write a learning contract and then upload it to the Web site by Monday.’”

8. Providing reminders of assignment-related work: Such reminders were related to meeting times, assignment content and progress, and distribution of assignments. For example, “When uploading the assignment, don’t forget to list the filename as ‘Group 2’.”
9. Explaining personal problems: Students explaining personal problems typically posted excuses or reasons for being unable to participate in group discussions or unable to finish assigned work on time. For example, “Sorry, I was sleepy last night, so I forgot to upload the file.”
10. Explaining the problems of others: Such behaviors were performed to tell group members why someone could not participate in group discussions or complete their assignment on time. For example, “Teresa has a class until 1:00 p.m., so, she will come later.”
11. Solving problems: Such behaviors were performed to work out problems that could hinder group progress. The most frequently encountered problems were a group member forgetting to upload an assignment or not distributing an assignment to all group members. For example, “The deadline is coming, but John has not uploaded his file. I have just uploaded the file by myself.”
12. Setting schedules: Such behaviors occurred when no group members had proposed a specific time for discussions. For example, “We should discuss our project topic this Friday.”
13. Assigning work: Such behaviors included asking group members to be responsible for certain work or asking for volunteers to complete work. For example, “Is there any volunteer to complete the learning contract?”

These online behaviors were counted (Table 2) to determine their frequency. Of these 13 behaviors, B6 (constructing a positive atmosphere) was the most frequent, followed by B2 (providing opinions for group assignments) and B8 (providing reminders of assignment-related work). The B10 behavior (explaining the problems of others) was the least common.

Table 2. The employed 13 online behaviors

	Type of online behavior												
	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13
Count	19	80	16	36	18	90	10	80	48	3	8	8	10

Note. B1 to B13 represent type 1 to type 13 online behaviors, respectively.

Analyses of online roles

Based on the data obtained in this study, the findings in related studies (Agre, 1998; Lin et al., 2007), and the definition that a role may comprise a set of behaviors (Corsini, 2002), the researcher tried to combine the aforementioned 13 behaviors into some ‘meaning-making’ roles. Finally, the following eight roles were identified.

1. Supervisors (R1): This role comprises B1 and B3. This role was the key to good group functioning. Supervisors gave suggestions about creating high-quality work, requested opinions from group members, set discussion schedules and assigned work to group members.
2. Information providers (R2): This role consists of B4. These group members typically provided and shared information related to assigned work.
3. Group instructors (R3): This role consists of B5. These group members attempted to clarify misconceptions.
4. Atmosphere constructors (R4): This role consists of B6. These group members attempted to construct a positive and harmonious atmosphere of support, caring, and cooperation.
5. Opinion providers (R5): This role is composed of B2. These group members provided opinions that contributed to group work.
6. Reminders (R6): This role is composed of B8. These group members were responsible for reminding others about discussion times, assignment deadlines, and the details for completing group work.
7. Trouble-makers (R7): This role is composed of B9. These group members frequently caused problems that hindered the completion of group work via their absence from group discussions or inability to finish assigned work on time.
8. Problem solvers (R8): This role comprises B7, B10, and B11. These group members attempted to answer questions posed by group members as well as to correct and explain problems caused by group members.

To determine which online roles were employed most frequently, the roles of each participant were counted and the outstanding roles were analyzed from within-group and across-group perspectives. Table 3 presents the details concerning role counts for all participants.

Table 3. Counts and means for online roles

Participant	R1	R2	R3	R4	R5	R6	R7	R8
	Type of online role							
				Group 1				
G1-1	3	2	0	1	2	3	1	0
G1-2	1	0	0	0	0	1	0	0
G1-3	1	1	0	1	0	0	0	0
G1-4	0	0	0	0	1	0	0	0
G1-5	1	0	0	0	0	0	2	0
Mean	1.20	0.60	0.00	0.40	0.60	0.80	0.60	0.00
				Group 2				
G2-1	3	1	0	5	4	8	4	3
G2-2	4	2	0	9	2	16	0	2
G2-3	1	1	0	5	0	5	5	1
G2-4	0	0	0	1	0	0	0	0
G2-5	1	0	0	1	0	0	0	0
G2-6	2	0	0	3	0	3	0	0
Mean	1.83	0.67	0.00	4.00	1.00	5.33	1.50	1.00
				Group 3				
G3-1	2	0	0	1	3	0	5	0
G3-2	0	0	0	0	2	3	0	0
G3-3	1	4	0	2	1	1	3	0
G3-4	1	2	0	5	1	2	2	0
G3-5	4	0	0	1	1	1	3	2
G3-6	7	3	0	4	0	8	0	4
Mean	2.50	1.80	0.00	2.17	1.33	2.50	2.17	1.00
				Group 4				
G4-1	1	1	1	3	0	1	3	1
G4-2	0	0	0	1	0	1	1	0
G4-3	2	2	2	0	2	3	1	2
G4-4	0	0	0	4	1	1	1	0
Mean	0.75	0.75	0.75	2.00	0.75	1.50	1.50	0.75
				Group 5				
G5-1	2	0	0	6	4	1	0	0
G5-2	3	1	2	0	2	3	1	1
G5-3	1	2	0	4	4	2	2	2
G5-4	0	1	0	6	6	5	1	1
G5-5	1	0	1	0	7	1	0	1
Mean	1.40	0.80	0.60	3.20	4.60	2.40	0.80	1.00
				Group 6				
G6-1	5	4	5	8	8	3	3	1
G6-2	0	1	0	3	7	0	4	0
G6-3	4	4	4	8	8	4	0	0
G6-4	2	1	0	2	4	2	2	0
G6-5	0	2	1	5	8	1	3	0
G6-6	0	1	1	2	1	1	1	0
Mean	1.83	2.17	1.83	4.67	6.00	1.83	2.17	0.17
				Class				
Total	53	36	17	91	79	80	48	21
Mean	1.66	1.16	0.53	2.84	2.47	2.50	1.5	0.66

An outstanding role within a group was observed when the count of a participant for a certain role exceeded the group mean for that role, and an outstanding role across groups was observed when the count of a participant for a certain role exceeded the class mean for that role. For instance, when analyzed within the group, two participants in Group 3 (i.e., G3-5 and G3-6) were outstanding for R1 (supervisors). Their total number of participation for this role were, respectively, 4 and 7, which were higher than the group mean for this role (M = 2.50). However, when analyzed across groups, three participants in Group 3 (i.e., G3-1, G3-5, and G3-6) were outstanding for R1; their total numbers of participation for this role were 2, 4, and 7, respectively, and these numbers were higher than the class mean for this role (M = 1.66). Restated, in Group 3, participants G3-5 and G3-6 were supervisors, and, when examined in the context of the whole class, G3-1, G3-5, and G3-6 were supervisors. Following these calculations, Table 4 presents the numbers and distributions of outstanding roles for each group.

When analyzed within the group, the roles of information providers, opinion providers, and troublemakers were the most numerous (15, 14, and 14, respectively), and the number of group instructors was the lowest (Table 4). When analyzed across groups, the roles of supervisors and troublemakers were the most numerous (both 13), followed by positive atmosphere constructors, reminders, and problem solvers. Group instructor was the least common.

Table 4. Counts for outstanding roles in each group

	Type of online role							
Group	R1	R2	R3	R4	R5	R6	R7	R8
	Within the group							
G1	1	2	0	0	2	0	2	0
G2	3	3	0	3	2	2	2	3
G3	2	3	0	2	2	2	3	2
G4	2	2	2	2	2	1	1	2
G5	2	3	2	3	2	2	3	4
G6	3	2	2	3	4	3	3	1
Total	13	15	6	13	14	10	14	12
	Across groups							
G1	1	1	0	0	0	1	1	0
G2	3	1	0	1	1	4	2	3
G3	3	3	0	2	1	2	4	2
G4	1	1	2	2	0	1	1	2
G5	2	1	2	3	4	2	1	4
G6	3	3	4	4	5	2	4	1
Total	13	10	8	12	11	12	13	12

Note. R1 to R8 represent type 1 to type 8 online roles, respectively.

Online learning communities

This study attempted to define online learning community types based on two indices: collaboration and participation. These indices were evaluated based on the aforementioned roles. In terms of collaboration, the number of roles was counted to represent a group member's discussion frequency. Specifically, if most group members had similar discussion frequencies, the group was considered "high collaboration," whereas if the discussion frequencies of group members varied significantly, the group was considered "low collaboration." Based on this central idea, this study first summed up the roles contributing to collaboration within each group. Means and standard deviations were then calculated for each group. According to Baym (1998), sense-making message content contributes to the development of intracommunity trust. Troublemakers are clearly harmful to group collaboration; this role was therefore eliminated when summing up collaborative roles in each group. Specifically, the sum of cooperative roles = the sum of all roles − the sum of troublemakers (Table 5).

Table 5. The counts of online roles and collaborative roles

	Type of online role								Sum	
Group	R1	R2	R3	R4	R5	R6	R7	R8	Roles	Collaborative roles
G 1	6	3	0	2	3	4	3	0	21	18
G 2	11	4	0	24	6	32	9	6	92	83
G 3	15	9	0	13	8	15	13	6	79	66
G 4	3	3	3	8	3	6	6	3	35	29
G 5	7	4	3	16	23	12	4	5	74	70
G 6	11	13	11	28	36	11	13	1	124	111

Note. R1 to R8 represent type 1 to type 8 online roles, respectively.

Since the total online discussion count varied greatly between groups, directly comparing SDs between groups would have been inappropriate. Therefore, the coefficient of relative variability (CV) rather than SD was employed to compare individual differences within a group and further helped determine the degree of collaboration for each group. Notably, CV represents the ratio of SD to mean (CV = SD*100/M). The analytical results indicated that Group 1 and Group 2 had comparatively large CVs, 117.19 and 96.49, respectively. Thus, these groups were regarded as "low collaboration." The other groups were regarded as "high collaboration." (Table 6)

Table 6. The Ms, SDs, and CVs for collaborative roles

Group	*n*	Minimum	Maximum	Total	*M*	*SD*	CV
G 1	5	1	11	18	3.60	4.22	117.19
G 2	6	1	35	83	13.83	13.35	96.49
G 3	6	5	26	66	11.00	7.67	69.71
G 4	4	2	13	29	7.25	4.57	63.08
G 5	5	11	19	70	14.00	3.16	22.59
G 6	6	6	34	111	18.50	11.78	63.66

As collaboration was measured according to a within-group perspective, the degree of participation was defined from an across-group perspective. The rationale for this difference is that a participant may have had high participation in comparison with other group members but low participation in comparison with the entire class. This typically occurred when the entire group had low participation. On the other hand, a participant may have had low participation compared with that of his/her group but high participation compared with that of the entire class; this generally occurred when the entire group had high participation. Additionally, since being a trouble-maker is a participation type, this role was included when determining the participation for each group. Restated, the sum of all roles for each group was considered indicative of its participation. Thus, "high participation" in this study was defined as a mean role of a group higher than that of the class, and "low participation" was defined as a mean role of a group lower than that of the class. The analytical results demonstrated that Groups 2, 5, and 6 were classified as having "high participation", and Groups 1, 3, and 4 were classified as having "low participation" (Table 7).

Table 7. Group means and class means for online roles

	Type of online role										
Group	R1	R2	R3	R4	R5	R6	R7	R8	Total	*n*	Mean
G 1	6	3	0	2	3	4	3	0	21	5	4.20
G 2	11	4	0	24	6	32	9	6	92	6	15.33
G 3	15	9	0	13	8	15	13	6	79	6	13.17
G 4	3	3	3	8	3	6	6	3	35	4	8.75
G 5	7	4	3	16	23	12	4	5	74	5	14.80
G 6	11	13	11	28	36	11	13	1	124	6	20.67
Class	53	36	17	91	79	80	48	21	425	32	13.28

Note. R1 to R8 represent type 1 to type 8 online roles, respectively.

Based on the above analysis, a two (collaboration vs. participation) by two (high vs. low) model was proposed. Specifically, four online learning communities were identified: active collaboration, passive collaboration, individualized participation, and indifference (Figure 2). The four online learning community types and their distributions are as follows.

1. Active collaboration (high cooperation and high participation): Groups 5 and 6.
2. Passive collaboration (high cooperation and low participation): Groups 3 and 4.
3. Individualized participation (low collaboration and high participation): Group 2.
4. Indifference (low cooperation and low participation): Group 1.

As Table 7 shows, the active collaboration communities were typically high on R4 and R5; the passive collaboration communities were common in using R4, R6, and R7; the individualized participation community was high on R4 and R6; and the indifference community was high on R1.

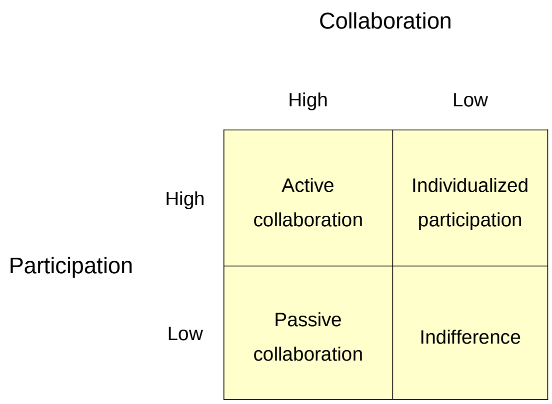

Figure 2. Types of online learning communities

Discussion

This study examines three questions concerning online behaviors, online roles, online learning communities, and their interrelationships. The analytical findings in this study indicate that the three questions are satisfactorily answered. Although this is an exploratory study, it evolves from a pilot study (Yeh, 2005) of 48 preservice teachers enrolled in the same course a year before this study was conducted. The identification of seven online roles in that pilot study provides a framework for this study. To generate a more comprehensive list of online roles than what is presented in the pilot study, this study deliberately starts by analyzing online behaviors. Moreover, to further clarify the relationships among online behaviors, online roles, and online learning communities, an elaborate instructional design is utilized and objective analyses based on online discussions are applied. As expected, the high frequencies of discussions (Table 1) suggest that the instructional design in this study successfully motivates participants to take advantage of online discussions, especially asynchronous discussions. Such participation is essential for objectively analyzing online behaviors and roles of participants as well as for learning community types.

Formation of an online learning community depends on the effectiveness of online learning behaviors (Palloff & Pratt, 1999) and the meaningfulness of exchanged messages (Baym, 1998; Havelock, 2004). Based on these rationales, this study only considers meaningful messages when analyzing online behaviors. The messages not focused on the discussed topics or issues and those not representing personal thoughts (e.g., a simple answer, “Yes”) are screened out. Further, rather than focusing on a specific perspective such as collaborative strategies (Lin et al., 2007), this study analyzes participant behaviors from a holistic perspective. Accordingly, 13 online behaviors are identified. The most frequently utilized behaviors are constructing a positive atmosphere, providing opinions for group assignments, and providing reminders of assignment-related work.

Lin et al. (2007) found that group members recognize their functional roles in knowledge-related activities. Accordingly, Lin concluded that each functional role requires a corresponding behavior to act during the knowledge sharing and creation processes. However, Lin et al. did not further analyze the corresponding relationships between roles and behaviors. Analytical findings in this study provide empirical and descriptive evidence supporting the conclusions obtained by Lin et al. By further integrating the 13 behaviors, the empirical evidence in this study shows that 8 important roles exist in online learning communities, and all participants play multiple roles during online discussions. The analytical results also demonstrate that, although some roles are composed of multiple behaviors, some comprise only one behavior. Moreover, this study analyzes the outstanding roles from different perspectives. From the within–group perspective, the most frequently utilized roles are information providers, opinion providers, and trouble-makers; on the other hand, the most frequently used roles determined using the across-group perspective are supervisors, trouble-makers, positive atmosphere constructors, reminders, and problem solvers. Among these roles, trouble-makers clearly hinder the formation and functioning of online learning communities. Unfortunately, this role typically exists in online learning communities, as the analytical findings in this study suggest. In the within–group context, “group instructor” is the least common role. This analytical finding is expected. Group instructors assist in resolving misconceptions and organizing gathered information. Although such a role is critical for knowledge construction in online settings, not everyone can play this role (Ludwig-Hardman & Woolley, 2000; Waltonen-Moore et al., 2006). As Chang et al. (2007) suggested, identifying participants by analyzing “roles” is essential for understanding the interpersonal behaviors of network community members. The findings obtained in this study are valuable for further analyses of online learning communities.

To define online learning community types, this study employs two indices—collaboration and participation—which have been suggested by many researchers (Agre, 1998; Baym, 1998; Collison, Elbaum, Haavind & Tinker 2000; Havelock, 2004; Ligoria, 2001; Lin et al., 2007). The index of collaboration is derived from the sum of collaborative roles (the sum of all roles − the sum of troublemakers) and CVs, while the index of participation is evaluated based on the mean of total online roles. The analytical results reveal the following four online learning communities: active collaboration, passive collaboration, individualized participation, and indifference. It is also found that while R4 (atmosphere constructors) is commonly found in active collaboration, passive collaboration, and individualized participation communities, R5 (opinion providers) seems to be the key role for distinguishing the active collaboration communities from the other communities. Moreover, R1 (instructors) is exclusively eminent in the indifference community. More specifically, the behavior of constructing a positive atmosphere is commonly used in the active collaboration, passive collaboration, and individualized participation communities; the behavior of providing opinions for group assignment is critical for establishing the active collaboration communities; and the behavior of providing opinions for group functioning and that for encouraging opinions about/responses to group assignments are eminent in the indifference community. It is also found in this study that the active collaboration communities (Group 5 and Group 6) have best performance in the assigned tasks while the indifference community (Group 1) has the worst performance when evaluated by their final grades. When examining the online roles in Table 7, it is determined that the active collaboration communities have all the eight types of roles although the frequencies of these roles are high on R4 (atmosphere constructors) and R5 (opinion providers) and low on R8 (problem solvers). On the other hand, the indifference community is high on R1 (supervisors) and R6 (reminders), but is missing on R3 (group instructors) and R8 (problem solvers). These findings are in line with findings of Lin et al: 'encouragers' exist in both the inferior and superior group, the superior group consists of a greater variety of roles than the inferior group, and the superior group habitually cooperates while the inferior group does not. Accordingly, the relationships among online behaviors, online roles, and types of online learning communities are closely related.

Moreover, the analytical findings in this study suggest that collaborative PBL is a useful tool for exploring online learning communities when instructional activities are well designed. The finding that most groups have frequent online discussions also supports the conclusion obtained by Hann, Glowacki-Dudka, and Conceicao-Runlee (2000), who advocated that cooperative PBL contributes to the formation of online learning communities.

Conclusion and suggestions

To date, no study has clearly identified the important online roles and their corresponding behaviors, nor has a study defined the online learning community types from a holistic perspective. Moreover, objective indices have not been proposed for categorizing online learning communities. This study therefore attempts to pioneer an examination of these areas. To achieve this goal, an 18-week instructional program is employed and the findings are inspiring. The principal findings are as follows.

First, 13 important online behaviors and 3 commonly used online behaviors (constructing a positive atmosphere, providing opinions for group assignments, and providing reminders of assignment-related work) are identified. Second, eight online roles and three common online roles (information providers, opinion providers, and troublemakers) are identified; moreover, the eight roles and their corresponding relationships with online behaviors are elucidated. Third, a two (collaboration vs. participation) by two (high vs. low) model is proposed and four online learning community types (active collaboration, passive collaboration, individualized participation, and indifference) are recognized. These types of online learning communities should be representative, for they carefully evolve from a pilot study, an elaborate instructional design, and, most importantly, specific objective criteria based on online behaviors and online roles. Based on this elaborate evolving process, it is strongly believed that online behaviors, online roles, and online learning communities are closely related.

To conclude, the analytical results of this study are valuable since the instructional design and analyses in this study are deliberately constructed and applied; however, the number of online behaviors and online roles may vary with different discussion content and different participants. Therefore, in addition to replicating the analytical results of this study in a different context, future studies may compare and contrast online behaviors, roles, and communities across various contexts. Additionally, realizing what online learning community type is most beneficial to learners would enhance the effectiveness of online learning. Consequently, further study can verify the relationship between learning effects and the online learning community types identified in this study.

Acknowledgments

The author would like to thank the National Science Council of the Republic of China, Taiwan for financially/partially supporting this research under Contract No. NSC-94-2520-S-004-001.

References

Agre, P. E. (1998). Designing genres for new media: Social, economic, and political contexts. In S. G. Jones (Ed.), *Cybersociety 2.0: Revisiting computer-mediated communication and community* (pp. 69-99). Thousand Oaks, CA: Sage.

Baym, N. K. (1998). The emergence of online community. In S. G. Jones (Ed.), *Cybersociety 2.0: Revisiting computer-mediated communication and community* (pp. 69-99). Thousand Oaks, CA: Sage.

Chang, B., Cheng, N. H., Deng, Y. C. & Chan, T. W. (2007). Environmental design for a structured network learning society. *Computers & Education, 48*, 234-249.

Chang, C. K., Chen, G. D. & Li, L. Y. (2006). Constructing a community of practice to improve coursework activity. *Computers & Education, 50*, 235-247.

Cho, H, Gay, G., Davidson, B, & Ingraffea, A. (2007). Social networks, communication styles, and learning performance in a CSCL community. *Computers & Education, 49*, 309-329.

Collison, G., Elbaum, B., Haavind, S. & Tinker, R. (2000). *Facilitating online learning: Effective strategies for moderators*. Madison, WI: Atwood Publishing.

Corsini, R. (2002). *The dictionary of psychology*. New York, NY: Brunner-Routledge.

Guzdial, M. & Turns, J. (2000). Effective discussion though a computer-mediated anchored forum. *Journal of the Learning Sciences, 9*, 437-469.

Hann, D., Glowacki-Dudka, M. & Conceicao-Runlee, S. (2000). *147 Practical tips for teaching online groups: Essentials of web-based education*. Madison, WI: Atwood Publishing.

Havelock, B. (2004). Online Community and Professional Learning in Education: Research-Based Keys to Sustainability. *AACE Journal, 12*(1), 56-84.

Jones, A. and Issroff, K. (2005). Learning technologies: Affective and social issues in computer-supported collaborative learning. *Computers & Education, 44*, 395-408.

Ligoria, M. B. (2001). Integrating communication formats: Synchronous versus asynchronous and text-based versus visual. *Computers & Education, 37*, 103-125.

Lin, F., Lin, S. & Huang, T. (2007). Knowledge sharing and creation in a teachers' professional virtual community. *Computers & Education, 50*, 742-756.

Ludwig-Hardman, S. & Woolley, S. (2000). Online learning communities: Vehicles for collaboration and learning in online learning environments. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunication, 2000,* 1556-1558.

Moller, L. (1998). Designing communities of learners for synchronous distance education. *Educational Technology Research and Development, 46*(4), 115-122.

Palloff, R. & Pratt, K. (1999). *Building learning communities in cyberspace*. San Francisco, CA: Jossey-Bass Publishers.

Salomon, G. (1993). *Distributed cognitions: Psychological and educational consideration.* Cambridge: Cambridge University Press.

Salomon, G. (1998). Novel constructivist learning environments and novel technologies: some issues to be concerned with. *Research Dialogues in Learning and Instruction, 1*(1), 3-12.

Waltonen-Moore, S., Stuart, D., Newton, E., Oswald, R. & Varonis, E. (2006). From virtual strangers to a cohesive online learning community: The evolution of online group development in a professional development course. *Journal of Technology and Teacher Education, 14*, 287-311.

Wang, M. J. & Poole, M. (2004). Nurturing a dynamic online learning community among teens. *The International Journal of Learning*, 9, 859–870.

Yang, F., Wang, M., Shen, R. & Han, P. (2007). Community-organizing agent: An artificial intelligent system for building learning communities among large numbers of learners. *Computers & Education, 49*, 131–147.

Lee, J., Cerreto, F. A., & Lee, J. (2010). Theory of Planned Behavior and Teachers' Decisions Regarding Use of Educational Technology. *Educational Technology & Society*, *13* (1), 152–164.

Theory of Planned Behavior and Teachers' Decisions Regarding Use of Educational Technology

Jung Lee, Frank A. Cerreto and Jihyun Lee*

The Richard Stockton College of New Jersey, Pomona, NJ, USA // *College of Education, Seoul National University, Seoul, Korea // leej@stockton.edu // Frank.Cerreto@stockton.edu // leeji1@snu.ac.kr

ABSTRACT

According to Ajzen's Theory of Planned Behavior (TPB), behavioral intention (BI) is predicted by attitude toward the behavior (AB), subjective norm (SN), and perceived behavioral control (PBC). Previous studies using the TPB to explain teachers' intentions to use technology have resulted in inconsistent findings. This inconsistency might be due to overly broad definitions of the target behavior. To investigate this potential weakness, we defined a specific target behavior, using computers *only* to create and deliver lessons, and then used the TPB to investigate teachers' decisions. An elicitation study was used to identify teachers' salient beliefs and develop a closed-ended questionnaire. Results of the closed-ended questionnaire revealed that AB, SN, and PBC all were significant predictors of teachers' intentions. However, AB had twice the influence of SN and three times that of PBC. This finding suggests that teachers must have positive attitudes about using computers to create and deliver lessons. They are less concerned about what others think of this practice, and far less bothered by internal or external constraints. Results provide specific information that can be used to design effective teacher development programs and remind TPB researchers of the importance of using specific definitions of the target behavior.

Keywords

Theory of Planned Behavior, Behavioral intention, Technology usage, Teacher beliefs

Problem Statement and Theoretical Foundation

It is generally accepted that the use of technology in schools has altered, and continues to transform the educational landscape dramatically, fueling changes in content, pedagogy, and assessment (US DOE, 2004). In order to capitalize on the potential benefits of technology in the classroom, governments throughout the world have instituted initiatives intended to increase its use (e.g., Rha & Yoshida, 2005; US Web-based Education Commission, 2000).

These initiatives generally recognize the need for effective, continuous teacher development programs designed to help teachers integrate technology into their teaching. However, most recommendations focus mainly on teacher competence with technology. For example, according to the US Web-based Education Commission (2000), teachers must be, "able to apply it [technology] appropriately, and conversant with new technological tools, resources, and approaches (p. 39)."

We will argue that teachers' competence is only one of several factors determining their decisions regarding the use of educational technology. Other influences might include the value they attribute to the use of technology. Regardless of their perceived self competence, teachers may not use technology if they do not value it in their teaching. Another possible influential factor is the opinions of significant others. If, for example, a teachers' supervisor strongly promotes the use of technology in the classroom, this teacher might be inclined to please the supervisor by using technology, despite any perceived personal incompetence or uncertainly of the value.

As a result, designing professional development programs without taking into account other factors limits their potential impact. Moreover, monetary decisions regarding support for technology initiatives must be based on consideration of all factors that determine teachers' decisions to use educational technology. Finally, from a research perspective, it is important to establish the extent to which empirical findings support intuition or conventional wisdom. What then are the primary factors that underlie teachers' intentions to utilize technology in their classrooms, and what are their relative strengths?

Icek Ajzen's (1985) Theory of Planned Behavior (TPB), an explanatory model for a wide variety of behavioral intention, can be used to address this question. According to the TPB, volitional human behavior is immediately preceded by intention to engage in this behavior (see Figure 1). Behavioral intention is predicted, in turn, by three main determinants: attitude toward the behavior (AB), subjective norm (SN), and perceived behavioral control

(PBC). The extent to which individuals view a particular behavior positively (attitude), think that significant others want them to engage in the behavior (subjective norm), and believe that they are able to perform the behavior (perceived behavioral control), serve as direct determinants of the strength of their intention to carry out the behavior.

Each of these three direct determinants of behavioral intention is influenced, in turn, by an indirect determinant. Indirect determinants are based on a set of salient beliefs and evaluations of these beliefs. Measures of the indirect determinants embody expectancy-value theory (Fishbein & Ajzen, 1975). This theory posits that attitudes are developed and revised according to assessments about beliefs and values. This idea was applied to the calculation of the three indirect determinants of the TPB as follows (Ajzen, 1985):

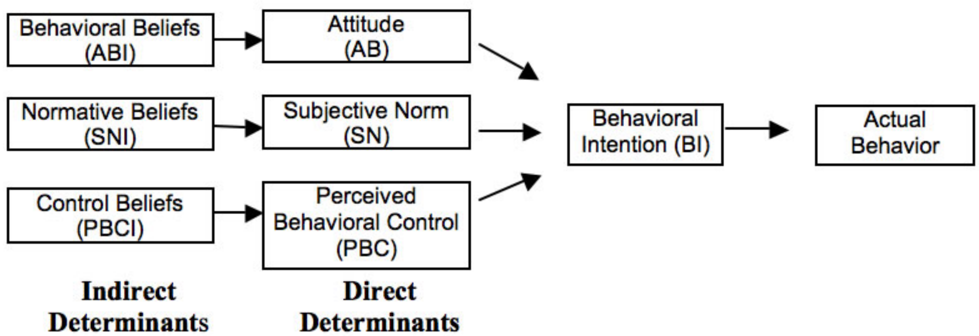

Figure 1. Theory of Planned Behavior (Adapted from Ajzen (1985))

Salient behavioral beliefs (BB) about the outcomes of a particular behavior, weighted by their outcome evaluations (oe), form an indirect measure of an individual's attitude toward the behavior (ABI). Salient normative beliefs (NB) about whether important others approve of the behavior, weighted by the motivation to comply (mc) with these perceived norms, constitute an indirect measure of subjective norm (SNI). Salient control beliefs (CB) about facilitators of or obstacles to performing the behavior, weighted by their control power (cp), comprise an indirect measure of perceived behavioral control (PBCI). The three indirect measures are given by:

$$ABI = \sum_i (oe)_i (BB)_i \qquad SNI = \sum_j (mc)_j (NB)_j \qquad PBCI = \sum_k (cp)_k (CB)_k$$

The TPB has been used successfully to understand a wide variety of human behaviors, such as weight loss behavior (Schifter & Ajzen, 1985), and smoking cessation (Godin, Valois, Lepage, & Desharnais, 1992). Results of these studies have provided specific information used to design effective programs aimed at these behaviors (See Ajzen, n.d., for a comprehensive list of TPB studies). For example, because Godin et al. (1992) found perceived behavioral control to be a relatively strong predictor of both intention and behavior, they recommended programs that help smokers develop their will-power and inform them of the effort required in order to modify smoking behavior.

The TPB has also been used to explain teachers' intentions and behavior in the classroom (e.g., Crawley, 1990; Zint, 2002). In particular, the TPB has been utilized in predicting K-12 teachers' intentions toward educational technology usage (Czerniak, Lumpe, Haney, & Beck, 1999; Salleh & Albion 2004; Sugar, Crawley, & Fine, 2004). Results of such studies have the potential to help guide approaches to fostering teacher technology use.

Besides the TPB, several other models have been used to predict intentions to use technology, including the Theory of Reasoned Actions (Fishbein and Ajzen, 1975) and the Technology Acceptance Model (Davis, Bagozzi, and Warshaw, 1989). The Theory of Reasoned Actions, a precursor to the TPB, includes only two of the three direct determinants of behavioral intention, attitude toward the behavior and subjective norm. The Technology Acceptance Model uses perceived usefulness, instead of subjective norm, as the second determinant of behavioral intention. All three models postulate that behavior is predicted by behavioral intention.

Davis, Bagozzi, and Warshaw (1989) compared the relative effectiveness of the Theory of Reasoned Actions and the Technology Acceptance Model on MBA students' intention to use a word processing program and their subsequent usage and found that, while both models predicted behavioral intention well, perceived usefulness was a relative strong predictor, accounting for more than half of the variance in behavioral intention. Mathieson (1991) compared Technology Acceptance Model and TPB in predicting undergraduate students' intention to use an information system. Both models were found to be effective, and although the Technology Acceptance Model was easier to use, the TPB provided more specific guidance to developers. Because those who are interested in teachers' intentions to use technology are often searching for specific information to guide program development, the TPB is a wise choice.

Unfortunately, attempts to use the TPB to explain teachers' intentions to use technology have resulted in inconsistent findings. For example, Czerniak, Lumpe, Haney, and Beck (1999) concluded that subjective norm and perceived behavioral control were the only two statistically significant predictors of behavioral intention. Sugar, Crawley, and Fine (2004) found that the only significant predictor of behavioral intention was attitude toward the behavior. Salleh and Albion (2004) reported that only attitude and subjective norm were significant predictors of intention.

There are many possible explanations for this inconsistency. Our argument is that these conflicting findings may have resulted from insufficient granularity in the definitions used for the targeted behavior. These three studies applied the TPB to describe teachers' beliefs and intentions regarding the integration of electronic technology in general terms only. They did not take into account the fact that many different technologies exist, and there are many different ways for teachers to utilize a specific technology in the classroom. For example, Czerniak et al. (1999) defined the target behavior as using a wide variety of technologies to foster student learning. Similarly, in their study, Sugar et al. (2004) defined the behavior of interest as, "adopting at least one new technology into a lesson by the end of the next school year (p. 204)." Salleh and Albion (2004) used the general term, Information and Communication Technology, to describe the behavior. These definitions allow for many different technologies, ranging from electronic computers to physical manipulatives, not to mention a broad array of possible uses for each one.

According to Ajzen (2006), when using the TPB, the action comprising the behavior must be defined at an appropriate level of specificity to allow for useful generalization. Ajzen argued, for example, that using a definition such as walking on a treadmill as opposed to exercising would yield more useful results because the reasons individuals decide whether to exercise may depend on the specific form of exercise. Mathieson (1991) also pointed out that the TPB focuses on "specific beliefs that are specific to each situation (p. 178)," providing specific information and insight into an individual's or a group's predispositions.

In particular, teachers' attitude, subjective norm, and perceived behavioral control, and the relative importance of these three factors as predictors of behavioral intention might be very different for different technologies, such as the use of educational sources on the World Wide Web as opposed to the use of online conferencing systems. Relatively vague behavioral definitions may explain discrepancies in findings among previous studies of teacher's intention to utilize technology.

Moreover, from a practical perspective, many, if not most technology initiatives, focus on particular technological solutions to particular educational problems. In practice, teachers do not make global decisions about the place of technology in their classrooms, but rather, they make local decisions about whether or not they will adopt a particular, often emerging technology. Therefore, in order to be fruitful, any study of teachers' intentions, including those that use the TPB, must focus on particular uses of technology.

Purpose and Research Questions

In order to address the aforementioned discrepancy in findings of previous studies, we applied the TPB to investigate teachers' intentions to utilize a specific technology in a specific way. By defining the target behavior at an appropriate level of specificity, we expect to obtain more accurate insight into the factors that influence teachers' intentions to integrate a particular technological approach into their classrooms.

The primary purpose of this study is to use the TPB to examine the underlying beliefs and the relative strengths of the three direct determinants (AB, SN, and PBC) of teachers' intentions to utilize technology in a specific way. For this study, the target behavior is defined as using computers to create and deliver teachers' lessons by using presentation software, such as PowerPoint, during the next month.

In particular, we address the following research questions:

1. Underlying Beliefs: What salient teacher beliefs (behavioral, normative, and control) underlie each of the three indirect determinants of teachers' intention to use computers to create and deliver lessons?
2. Direct and Indirect Determinants of Intention:
    a. Which of the three indirect determinants of intention are statistically significant predictors of their corresponding direct determinants of teachers' intentions to use computers to create and deliver lessons?
    b. To what extent does each statistically significant indirect determinant predict its associated direct determinant of teachers' intention to use computers to create and deliver lessons?
    c. Which of the three direct determinants of intention are statistically significant predictors of teachers' intentions to use computers to create and deliver lessons?
    d. For the statistically significant direct determinants of intention, what are their relative strengths vis-à-vis teachers' intentions to use computers to create and deliver lessons?

Research Method

The Republic of Korea is one of the most technologically advanced societies in the world. According to the 2008 Organization for Economic Cooperation and Development (OECD) report, Korea's broadband subscription rate is 32%, one of the highest rates among countries (OECD, 2008). Also according to the 2003 Program for International Student Assessment (PISA), Korea ranked fourth among 40 nations in student access to computers, averaging fewer than four students per computer (OECD, 2006). With such access, the potential impact of computer technology on education is vast. Consequently, Korea serves as an ideal setting for examining teachers' intentions to utilize technology.

In order to insure construct validity, guidelines for conducting TPB studies given by Ajzen (2006), Ajzen and Fishbein (1980), and Francis et al. (2004) were followed. First, a preliminary, elicitation study was conducted in order to identify participants' salient beliefs regarding the use of presentation software in classroom lessons. The results of the elicitation study were then used to develop measures of behavioral, normative, and control beliefs that were then included, along with direct measures of behavioral intention, attitude, subjective norm, and perceived behavioral control, in the construction of a closed-ended questionnaire.

Because the participants are Korean, all research instruments were administered in Korean. However, following the procedures described in a manual for developing TPB questionnaires (Francis et al., 2004), the documents were first written in English and then translated by two of the authors, who are native Korean speakers. A third native Korean speaker reviewed final drafts. Back-translation validation procedures, as described in Francis et al. (2004), were also used.

Elicitation Study

Participants

The academic preparation of secondary teachers is the same for middle school and high school teachers, but differs significantly from that of elementary school teachers. Because this difference may be related to differences regarding the research questions, and because the TPB is intended for use with homogeneous groups, both the elicitation study and the questionnaire study focused on secondary teachers alone.

The elicitation study was conducted with 34 middle and high school teachers in the Republic of Korea, in March, 2007, in order to identify teachers' relevant salient beliefs. These teachers were purposely selected to represent various subjects, grade levels, teaching experience, and expertise with technology. Table 1 contains participants' demographic information.

Table 1. Elicitation Study Participant Information

Category		Number
Location		
	Seoul	26
	Busan	3
	Deajeon	5
Gender		
	Male	16
	Female	18
Age		
	25-29	6
	30-34	9
	35-39	7
	40-44	6
	45-49	6
School Type		
	High School (HS)	24
	Middle School (MS)	3
	Vocational HS	3
	Science HS	4
Subject		
	Math	7
	Language	5
	Science	6
	Social Science	7
	Korean	4
	Other	5
Teaching Experience		
	0-4	4
	5-9	10
	10-14	6
	15-19	5
	20-14	9
Technology Level		
	Low	4
	Middle	22
	High	8

Procedure

Participants were asked to write answers to open-ended questions regarding their beliefs about the use of presentation software to create and present classroom lessons. In order to educe behavioral beliefs, participants were asked to specify advantages and disadvantages of using presentation software. They were asked to list individuals or groups who would approve or disapprove their use of presentation software in order to provide data on their normative beliefs. Finally, in order to elicit control beliefs, participants were asked to enumerate factors or circumstances that would facilitate or hinder their use of presentation software.

Analysis

In order to answer Research Question 1, two of the authors analyzed the responses independently, grouping similar responses into categories, labeling the categories, and noting their frequencies. All three authors met to finalize labels and reach consensus on discrepant cases. Labels that occurred most often were selected for inclusion in the subsequent closed-ended questionnaire.

Closed-ended Questionnaire Study

Instrument

The closed-ended questionnaire was developed following procedures described in Ajzen (2006) and Francis et al. (2004). In addition to background questions, the questionnaire contained both direct measures of behavioral intention, attitude, subjective norm, and perceived behavioral control, as well as indirect measures (behavioral beliefs, normative beliefs, and control beliefs). Standard scaling procedures were used to construct measures. Except for background questions, all items used a seven-point Likert scale, and items measuring various constructs were interspersed.

As recommended in Ajzen (2006), in order to improve internal consistency of the direct measures, items were constructed with the particular behavior and population in mind. Items on the attitude scale included two types: those that are instrumental (e.g., valuable, beneficial), and those that are experiential (e.g., pleasant, enjoyable). Items on the perceived behavioral control scale embody capability or controllability of performing the behavior.

As mentioned earlier, results of the elicitation study were used to develop each of the three indirect measures. As recommended by Ajzen and Fishbein (1980), those beliefs constituting a majority of the beliefs obtained in the elicitation study were selected for inclusion in the closed-ended questionnaire. Items were constructed for each identified behavioral belief and its outcome evaluation, each normative believe and the motivation to comply with it, and each control belief and its power.

Before administering the closed-ended questionnaire, it was piloted with 20 graduate students in the education department at a large university in Korea, including 10 middle school or high school teachers, and reviewed by two university faculty members. Based on the pilot study and review, minor changes were made to questionnaire instructions and a few items, and subsets of scales that exhibited high internal consistency were selected for the final version.

Reliability of each construct in the final questionnaire was calculated using Cronbach alpha procedures, and all scales were found to have acceptable internal consistency (alpha $\geq$ 0.6) based on guidelines provided by Francis et al. (2004). Table 2 contains reliability data for each of the 10 constructs.

Table 2. Reliability Values from Closed-Ended Questionnaire

Variable	Cronbach's alpha
Behavioral Intention (BI)	0.94
Attitude Toward the Behavior (AB)	0.71
Subjective Norm (SN)	0.73
Perceived Behavioral Control (PBC)	0.73
Behavioral Beliefs (BB)	0.77
Outcome Evaluations (oe)	0.94
Normative Beliefs (NB)	0.79
Motivation to Comply (mc)	0.70
Control Beliefs (CB)	0.60
Control Power (cp)	0.75

Procedure

Stratified sampling was used to select 11 schools for the questionnaire, based on the relative student population size in each of several major geographic regions of Korea, and intended to represent urban, suburban, and rural areas. Based on the number of teachers in each school, a predetermined number of questionnaires was sent via post. At each school, a senior administrator was asked to distribute questionnaires to teachers and then collect and return all completed and blank questionnaires via post. A detailed script and instructions for administering the surveys were provided to the administrators. The script included a description of the purpose of the research study as well as assurances of confidentiality and safety for participants.

Questionnaires were administered to 149 middle school and high school teachers in Korea, in May, 2007. A total of 137 questionnaires were returned, representing a return rate of 91.9 percent. Assuming a moderate effect size for TPB studies (Francis et al., 2004), a sample size of 137 resulted in acceptable statistical power.

Data Analysis

In order to address Research Question 2, a two-stage regression procedure was used (Francis et al., 2004). After item analysis was performed in order to establish internal consistency and appropriate diagnostics for linearity, normality and homoscedasticity were performed, multiple regression was conducted using the direct determinants of attitude, subjective norm, and perceived behavioral control as predictors of intention. Finally, regression was performed with each of the indirect determinants and its associated direct determinant. As described earlier, the sum of the products of each behavioral belief and its outcome evaluation was used as the predictor of attitude toward the behavior, and similarly for subjective norm and perceived behavioral control. SPSS Version 12.0 for Windows was used to compute all statistics for this report.

Results

Underlying Beliefs: Research Question #1

Middle and high school teachers who participated in the elicitation study expressed a variety of behavioral, normative, and control beliefs regarding the use of computers to create and deliver lessons. Table 3 contains a summary of the most commonly held beliefs.

Teachers' behavioral beliefs regarding the use of computers to create and deliver lessons gravitate toward two areas: better teaching, and improved student behaviors. Reactions were generally positive, with perceived advantages of using computers outweighing disadvantages. Following are a few samples of specific teacher comments about attitudes and, in parentheses, the categories to which they were assigned in Table 3:

> "Because of reducing writing time on the blackboard, I can give more detailed explanations." (quality of teaching)
>
> "When presenting computer graphics, it could deliver wrong information. For example, in the case of y=x, if the graphic angle is not exactly 45 degree on PPT, students may get the wrong information." (student achievement)
>
> "Visualized computer-based presentations make students pay attention to teacher's instruction." (student attention)

The most important others whose opinions teachers consider include school administrators, students and their parents. Teachers who cited school administrators generally reported that these administrators support the use of computers to create and deliver lessons. However, those who cited students or their parents had conflicted views reporting that, for example, students who are interested in computers encourage their use to create and deliver lessons, but those who are not discourage it.

In order to use computers effectively to create and deliver lessons, teachers cite the importance of both external (reliable hardware and software) and internal (skills) factors. For both types of factors, some teachers reported inhibitory effects, whereas in others reported enabling effects.

Table 3. Summary of Salient Beliefs from Elicitation Study

Behavioral Beliefs	Normative Beliefs	Control Beliefs
• quality of teaching • student achievement • student attention	• school administrators • students • students' parents	• reliable hardware and software • skills • training and support • time to create

Direct and Indirect Determinants of Intention: Research Question #2

Summary statistics for the seven main constructs measured in the closed-ended questionnaire are provided in Table 4. For each of the four direct measures (BI, AB, SN, PBC), the theoretical minimum and maximum scores are -21 and 21, respectively. For the each of the three indirect measures (ABI, SNI, PBCI), the theoretical minimum and maximum scores are -63 to 63, respectively. All intercorrelations are considerably less than 1, indicating that discriminant validity was achieved. We also considered the assumption for TPB models that predictive factors in the model are correlated. As shown in Table 4, pairwise correlations among all predictors in the model were statistically significant at the 0.05 level.

Table 4. Descriptive Statistics and Correlations (N= 137)

	ABI	AB	SNI	SN	PBCI	PBC	BI	Mean	SD
Indirect Attitude toward the Behavior (ABI)	--							27.33	15.735
Direct Attitude toward the Behavior (AB)	-.534**	--						14.56	3.410
Indirect Subjective Norm (SNI)	-.360*	.443**	--					-1.12	17.852
Direct Subjective Norm (SN)	-.240**	.657**	.662**	--				12.68	3.306
Indirect Perceived Behavioral Control (PBCI)	.474**	-.708**	-.330**	-.612**	--			19.33	18.454
Direct Perceived Behavioral Control (PBC)	-.402**	.583**	.302**	.558**	-.687**	--		15.17	3.392
Behavioral Intention (BI)	-.434**	.799**	.479**	.661**	-.713**	.604**	--	13.21	4.752

Note. **p* < .05, ** *p* < .01, ****p* < .001.

All direct measures in the model are positive. Teachers in the study intended to use computers to create and deliver lessons, they had positive attitudes about its use, significant others endorse this use, and teachers expressed confidence they had the necessary capability.

The situation with regard to indirect measures is less consistent. Teachers in the study exhibited positive beliefs about the outcomes associated with using computers to create and deliver lessons. In addition, they expressed the belief that they possessed the internal and external resources needed to do so. However, teachers reported that opinions of others regarding their use of computers were neutral, overall.

Data diagnostics were conducted in order to ascertain whether assumptions underlying the validity of conclusions based on the regression analysis were met. A preliminary examination of histograms and normality plots suggested that all seven variables were normally distributed. Subsequent analyses were conducted using the Kolmogorov-Smirnov test, with the Lilliefors correction (Lilliefors, 1967) and the Jarque-Bera test (Jarque and Bera, 1987). The results of these tests confirmed that none of the variables differs from normality at the 0.05 significance level (Table 5).

Table 5. Normality Tests

Variables	Kolmogorov-Smirnov Statistic	Jarque-Bera Statistic
ATTI	0.074	1.283
SNI	0.074	1.171
PBCI	0.053	1.720
ATTD	0.076	0.835
SND	0.075	1.031
PBCD	0.072	1.676
BI	0.076	4.262

An examination of scatter plots provided strong evidence of linearity and multivariate normality. Ramsey's RESET test (Ramsey, 1969) provided formal support for the assumption of linearity and the specification of the models with all results failing to reject the null hypothesis at the 0.05 significance level (Tables 6-9). A scatterplot of the standardized residuals versus the predicted values from the regression analysis confirmed the assumption of homogeneity of variance-covariance. Formal tests of heteroscedasticity, using White's test (White, 1980) with the number of predictors as the degrees of freedom, were conducted. All results failed to reject the null hypothesis at the 0.05 significance level, and, therefore, supported the assumption of homoscedasticity for all regressions (Tables 6-9).

Tables 6-8 contain the results when each direct determinant was regressed on its indirect counterpart. The indirect determinant of attitude toward the behavior (ABI) was a significant predictor of the direct determinant (AB), $F(1,135) = 48.610$, $p < 0.001$, and accounted for 26.5 percent of its variance. The indirect determinant of subjective norm (SNI) had a significant influence on the direct determinant (SN), $F(1, 135) = 113.017$, $p < 0.001$, and accounted for 45.6 percent of its variance. The indirect determinant of perceived behavioral control (PBCI) was a significant predictor of the direct determinant (PBC), $F(1,135) = 114.281$, $p < 0.001$, and accounted for 45.8 percent of its variance.
Regarding Research Question 2a and b, each of the three indirect determinants of the theory constructs was significantly and strongly related to its corresponding direct determinant, further supporting the model, and providing additional support for the measures' validity.

Table 6. Regression Analysis: Predicting Attitude Toward the Behavior, AB (N = 137)

	R^2	S.E.	F	B	S. E. B	β	White's Statistic	Ramsey's RESET Statistic
Indirect Attitude toward the Behavior (ABI)	0.265	2.935	48.610***	0.112	0.16	0.515	1.781	1.586

Note. *$p < .05$, ** $p < .01$, ***$p < .001$.

Table 7. Regression Analysis: Predicting Subjective Norm, SN (N = 137)

	R^2	S.E.	F	B	S. E. B	β	White's Statistic	Ramsey's RESET Statistic
Indirect Subjective Norm (SNI)	0.456	2.448	113.017***	0.125	0.012	0.675	3.699	1.724

Note. *$p < .05$, ** $p < .01$, ***$p < .001$.

Table 8. Regression Analysis: Predicting Perceived Behavioral Control, PBC (N = 137)

	R^2	S.E.	F	B	S. E. B	β	White's Statistic	Ramsey's RESET Statistic
Indirect Perceived Behavioral Control (PBCI)	0.458	2.505	114.281***	0.124	0.012	0.677	5.617	0.123

Note. *$p < .05$, ** $p < .01$, ***$p < .001$.

Table 9. Regression Analysis: Predicting Behavioral Intention (N=137)

	R^2	S.E.	F	B	S. E. B	β	White's Statistic	Ramsey's RESET Statistic
Model	0.700	2.630	103.644***				16.577	0.211
Attitude toward the Behavior (AB)				0.793	0.094	0.569		
Subjective Norm (SN)				.0329	0.096	0.229		
Perceived Behavioral Control (PBC)				0.201	0.084	0.144		

Note. *$p < .05$, ** $p < .01$, ***$p < .001$.

Table 9 contains the results of behavioral intention regressed on the three direct predictors in the model, which show that at least one of the direct determinants influenced behavioral intention, $F(3,133) = 103.644$, $p < 0.001$. In response to Research Question 2c, the analysis revealed that all three direct determinants – attitude toward the behavior, $t(133) = 8.481$, $p < 0.001$, subjective norm, $t(133) = 3.446$, $p = 0.001$, and perceived behavioral control, $t(133) = 2.386$, $p < 0.05$ – were statistically significant predictors of teachers' intentions to use computers to create and deliver lessons. Together, the three determinants accounted for 70 percent of the variance in teachers' intentions. This finding is in contrast to those of Czerniak et al. (1999), Sugar et al. (2004), and Salleh and Albion (2004). In none of those studies were all three direct determinants found to be significant predictors of behavioral intention.

Figure 2 represents the pathways, including beta values, found in the regression analysis. The betas from regression model were used to determine the relative weights of each factor.

Regarding Research Question 2d, of the three direct determinants, attitude toward the behavior had the most substantial impact ($\beta = 0.569$) on teachers' intentions to use computers to create and deliver lessons, producing a change of 0.569 units in behavioral intention for each unit change in attitude. This influence on intention is more than twice that of subjective norm ($\beta = 0.229$) and more than three times that of perceived behavioral control ($\beta = 0.144$). This finding suggests that teachers' decisions about using computers to create and deliver lessons are influenced strongly by their view of its value, moderately by the opinions of significant others, and weakly by teachers' perceived ability to do so.

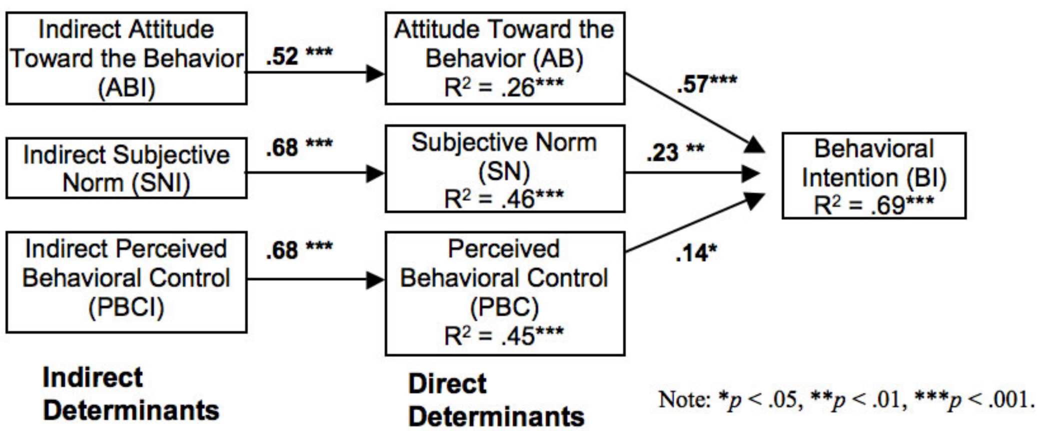

Figure 2. Path diagram of TPB model of teachers' use of computers to create and deliver lessons

Discussion

Regarding Research Question 1, underlying salient beliefs, it is interesting to compare and contrast the results of this study and those of Czerniak et al. (1999), Sugar et al. (2004), and Salleh and Albion (2004). All of the modal salient beliefs in attitude, subjective norm, and perceived behavioral control identified in this study were also present in Czerniak et al. (1999)and all but one in Sugar et al. (2004) (Salleh and Albion (2004) did not report salient beliefs.). However, the first two studies identified several additional common beliefs. For example, under attitudes, both Czerniak et al. (1999) and Sugar et al. (2004) reported preparing students for the future and helping them to acquire new skills as common responses. These additional salient beliefs are undoubtedly an artifact of the relatively general definition of the target behavior used in the other studies. Given that the current study focused specifically on teachers' use of computers to create and deliver lessons, it is not surprising that these other beliefs were not observed.

Regarding Research Question 2, indirect and direct predictors of intention, recall the disagreement as to which factors serve as significant predictors of teachers' intention to use technology. Czerniak et al. (1999) found that SN and PBC were the only two significant predictors of teachers' intentions to use technology. Sugar et al. (2004) identified AB as the only significant predictor of this behavior. Salleh and Albion (2004) found that only AB and SN were significant predictors of the behavior. The main finding of this study, embodied in Research Question 2c, provides a possible resolution to this paradox, by demonstrating that AB, SN, and PBC all served as significant antecedents to teachers' intentions to use computers to create and deliver lessons. This finding demonstrated our claim that providing a clear and specific definition of the target behavior, as opposed to general definitions, could lead to more meaningful conclusions that are consistent with the TPB.

Once the significance of all three direct determinants was established, we were able to examine their relative strengths. In response to Research Question 2d, AB had more than twice the influence of SN and more than three times the influence of PBC on teachers' intentions to use computers to create and deliver lessons. This finding suggests that teachers must believe positive educational outcomes will follow in order for them to intend to use computers to create and deliver lessons. They are less concerned about what others think of this practice, and far less bothered by any internal or external constraints that may exist.

Conclusion

This study has both theoretical and practical importance. With regard to the TPB, we refined the application of a widely used social psychological theory by reemphasizing the importance of providing specific definitions of the target behavior.

Findings provide practical information to two groups of individuals interested in the effective integration of computer technology in the classrooms. First, these findings give specific guidance to individuals who design and implement technology initiatives. In particular, findings from the elicitation study lead to specific recommendations for developers of teacher development programs. For example, because several teachers expressed concern that using computers to create and deliver lessons requires too much time, designers of teacher development programs should emphasize methods to improve efficiency.

Second, the findings will aid decision makers in determining where resources should be targeted in order to optimize their allocation. Attitude toward the behavior was found to have much greater influence on teachers' intentions to use computers to create and deliver lessons than either subjective norm or perceived behavioral control. According to these findings, teachers base their decisions primarily on their evaluation of the potential benefits, with less regard for the opinions of others and little concern over internal and external resources. Therefore, resources directed toward teacher development programs should be allocated accordingly.

The concern for internal consistency for direct measures discussed earlier does not apply to indirect measures in the TPB because individuals can (and often do) hold both positive and negative beliefs about any particular behavior. Therefore, alternate measures of reliability, such as test-retest studies, are recommended for indirect measures (Ajzen, 2006; Francis et al., 2004). Unfortunately, participant access constraints did not allow for completion of a test-retest study of indirect (belief-based) measures in this study. As an alternative, reliability analysis was conducted, with satisfactory Cronbach alpha values obtained for all three indirect measures, as was the case with the direct measures. Although this approach provides strong evidence of reliability for the direct measures, it gives weaker reliability support for indirect measures and thus may limit this study's conclusions.

Our primary interest in this study is in the direct and indirect factors determining teachers' intentions to utilize technology. The ultimate goal of many, of course, is that teachers will actually use technology effectively in their classrooms. However, for reasons outlined earlier, unless teachers "buy into" the idea, efforts to bring technology into the schools will have limited effectiveness. Our decision to focus on behavioral intention and not the behavior itself rests on solid theoretical and empirical ground. Like many other models of behavior, The TPB postulates that behavioral intention is the immediate antecedent of volitional behavior. Empirical studies have validated the strength of this intention-behavior link in the TPB model (e.g., Ajzen & Madden, 1986) as well as in other models of behavior (Davis et al., 1989; Sheppard, Hartwick, & Warshaw, 1988).

This study demonstrated that precise definitions must be used in order to determine the predictors of teachers' intentions to use technology in specific ways. A logical next step would be to replicate study with other specific uses of technology in order to ascertain what differences exist among them when the TPB is used as an explanatory model. The authors have completed a study comparing and contrasting the results among three different uses of technology that reveals significant differences with respect to significance and influence of the three direct determinants of teachers' intentions to use different forms of educational technology (Lee, Cerreto, & Lee, in press).

Finally, in order to establish the generalizabilty of the results, the study should be replicated in other geographic locations and with elementary school teachers. Findings of these follow-up studies would help us to identify which findings can be applied to which populations.

References

Ajzen, I. (1985). From intentions to action: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.) *Action-control: From cognition to behavior*, Heidelberg: Springer, 11-39.

Ajzen, I. (2006). Constructing a TpB questionnaire: Conceptual and methodological considerations, Retrieved March 22, 2007, from http://www.people.mass.edu/aizen/tpb.html.

Ajzen, I. (n.d.). *The Theory of Planned Behavior: A Bibliography*. Retrieved June 23, 2007, from http://people.umass.edu/aizen/tpbrefs.html.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Ajzen, I & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions and perceived behavioral control. *Journal of Experimental Social Psychology, 22*(5), 453-474.

Crawley, F. E. (1990). Intentions of science teachers to use investigative teaching methods: A test of the theory of planned behavior. *Journal of Research in Science Teaching. 27*(7), 623-716.

Czerniak, C.M., Lumpe, A. T., Haney, J.J., & Beck, J. (1999). Teachers' beliefs about using educational technology in the science classroom. *International Journal of Educational Technology, 1*(2), Retrieved February 12, 2007, from http://www.outreach.uiuc.edu/ijet.

Davis, F., Bagozzi, P., & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982-1003.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Francis, J. J., Eccles, M. P., Johnston, M., Walker, A., Grimshaw, J., Foy, R., et al. (2004). Constructing questionnaires based on the theory of planned behavior: A manual for health services researchers. Centre for Health Services Research: Newcastle Upon Tyne, United Kingdom. Retrieved January 25, 2007, from http://www.rebeqi.org/ViewFile.aspx?itemID=212

Godin G., Valois P., Lepage L., & Desharnais R.(1992). Predictors of smoking behaviour: an application of Ajzen's theory of planned behaviour. *British Journal of Addiction, 87*(9), 1335-1343.

Jarque, C. M., and Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review, 55*(2), 163-172.

Lee, J., Cerreto, F. A., & Lee, J (in press). Teachers' intentions toward technology usage: Do different uses lead to different determinants? In C. Maddux (Ed.), *Research Highlights in Information Technology and Teacher Education 2009.* Chesapeake, VA: Society for Information Technology and Teacher Education

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov tests for normality with mean and variance unknown. *Journal of the American Statistical Association, 62*(318), 399-402.

Mathieson, K. (1991). Predicting user intentions: Comparing the technology acceptance model with the theory of planned behavior. *Information Systems Research, 2*(3), 173-191.

OECD (2006). Are students ready for a technology-rich world?: What PISA studies tell us. OECD, Paris. Retrieved June 1, 2009, from http://www.oecd.org/dataoecd/28/4/35995145.pdf

OECD (2009). Broadband subscribers per 100 inhabitants (Dec. 2008). Retrieved June 1, 2009, from http://www.oecd.org/dataoecd/21/35/39574709.xls

Ramsey, J. B. (1969). Test for specification error in classical linear least squares regression analysis. *Journal of the Royal Statistical Society, Series B (Methodological), 31*(2), 350-371.

Rha, I. & Yoshida, A. (2005). A comparative study on ICT policy in education in Korea and Japan. *Educational Technology International, 6*(1), 3-39.

Salleh, S. & Albion, P. (2004). Using the theory of planned behaviour to predict Bruneian teachers' intentions to use ICT in teaching. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2004* (pp. 1389-1396). Chesapeake, VA: Association for the Advancement of Computing in Education.

Schifter, D. E. & Ajzen, I. (1985). Intention, perceived control, and weight loss: An application of the theory of planned behavior. *Journal of Personality and Social Psychology*, *49*(3), 843-851.

Sheppard, B. H., Hartwick J., & Warshaw, P.R. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research. 15*(3), 325-343.

Sugar, W., Crawley, F., & Fine, B. (2004). Examining teachers' decisions to adopt new technology. *Educational Technology and Society*, *7*(4), 201-213.

US Department of Education (2004). *Toward a new golden age in American education--How the Internet, the law and today's students are revolutionizing expectations*. Retrieved September 21, 2007, from http://www.ed.gov/about/offices/list/os/technology/plan/2004/index.html.

US Web-based Education Commission (2000). *The power of the Internet for learning: Final report of Web-Based Education Commission*. Retrieved April 1, 2007, from http://www.ed.gov/offices/AC/WBEC/FinalReport/index.html.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838.

Zint, M. (2002). Comparing three attitude-behavior theories for predicting science teachers' intentions. *Journal of Research in Science Teaching, 39*(9), 819-844.

Almekhlafi, A. G., & Almeqdadi, F. A. (2010). Teachers' Perceptions of Technology Integration in the United Arab Emirates School Classrooms. *Educational Technology & Society, 13* (1), 165–175.

Teachers' Perceptions of Technology Integration in the United Arab Emirates School Classrooms

Abdurrahman Ghaleb Almekhlafi and Farouq Ahmad Almeqdadi[*]
College of Education, P.O Box 17551, Al-Ain, UAE // almekhlafi@uaeu.ac.ae
[*]Emirates College for Advanced Education, P. O. Box 126662 Abu Dhabi, UAE // falmeqdadi@ecae.ac.ae

ABSTRACT

Technology is a growing part of any society today. Educational technology has become a cornerstone for any country's efforts to improve students' performance at K-12 schools. It has become the focus of educators worldwide. However, research studies investigating technology integration, particularly at the United Arab Emirates (UAE) K-12 schools, focus on quantitative data collection methodology. This study investigated technology integration at UAE Model schools using a mixed method of data collection consisting of focus group interviews and a questionnaire. Study sample consisted of 40 female and 60 male teachers from two schools in Al-Ain Educational Zone, Abu Dhabi. Study results showed that teachers at both schools are integrating technology in their classes' activities. They use a variety of technologies to promote students' learning. However, methods of integration by male teachers differed in some cases compared to their female colleagues. Implications for technology integration in the UAE context are discussed.

Keywords

Technology Integration, Teachers' Perceptions, UAE Schools

Introduction

Technology integration in the classroom has become an important aspect of successful teaching. It has triggered many researchers to investigate different aspects of such integration (e.g., Kotrlik & Redmann, 2005; Bauer and Kenton, 2005; Judson, 2006; Totter et al., 2006; ChanLin et al., 2006; Zhao, 2007; Gulbahar, 2007; Anderson and Maninger, 2007; Abbit and Klett, 2007; & Wood and Ashfield, 2008). This is because it allows students to learn more in less time and allows schools to focus on global learning environments if used appropriately. In addition, it could be an effective teaching tool when used to engage all students in the learning process (Almekhlafi, 2006a, 2006b).

Research shows that there are increasing number of computers being used at home and an increasing number of technological devices available to schools (Goddard, 2002). Research documented teachers' use of computers for different purposes and objectives (e.g., Guha, 2000; Yildirim, 2000; & Rowand, 2000). Some teachers use computers for instructional purposes while others use them for both personal and instructional goals. This study investigates teachers' perceptions of tilizing of computers and other technologies for teaching and learning.

Literature Review

Technology use in education is becoming an increasingly important part of higher and professional education (Wernet, Olliges, & Delicath, 2000; & Almekhlafi, 2006a, 2006b). Technology not only gives learners the opportunity to control their own learning process, but also provides them with ready access to a vast amount of information over which the teacher has no control (Lam & Lawrence, 2002).

According to Rowand (2000), a survey based on a National Center for Education Statistics (NCES, 2000), found that 39% of teachers indicated that they used computers or the Internet to create instructional materials, 34% for administrative record keeping, less than 10% reported to access model lesson plans or to access research and best practices. Novice teachers were more likely to use computers or the Internet. Similarly and according to a report released by the U. S. Department of Education, NCES (2000), novice teachers were more likely to use computers or the Internet to accomplish various teaching objectives. Teachers with at most nine years of teaching experience were more likely compared teachers with 20 or more years of experience to report using computers or the Internet to communicate with colleagues.

Because technology integration is a very broad concept and has several aspects and implications, researchers categorized the previous studies into four different categories:

(1) Technology Integration and its Impact on Students and Teachers

A number of researchers have explored technology integration projects worldwide and reported positive impact on teaching and learning for teachers using technology (e.g., Holinga, 1999; Guha, 2000; Sandholtz, 2001; Manzo, 2001; Sherry et al., 2001; Hong and Koh, 2002; Zorfass and Rivero, 2005, & Almekhlafi, 2006a, 2006b). For example, Guha (2000) reported significant differences and positive correlations between teachers' present computer training, level of comfort, and computer usage in the classroom as compared to their previous training, comfort level, and usage.

Manzo's (2001) study found that many of the students who are drawn to Electronic Arts Class were struggling in most of their other classes. Once they saw what they could do with technology, they began to appreciate the importance of doing well in all subjects. Similarly, Sherry et al. (2001) studied the WEB Project. Their findings of a survey assessing the grant's impact on student achievement suggest that teachers should emphasize the use of meta-cognitive skills, application of skills, and inquiry of learning as they infuse technology into their academic content areas.

(2) Factors Influencing Teachers' Technology Integration in the Classroom

Technology integration at schools and factors affecting such integration has drawn the attention of many researchers and has been of high interest to them. A number of studies and projects have been conducted to explore teachers' use of technology and factors hindering such use (e.g., Becker and Ravitz, 2001; Redmann and Kotrlik, 2004; Kotrlik and Redmann, 2005; Bauer and Kenton, 2005; Judson, 2006; Totter et al., 2006; ChanLin et al., 2006; Zhao, 2007; Gulbahar, 2007; & Anderson and Maninger, 2007).

Bauer and Kenton (2005) found that teachers, who were highly educated and skilled with technology, were innovative and adept at overcoming obstacles, but they did not integrate technology on a consistent basis both as a teaching and learning tool. Results suggest that schools have not yet achieved true technology integration. Gulbahar (2007) concluded that teachers and administrative staff felt themselves competent in using ICT available at the school; they reported a lack of guidelines that would lead them to successful integration. On the other hand, students reported that ICT is not utilized sufficiently in their classes.

Zhao (2007) conducted a qualitative research to investigate the perspectives and experiences of 17 social studies teachers following technology integration training. The research indicated that teachers held a variety of views towards technology integration. These views influenced their use of technology in the classroom. Most teachers were willing to use technology, expressed positive experiences with technology integration training, increased their use of technology in the classroom, and used technology more creatively.

On the other hand, numerous studies have been carried out to identify factors facilitating or prohibiting technology integration in the classroom, particularly computers. Some studies focus on the availability of computers in the classroom, sharing of resources, a supportive administration, and a strong support staff as the primary influencing factors. As an example, the Becker and Ravitz (2001) study showed that computer use among teachers is related to more constructivist views and practices and to changes in practice in a more constructivist-compatible direction. In addition, other research studies suggest that there is a relationship between a teacher's student-centered beliefs about instruction and the nature of teacher's technology-integrated experiences (Judson, 2006; & Totter et al., 2006).

Similarly, ChanLin et al. (2006) conducted a study to identify the factors affecting eight teachers' use of technology in creative teaching practicies. The identified factors were classified into four categories: environmental, personal, social and curricular issues. Besides Chanlin's study, Anderson and Maninger (2007) investigated the changes in and factors related to students' technology-related abilities, beliefs, and intentions. Statistically significant changes were found in students' perceived abilities, self-efficacy beliefs, value beliefs, and intentions to use software in their future classrooms. Students' self-efficacy, value beliefs, and intentions were moderately correlated with each other. Abilities were correlated with self-efficacy and computer access. The best predictors of intentions were self-efficacy beliefs, gender, and value beliefs.

(3) Teachers' Perceptions of Technology Integration and Gender Differences

Teachers' perspectives of their use of instructional technology, understanding of this technology, and feelings about the support structure associated with this equipment have been examined with the findings suggesting that teachers believe technology is an integral part of the process of educating their students. Pertaining to gender differences in technology integration, the literature showed that there were some differences between male and female teachers in technology use, while other studies did not (e.g., Shashaani, 1997; Bhargava et al., 1999; and Hong & Koh, 2002).

The results of Shashaani's study (1997) showed that female students were less interested in computers and less confident than male students. The results also showed that males were more experienced than females and females' attitudes improved after taking the course. Bhargava et al. (1999) studied gender discrepancy in both classroom access and use. The findings showed that there were significant differences between males and females and these differences were due to biased classroom practices, lack of female role models, and home computer gender gaps.
Following the same path, Hong and Koh (2002) found that female teachers were more anxious than male teachers toward hardware. They also found that the overall computer anxiety levels of male teachers were not significantly different from the anxiety levels of female teachers. Only for the hardware anxiety domain was significant differences detected between male and female teachers.

(4) Technology Integration Barriers

A number of barriers that hinder technology integration have been documented (Flores, 2002; Earle, 2002; & Brinkerhof, 2006). According to Flores (2002), teachers face many barriers in their quest to incorporate technology. In addition to time scheduling for technology use and administrative support, equity is another important issue. The introduction of technology is particularly hard when there are few resources.

Earle (2002) pointed out some barriers to the integration of technology in the classroom including both restraining forces that are extrinsic to teachers such as access, time, support, resources, and training and forces that are intrinsic such as attitudes, beliefs, practices, and resistance. More recently, Brinkerhof (2006) pointed out that barriers are grouped into four main categories: resources, institutional and administrative support, training and experience, and attitudinal or personality factors.

Statement of the problem

Due to the role of technology in the advancement of society in general and educational sector in particular, effective technology integration into teaching and learning has become the focus of many educators. However, most research studies conducted so far focus on quantitative data collection methodology such as surveys. This method of data collection does not always give a true picture of technology integration in the classroom. This is particularly true if teachers and students were not voluntarily participating in the study. In such a case, they may fill in questionnaires without giving enough thought to content. Hence, study results are affected and do not reflect reality. Therefore, the need to investigate technology integration using a mixed-methodology is a must. This study aims at investigating technology integration at UAE K-12 schools using a mixed-method for data collection. Such research in the United Arab Emirates has not been conducted as literature search did not result in any studies. This study aims at investigating teachers' perceptions of their technology integration competencies, barriers obstructing such integration, and incentives to increase it, in addition to other related issues.

Methodology

Participants

The participants were 100 (Grades 6-9)) teachers from two model schools in Al-Ain educational zone, Abu Dhabi, United Arab Emirates. Forty of the participants were female, while the rest were male teachers. All teachers at both schools had between 5 and 15 years of teaching experience. All had experience using technology in their classes as it is mandated by model schools. Both schools have good technology infrastructures available for teachers.

Research Questions

This study aimed at answering the following questions
1. How do teachers perceive their competencies to technology integration?
2. How do teachers perceive obstacles and incentives related to successful classroom technology integration?
3. How do teachers perceive their students' classroom usage of technology?
4. To what extent do teachers perceive their classroom use of technology tools?
5. What is the difference in perception of male and female teachers in technology integration?

Data Collection

To answer these questions, the study used multiple research tools, a questionnaire and focus group interviews. The aim of these tools is to investigate teachers' perceptions of technology integration and actual classroom practices. In addition, the use of these tools enables researchers to validate study results, and hence get more reliable findings.

a) A questionnaire focusing on teachers' perception of technology integration was developed. It consisted of a number of subthemes that investigated teachers' perceptions of their technology competencies and usage, students' usage of technology, problems hindering technology integration, and incentives that motivate teachers to integrate technology. The face validity of the questionnaire was established by refereeing it by a panel of university professors with different specializations, including educational technology. The questionnaire validity using Cronbach's Alpha was 0.94. The questionnaire used a five-point Likert scale extending from 5 (very high or strongly agree) to 1 (very low or strongly disagree). The questionnaire was distributed to all teachers at participating schools. Response rate was around 75%.
b) Focus group interviews were conducted with the teachers at both schools. The aim of these interviews was to collect detailed data on technology integration methods, problems hindering such integration, and incentives that increase this integration in the class. Two focus group interviews were conducted with about 20 teachers from the two schools representing different subjects, namely Islamic and Arabic Studies, Social Studies, Science, Math, and English.

Data Analysis

Data gathered from questionnaire items were analyzed using SPSS 15.0. Descriptive statistics, a multivariate analysis, and analysis of variance (ANOVA) were used. In addition, the researchers analyzed these items using "Item Analysis" method in order to get a deep understanding of the results from the questionnaire. On the other hand, data collected from focus group interviews were analyzed using the phenomenographic approach to data analysis, which classified expressions used by participants according to similarities and differences (Levin and Wadmany, 2006).

Results and Discussion

To answer question 1 "How do teachers perceive their competencies to technology integration?", results indicated that teachers highly regard their competencies in technology integration. The mean scores ranged from 4.0 to 4.8 on a 5-point scale (see Table 1). This high perception by teachers might be due to the fact that technology integration in classrooms is a part of teacher evaluation, particularly at model schools. Investigating the items in details, the highest mean scores were for items that are related to teachers' ability to use hardware and software, using technology to locate, evaluate, and collect information from a variety of sources, and content-specific tools.

These results conform to Bauer and Kenton (2005), where they found that teachers were highly skilled with technology and had the competencies required from successful technology integration. In addition, they were also supported by Zhao (2007) who investigated the perspectives and experiences of 17 social studies teachers following technology integration training. Four major categories of technology-related activities were observed among participants: (a) teacher-centered, (b) structured inquiry, (c) teacher-student negotiated, and (d) student-centered. Most teachers were willing to use technology, expressed positive experiences with technology integration training, increased their use of technology in the classroom, and used technology more creatively.

Enforcing these results, focus group interviews yielded some recommendations by teachers in order to enhance their technology competencies and hence result in successful and effective technology integration in the classroom. Male teachers recommend the following: (1) using computer labs as they give teachers the freedom and flexibility to prepare class materials required for the whole course, (2) providing teachers with appropriate professional development in the form of workshops on technology integration, (3) matching technology with curriculum goals so that technology integration enhances teaching and learning, and (4) giving enough freedom for teachers in the coverage and selection of materials to focus on quality rather than quantity.

ChanLin et al. (2006) supported the above findings, where teachers' perceptions about technology use were studied in order to identify the factors affecting their use of technology in their teaching. Two major issues were explored. First, the researchers studied how teachers integrated technology into creative teaching; they then identified the factors that influenced teachers' use of technology in teaching. The identified factors were classified into four categories: environmental, personal, social and curricular issues.

Table 1: Teachers' Perceptions of their Competencies to Technology Integration

	M	SD
I am proficient in the use of common input and output devices; I can solve routine hardware and software problems; I can make informed choices about technology systems, resources, and services.	4.8	0.4
I can use technology to locate, evaluate, and collect information from a variety of sources.	4.6	0.5
I can use technology tools and information resources to increase productivity, promote creativity, and facilitate academic learning.	4.5	0.6
I can use content-specific tools (e.g., software, simulation, environmental probes, graphing calculators, exploratory environments, Web tools) to support learning and research.	4.5	0.6
I can collaborate in constructing technology-enhanced models, preparing publications, and producing other creative works using productivity tools.	4.5	0.6
I can use technology tools to process data and report results.	4.4	0.6
I have a strong understanding of the nature and operation of technology systems.	4.3	0.6
I understand the legal, ethical, cultural, and societal issues related to technology.	4.2	0.7
I can choose learning and technology resources.	4.1	0.7
I can use technology resources to facilitate higher order and complex thinking skills, including problem solving, critical thinking, informed decision-making, knowledge construction, and creativity.	4.1	0.8
I can troubleshoot common computer problems.	4.0	0.9
I can use technology in the development of strategies for solving problems in the real world.	4.0	0.7
I have knowledge to discuss health and ethical issues related to technology.	4.0	0.8
I can use technology tools and resources for managing and communicating information (e.g., finances, schedules, addresses, purchases, correspondence).	4.0	0.8
I can evaluate and select new information resources and technological innovations based on their appropriateness to specific tasks	4.0	0.7
I can use a variety of media and formats, including telecommunications, to collaborate, publish, and interact with peers, experts, and other audiences.	4.0	0.8
I can discuss diversity issues related to electronic media.	4.0	0.8

To answer question 2 "How do teachers perceive obstacles and incentives related to successful technology integration in the classroom?, results showed that teachers perceive time and curriculum as two major obstacles that hinder their technology integration in their classrooms (see Table 2).

Deep analysis of focus group interviews showed other barriers that hinder technology integration. Male teachers indicated that there is a lack of training on how to integrate technology effectively. Most teachers depend on self-learning. They need to be involved in subjects that enable them to learn technology integration techniques and strategies so they can use it successfully in their classes. Another barrier is parents' and teachers' negative attitudes toward the importance and benefits of technology for learning and teaching.

Female teachers pointed out that a large number of students, technical problems, and expensive tools are the common problems that negatively affect the effectiveness of technology. They suggested that schools should provide teachers with affordable and/or free professional development subjects. Furthermore, there should be collaboration between schools where teachers can exchange ideas and successful technology integration techniques. Finally, they suggested that the curriculum should be accompanied with technology-enhanced materials such as CDs and videos. These findings conform to Shelly et al. (2002).

Pertaining to incentives (see Table 2), results showed that participants consider having a free or discounted computer as a major incentive for them. This might be because they need computers at home to enable them to work on technology integration activities at their own pace and time. Other incentives that had high mean scores are participation in workshops, having additional resources, positive evaluations, and recognition by school or school zone. In fact, professional development for teachers and having enough technology resources are crucial for successful technology integration in the classroom. These results are supported by the results of a number of research studies such as Roberts and Ferris, 1994; Slough and Chamblee, 2000; Flores, 2002; Earle, 2002; Zorfass and Rivero, 2005; & ChanLin et al., 2006.

Roberts and Ferris (1994) stated that barriers to technology integration included lack of knowledge of available hardware and software, time commitment, and the risk of using technology. Similarly, Slough and Chamblee (2000) argued that a view of technology as something unstable and always changing presents a major barrier to its use in the classroom. Moreover, Flores (2002) concluded that teachers face many barriers in their quest to incorporate technology such as time scheduling for technology use and administrative support, equity, and the lack of resources. Earle (2002) pointed out to extrinsic barriers to technology integration such as access, time, support, resources, and training and forces that are intrinsic to teachers such as attitudes, beliefs, practices, and resistance. On the other hand, ChanLin et al. (2006) supported these findings by identifying the factors that influenced teachers' use of technology in teaching. These factors were classified into four categories: environmental, personal, social and curricular issues.

Table 2: Teachers' Perceptions of Obstacles and Incentives Related to Successful Technology Integration in Classroom

Variable	Mean	SD
Obstacles		
The teacher does not have much time to prepare and implement them	3.4	1.2
Curricula are not ready to use such new technologies	3.0	1.3
Not enough encouragement to use them	2.7	1.3
Qualified staff for the labs are not available to help	2.5	1.3
Equipped labs are not available in schools	2.1	1.2
Technologies are not available in schools	2.0	1.1
Incentives		
Free or discounted computers for their own use	3.6	1.6
Participation in special workshops	3.0	1.2
Additional resources for their classroom	3.0	1.2
Positive evaluations	3.0	1.3
School or educational zone recognition program	3.0	1.2
Free software.	2.8	1.6
Release time	2.7	1.4
Salary supplement	2.3	1.5
Mentor teacher designation (or similar designation)	2.3	1.3

To answer question 3 "How do teachers perceive their students' usage of technology in the classroom?", results showed that teachers had high perception of students' usage of technology (see Table 3). They reported high usage of technology for interaction and communication, independent learning, engagement in learning, and understanding of academic subjects. The mean score for each of these items was 4.0 on a 5-point scale. These results are supported by Holinga (1999) who studied how Project LINCOLN in Springfield, Illinois, changed children's education in an important and meaningful way. The result of the project showed that student achievement has improved across all grades.

To answer question 4 "To what extent do teachers perceive their usage of technology tools in the classroom?", results indicated that teachers use a number of technologies in their classrooms such as computers with different software, transparencies, the Internet, maps, OHP, and Flyers & Folded Papers (see Table 4). Mean scores for the usage of these tools were 3.3 or above. These results are supported by Ertmer et al. (1999) who found that teachers' perceptions of the role of technology are closely linked to how technology is used. Another study conforming the results of this study was conducted by Kotrlik and Redmann (2005), where results revealed that although teachers feel some anxiety when it comes to technology integration, they perceived that they are effective in using technology.

Table 3: Teachers' Perceptions of their Students' Usage of Technology in classroom

Variable	Mean	SD
Students are interacting and communicating differently with the help of technology	4.0	0.9
Students become more independent learners as a result of technology.	4.0	0.8
Students are more engaged in learning due to technology.	4.0	0.7
Student understanding of academic subjects has deepened due to technology use	4.0	0.8
Students use technology to improve their basic skills with computer programs.	3.8	0.8
Students are developing online research expertise.	3.8	0.9
Students do more school work when not in school	3.8	0.6
The primary student-related use of technology is to teach students how to use the technology itself.	3.7	0.9
Schools report that students have better grades and/or test scores since they began using technology	3.7	1.0
Students use technology in at least some of their regular classrooms.	3.6	1.1
Schools report an increase in attendance on days that students are scheduled to use technology.	3.2	1.0
Students use computers only in a lab	3.1	1.2
Schools have reported decreases in the student dropout rate attributed to the use of technology.	3.1	1.2
Students actively participate in distance learning with other schools.	2.7	1.3

Away from perceptions, data analysis for the first focus group interview with male teachers indicated that most male teachers believe that using technology is important, but not all the time. On the other hand, they indicated that technology has many advantages for the teaching-learning process. It saves class time, minimizes teachers' efforts, grasps students' attention, and makes learning interesting. Students' understanding is the most important factor that teachers could use to evaluate the effectiveness of using technology in their classrooms.

Most female teachers highly regard technology and are using different types of applications in their classes such as computers, visual projectors, and the Internet. Female teachers think that technology helps facilitate learning and teaching, increases student participation, and provides visual support for students of different learning styles.

Table 4: Teachers' Perceptions of their Usage of Technology Tools in Classroom

Variable	M	SD
Computer	4.6	0.5
Transparencies	4.6	0.6
Different Computer Software	4.5	0.7
Geographic maps	4.2	1.0
Internet	4.0	1.1
Over Head Projector	4.0	1.1
Flyers & Folded Papers	4.0	1.1
Electronic Mail	3.8	1.2
Posters	3.8	1.0
Video	3.7	1.2
Wood Manipulatives	3.7	1.2
Drawing Tools	3.6	1.1
Tools for Creating Models	3.6	1.2
Raw Materials & Real Things (e.g., Seeds, Buttons, Bean Pills,…)	3.6	1.2
Video	3.5	1.3
TV	3.4	1.4
Distance Learning Equipment and Infrastructure	3.3	1.1

To answer question 5 “What is the difference in perception of male and female teachers in technology integration?”, a multivariate analysis was run. Results indicated a significant difference between the two groups with a Hotelling’s trace value of 9.3 with a significant f. of 10.75. To locate the significant differences within subscales, a one way analysis of variance (ANOVA) was run. However, in order to control Type I error when conducting the analysis of Variance, the researchers adjusted α level (0.05) using Benfaroni modification method. The adjusted value of α is ≤ 0.005. Table 5 shows the items that yielded significant differences within the sub-themes.

As seen from the table, technology availability was a concern for female teachers more than it was for males. In spite of this fact, results showed that female teachers use different types of technologies more than male teachers do. The means scores for female teachers on technologies used are all above 4.4, while the mean scores for male teachers ranged from 2.5 to 3.5. This might indicate that female teachers integrate technology in their classrooms more than male teachers do.

On the other hand, Hong and Koh (2002) found that female teachers were more anxious than male teachers toward hardware. They also found that the overall computer anxiety levels of male teachers were not significantly different from the anxiety levels of female teachers. Only for the hardware anxiety domain were significant differences detected between male and female teachers.

Table 5: Differences Between Male and Female Teachers in their Perception of Technology Integration

	M	F	f.
Teachers Perception of their Competencies in Technology Integration			
I can use a variety of media and formats, including telecommunications, to collaborate, publish, and interact with peers, experts, and other audiences.	3.9	4.4	0.0
creating multimedia presentations.	5.0	4.8	0.0
using computers for on-line communication (e.g., e-mail).	4.7	4.2	0.0
designing web pages.	3.3	4.2	0.0
Obstacles			
Technologies are not available in schools	1.7	2.8	.00
Qualified staff for the labs are not available to help	2.2	3.1	.01
Technologies that Might be Used			
Video	2.8	4.6	.00
Over Head Projector	3.3	4.7	.00
TV	2.5	4.7	.00
Electronic Mail	3.7	4.6	.01
Internet	3.7	4.7	.01
Distance Learning Equipment and Infrastructure	3.0	4.1	.01
Video	2.8	4.6	.00
Over Head Projector	3.3	4.7	.00
TV	2.5	4.7	.00
Electronic Mail	3.4	4.6	.00
Internet	3.7	4.7	.01
Wood Manipulatives	2.9	3.9	.02
Video	2.8	4.6	.00
Models and 3D Pieces	3.8	4.7	.02
Posters	3.4	4.4	.00
Transparencies	3.5	4.6	.00
Drawing Tools	3.5	4.6	.00
Tools for Creating Models	3.0	4.4	.00
Raw Materials & Real Things (e.g., Seeds, Buttons, Bean Pills,…)	3.1	4.7	.00
Flyers & Folded Papers	3.8	4.4	.02
Results of Teachers' Beliefs about Technologies and Using them in Instruction			
Most students have so many other needs that technology use is a low priority	3.3	4.0	.04

From focus group interviews the following can be concluded: (1) female teachers have more experience, familiarity, and knowledge of technology resources and applications than male teachers, (2) male teachers think that technology should be a part of the curriculum plan and that they should receive rewards for their technology integration performance, (3) most male and female teachers are mainly focusing on the use of computers and transparencies in their classes, (4) both male and female teachers think technology should be used only when needed while teachers should use a variety of teaching methods, and (5) all teachers agree that lesson goals and the nature of the subject are the two factors that determine the type of technology the teacher should use.

Conclusion

Study results show that both male and female teachers at UAE Model Schools have high self perception of their abilities and competencies to integrate technology successfully in their teaching. In addition, results revealed that teachers integrate technology in their classes with different degrees and effectiveness in spite of the barriers that hinder such integration (e.g., technical problems, large number of students, lack of professional development training, lack of motivation and financial support, and negative teacher and parent attitudes toward the impact of technology on teaching and learning).

In order to increase effective technology integration, both male and female teachers recommend the following: (1) regular professional development workshops, (2) enhancing curriculum with technology-enhanced materials such as CDs and videos, (3) increasing collaboration between schools across the country, and (4) giving enough freedom for teachers in the selection and coverage of curriculum materials.

It is worth mentioning that when model schools at the UAE were inaugurated more than a decade ago, they had advantages over typical schools, particularly in their infrastructure and teacher professional development activities. Due to the success of these schools, most public schools around the country started to follow their path. As a result, these days the gap between model schools and public schools almost vanished when it comes to technology availability and teacher professional development. Most public schools, particularly in Abu Dhabi have more or less the same advantages of model schools, particularly when it comes to technology equipment and teacher training. Thus, the implication of this change is that the results obtained from this study can be easily generalized to other UAE public schools covering the same grade levels as model schools.

These results were consistent with other studies investigating the same issues (e.g., Slough and Chamblee, 2000; Guha, 2000; Flores, 2002; Earle, 2002; Shelly et al., 2002; Bauer and Kenton, 2005; Kotrlik and Redmann, 2005; Zorfass and Rivero, 2005; and ChanLin et al., 2006).

Based on the above findings, the researchers recommend the following to enhance teachers' skills and competencies in technology integration regardless of country or gender:

1. Enhance teachers' technology integration abilities and skills by delivering workshops about effective technology integration.
2. Provide teachers with state-of-the-art technology including hardware and software.
3. Provide teachers with incentives and awards for outstanding technology integration in their classrooms.
4. Provide teachers with some release time so that they can plan effectively for technology integration in teaching and learning.
5. Explore the use of technology in classrooms covering all school levels, including public and private schools.
6. Investigate the effect of technology integration on students' achievement and attitude.
7. Investigate technology integration in relationship to curriculum goals and outcomes.

References

Abbit, J. T., & Klett, M. D. (2007). Identifying influences on attitudes and self –efficacy beliefs towards technology integration among pre-service educators: *Electronic Journal for the integration of technology in Education*, *6*, 28-42.

Almekhlafi, A.G. (2006a). The effect of computer assisted language learning (CALL) on United Arab Emirates English as a foreign language (EFL) school students achievement and attitude. *Journal of Interactive Learning Research*, *17*(2), 121-142.

Almekhlafi, A.G. (2006b). Effectiveness of interactive multimedia environment on language acquisition skills of 6th grade students in the United Arab Emirates. *International Journal of Instructional Media*, *33* (4), 427, 241.

Anderson, S., & Maninger, R, (2007). Preservice teachers' abilities, beliefs, and intentions regarding technology integration. *Journal of Educational Computing Research*, *37* (2), 151-172.

Bauer, J., & Kenton, J. (2005). Toward technology integration in the schools: Why it isn't happening. *Journal of Technology and Teacher Education*, *13* (4), 519-546.

Becker, H. J., & Ravitz, J. L. (2001). Computer use by teachers: Are Cuban's predictions correct? *Paper presented at the American Educational Research Association*, Seattle. Retrieved February 15, 2009 from http://www.crito.uci.edu/tlc/indings/confer-ences-pdf/aera_2001.pdf.

Bhargava, A., Kirova-Petrova, A., & McNair, S. (1999). Computers, gender bias, and young children. *Information Technology in Childhood Education Annual*, January 1, 263-274.

Brinkerhof, J. (2006). Effects of a long-duration, professional development academy on technological skills, computer self-efficacy, and technology integration beliefs and practices. *Journal of Research on Technology in Education*, 39 (1), 22-44.

ChanLin, L., Hong, J., Horng, J., Chang, S., & Chu, H. (2006). Factors influencing technology integration in teaching: A Taiwanese perspective. *Innovations in Education and Teaching International*, *43* (1), 57-68.

Earle, R. S. (2002). The Integration of instructional technology into public education: promises and challenges. *Educational Technology*, *42* (1), 5-13.

Ertmer, P. A., Addison, P., Lane, M., Ross, E., & Woods, D. (1999). Examining teachers' beliefs about the role of technology in the elementary classroom. *Journal of Research on Computing in Education*, *32* (1), 54–72.

Flores, A. (2002). Learning and teaching mathematics with technology. *Teaching Children Mathematics*, *8* (6), 308-325.

Goddard, M. (2002). What do we do with these computers? Reflections on technology in the classroom. *Journal of Research on Technology in Education*, *35* (1), 19-26.

Guha, S. (2000). *A Comparative analysis of present and preferred situations of elementary grade teachers in using computers for classroom instruction*, ERIC Document Reproduction Service No. ED440089.

Gulbahar, Y. (2007). Technology planning: A Roadmap to successful technology integration in schools. *Computers and Education*, *49* (4), 943-956.

Holinga, M. J. (1999). Project LINCOLN: improving and enhancing student learning. *Learning and Leading with Technology*, *26* (7), 54-80.

Hong, K., & Koh, C. (2002). Computer anxiety and attitudes toward computers among rural secondary school teachers: A Malaysian perspective. *Journal of Research on Technology in Education*, *35* (1), 27-46.

Judson, E. (2006). How teachers integrate technology and their beliefs about learning: Is there a connection? *Journal of Technology and Teacher Education*, *14* (3), 581-597.

Kotrlik, J., & Redmann, D. (2005). Extent of technology integration in instruction by adult basic education teachers. *Adult Education Quarterly: A Journal of Research and Theory*, *55* (3), 200-219.

Lam, Y., & Lawrence, G. (2002). Teacher-student role redefinition during a computer-based second language project: Are computers catalysts for empowering change? *Computer Assisted Language Learning*, *15* (3), 295-315.

Levin, T., & Wadmany, R. (2006). Teachers' beliefs and practices in technology-based classrooms: A Developmental view. *Journal of Research on Technology in Education*, *39* (2), 157–181.

Manzo, K. K. (2001). Academic record. *Education Week*, 20(35), 22-35. Washington.

NCES (2000). *Internet access in U. S. public schools and classrooms: 1994–99*. Washington, DC: NCES 2000–086.

Redmann, D., & Kotrlik, J. (2004). Technology integration into the teaching- learning process by business education teachers. *Delta Pi Epsilon Journal*, *46* (2), 76-91.

Roberts, N., & Ferris, A. (1994). Integrating technology into a teacher education program. *Journal of Technology and Teacher Education*, *2* (3), 215-225.

Rowand, C. (2000). *Teacher use of computers and the internet in public schools. Stats in Brief*, ERIC Document Reproduction Service No. 442463.

Sandholtz, J. H. (2001). Learning to teach with technology: A Comparison of teacher development programs. *Journal of Technology and Teacher Education*, *9*, (3), 349.

Shashaani, L. (1997). Gender differences in computer attitudes and use among college students. *Journal of Educational Computing Research*, *16* (1), 37-51.

Shelly, B. G., Cashman, T.J., Gunter, R. E., & Gunter, G. A. (2002). *Teachers discovering computers: Integrating technology in the classroom (2nd Ed.)*, Boston, MA: Course Technology.

Sherry, L., Bilig, S., Jesse, D., & Acosta, D. W. (2001). Assessing the impact of instructional technology on student achievement. *T.H.E. Journal*, *28* (7), 40-43.

Slough, S. W., & Chamblee, G. E. (2000). Implementing technology in secondary science and mathematics classrooms: A perspective on change. In D. A. Willis, J. D. Willis, & J. Willis (Eds.), *Proceedings of the Society for Information Technology and Teacher Education International Conference*, 1021–1026, Charlottesville, VA: AACE.

Totter, A., Stutz, D., & Grote, G. (2006). ICT and schools: Identification of factors influencing the use of new media in vocational training schools. *The Electronic Journal of e-Learning*, *4* (1), 95–102.

Wernet, S. P., Olliges, R. H., & Delicath, T.A. (2000). Postcourse evaluation of WebCT (Web Course Tools) classes by social work students. *Research on Social Work Practice*, *10* (4), 487-504.

Wood, R.; & Ashfield, J. (2008). The use of the interactive whiteboard for creative teaching and learning in literacy and mathematics: a case study. *British Journal of Educational Technology*, *39* (1), 84-96.

Yildirim, S. (2000). Effects of an educational computing course on preservice and inservice teachers: A discussion and analysis of attitudes and use. *Journal of Research on Computing in Education*, *32* (4), 479-495.

Zhao, Y. (2007). Social studies teachers' perspectives of technology integration. *Journal of Technology and Teacher Education*, *15* (3), 311-333.

Zorfass, J., & Rivero, H. (2005). Collaboration is a key: How a community of practice promotes technology integration. *Journal of Special Education Technology*, *20* (3), 51-60.

Johnson, G. M. (2010). Internet Use and Child Development: Validation of the Ecological Techno-Subsystem. *Educational Technology & Society*, *13* (1), 176–185.

Internet Use and Child Development: Validation of the Ecological Techno-Subsystem

Genevieve Marie Johnson
Department of Psychology, Grant MacEwan College, Edmonton, Canada T5J 4S2 // gen.johnson@shaw.ca

ABSTRACT

Johnson and Puplampu recently proposed the *ecological techno-subsystem*, a refinement to Bronfenbrenner's theoretical organization of environmental influences on child development. The ecological techno-subsystem includes child interaction with both living (e.g., peers) and nonliving (e.g., hardware) elements of communication, information, and recreation technologies in immediate or direct environments. The theoretical techno-subsystem requires empirical validation. Parents of 128 children in first through sixth grade consented to cognitive developmental assessment of their children and completed questionnaires on children's use of the Internet at home and family socioeconomic characteristics. In general, indices of home Internet use accounted for more of the variance in children's cognitive development than did indices of socioeconomic status. The ecological techno-subsystem furthers our understanding of environmental influences on child development by emphasizing the impact of digital technologies on cognitive growth during childhood.

Keywords

Ecological techno-subsystem, Child development, Child cognition, Ecological theory

Introduction

According to the Corporation for Public Broadcasting (2002), the prevalence of Internet use among American 6 to 8 year old children doubled between 2000 and 2002 (from 27% to 60%, across all locations, at least one a week). Approximately 20% of Canadian 9 year old children access the Internet through their own personal computer (Media Awareness Network, 2006). The Office of Communication (2007) reported that 7% of British 10-year-olds have a webcam. In Australia, nine in ten families have home Internet connectivity and 75% have broadband access (Australian Communications and Media Authority, 2007). Trends indicate continued increase in the number of children accessing the Internet, the amount of time they spend online, and the complexity of their online behavior (Livingstone & Helpsper, 2007).

Historically, panic surrounds the introduction of new technologies, particularly in relation to children and youth (Johnson, 2006). For example, in the 19th century, "the telegraph enabled a young woman, against her father's wishes, to maintain a flirtation with a number of men on the wire" (Quigley & Blashki, 2003, p. 311). In the 21st century, there are two conflicting public anxieties surrounding children and the Internet; first, that the Internet may harm children, for example, by exposure to inappropriate content (Media Awareness Network, 2008) and, second, that children without Internet access are cognitively and socially disadvantaged (Jackson et al., 2006). Public anxiety surrounding the digital divide (Burnett & Wilkinson, 2005; Livingstone & Helpsper, 2007), increasingly complex school Internet literacy curriculum (Johnson, 2007a; Takahira, Ando, & Sakamoto, 2007), and social policy initiatives directed at enhancing childhood Internet access (Sandvig, 2003) reveal the extent to which Internet use is perceived as developmentally appropriate (if not required). Indeed, there is mounting evidence that using the Internet provides children with cognitive and social benefits (Greenfield & Yan, 2006).

Internet Use and Child Development

Particularly during periods of rapid development associated with childhood, Internet use stimulates cognitive and psychosocial development (Johnson, 2006; Young, 2007). Fish and colleagues (2008) investigated home computer experience and cognitive development among preschool children in inner-city Head Start programs. Data was collected from parents regarding the children's experience with computers in the home environment, including access to a computer, time spent on a computer, and types of computer programs used. Two hundred participating children were administered standardized tests of cognitive development. After controlling for parent's education and household income, children who had home computer access had significantly higher scores of cognitive development than did children who did not have home access. Frequency of children's computer use also related to cognitive development. The investigators concluded that early computer use at home was a positive influence on young children's cognitive development.

The Internet, although rich in graphic display, is primarily a text-based medium; “the more a child uses the Internet, the more he/she reads” (Jackson et al., 2007, p. 188). Li and Atkins (2004) concluded that computer exposure during the preschool years was associated with subsequent school readiness. Jackson and colleagues (2006) provided low income children with home-based Internet access and continuously recorded time online. “Findings indicated that children who used the Internet more had higher scores on standardized tests of reading achievement and higher grade point averages 6 months, 1 year, and 16 months later than did children who used the Internet less” (p. 429). Fuchs and Wößmann (2005) inferred, having controlled for socioeconomic status, “a negative relationship between home computer availability and academic achievement, but a positive relationship between home computer use for Internet communication” (p. 581). From a developmental perspective, Internet use stimulates cognitive processes involved in interpreting text and images (Johnson, 2006). Metacognitive processes such as planning, search strategies, and evaluation of information are exercised when navigating websites (Tarpley, 2001).

DeBell and Chapman (2006) concluded that Internet use promotes cognitive development in children, “specifically in the area of visual intelligence, where certain computer activities -- particularly games -- may enhance the ability to monitor several visual stimuli at once, to read diagrams, recognize icons, and visualize spatial relationships” (p. 3). Van Deventer and White (2002) observed proficient 10- and 11-year-old video gamers and noted extremely high levels of self-monitoring, pattern recognition, and visual memory. In a comprehensive review of the literature of the time (when interactive digital games were relatively unsophisticated), Subrahmanyam, Kraut, Greenfield, and Gross (2000) concluded that “children who play computer games can improve their visual intelligence” (p. 128). It should be noted, however, that playing video games has also been linked to childhood distractibility, over-arousal, hostility, and aggression (Anderson, Gentile, & Buckley, 2007; Funk, Chan, Brouwer, & Curtiss, 2006).

While Internet use during childhood has been associated with negative developmental outcomes, research increasingly suggests that the Internet provides children with more developmental advantages than disadvantages (Greenfield & Yan, 2006). Comprehensive theoretical description of the developmental impact of Internet use is required. The recently proposed *ecological techno-subsystem* (Johnson & Puplampu, 2008) provides a conceptual framework for understanding the effect of Internet use on child development.

Ecological Systems Theory and the Techno-Subsystem

Contemporary theories of child development assume that biological predispositions and environmental experiences, to varying combined degrees, result in social, emotional, and cognitive growth. Cognitive-developmental theories assume that neurological maturation and environmental experience result in individuals who are progressively more able to function effectively in their environments (Luria, 1976). A socio-cultural orientation to child cognitive development presupposes that “through participation in activities that require cognitive and communicative functions, children are drawn into the use of these functions in ways that nurture and scaffold them" (Vygotsky, 1986, pp. 6-7). Ecological systems theory (Bronfenbrenner, 1979) presents a particularly comprehensive view of environmental influences on development by situating the child within a system of relationships affected by multiple levels of the surrounding environment.

Bronfenbrenner (1979) organized the contexts of development into five nested environmental systems, with bi-directional influences within and among systems. The *microsystem* refers to the immediate environment and includes, most notably, home and school interactions. The *mesosystem* is comprised of connections between immediate environments (e.g., home-school interactions). The *exosystem* includes environmental settings that indirectly affect child development (e.g., the parent's workplace). The *macrosystem* refers to overarching social ideologies and cultural values. The *chronosystem* highlights the effect of time on all systems and all developmental processes. As his theory evolved, Bronfenbrenner (2005) proposed a bioecological perspective which views the child's own biology as part of the microsystem. Bronfenbrenner (1989) described human development as “the progressive, mutual accommodation, throughout the life course, between an active, growing human being, and the changing properties of the immediate settings in which the developing person lives, as this process is affected by the relations between these settings, and by the larger contexts in which the settings are embedded” (p. 188).

Ecological systems theory (Bronfenbrenner, 1979) emerged prior to the Internet revolution and the developmental impact of then available technology (e.g., television) was conceptually situated in the child's microsystem. Johnson and Puplampu (2008) recently proposed the *ecological techno-subsystem*, a dimension of the microsystem. As illustrated in Figure 1, the techno-subsystem includes child interaction with both living (e.g., peers) and nonliving (e.g., hardware) element of communication, information, and recreation technologies in immediate or direct environments. Since tools, by definition, extend human capabilities, interaction with increasingly complex tools requires increasingly complex cognitive processes (Johnson, 2008; Nickerson, 2005). The Internet extends human access to information and communication and provides cognitive scaffolding (e.g., search engines and e-directories) which allows for higher-order processes such as evaluation and application of information to solve real problems.

Research Issues and Questions: Validation of the Ecological Techno-Subsystem

The utility of the recently proposed ecological techno-subsystem in explaining child development has not been established nor investigated. From an ecological perspective, the techno-subsystem mediates bidirectional interaction between the child (i.e., bioecology) and the family (i.e., microsystem). Does the techno-subsystem contribute to increased understanding of the mechanisms of cognitive development during childhood? Which is the better predictor of cognitive development during childhood, -- indices of home Internet use (elements of the techno-subsystem) or family socioeconomic characteristics (elements of the microsystem)?

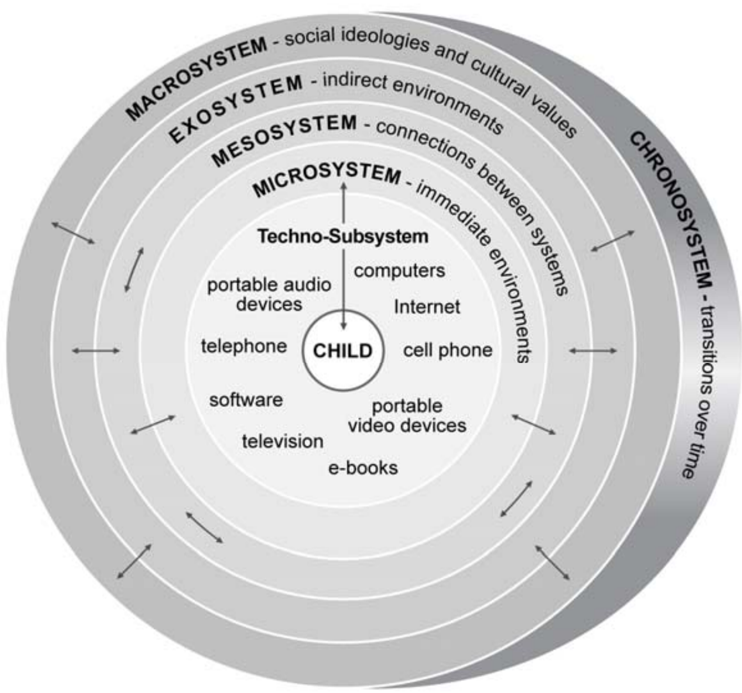

Figure 1. The Ecological Techno-Subsystem (Johnson & Puplumpu, 2008)

Methods

Participants

Parents of children in first through sixth grade (N = 151) attending an elementary school in suburban Western Canada were invited to participate in the study. Parents completed a questionnaire and consented to cognitive developmental assessment of their children. One hundred twenty-eight signed consent forms and completed parent questionnaires were returned to the school. Participating children (62 males and 66 females) ranged in age from 6 years, 4 months to 12 years, 5 months; 14.8% of the children were in first grade, 12.5% were in second grade, 15.6%

were in third grade, 25.0% were in fourth grade, 16.4% were in fifth grade, and 15.6% were in sixth grade. Twelve of the 128 children were funded for special needs (e.g., communication disorder, learning disability, behavior disorder, medical condition).

Measures

Three constructs, corresponding to three ecological systems/subsystems, were measured: child cognitive development (bioecology), indices of child use of the Internet at home (techno-subsystem), and family socioeconomic characteristics (microsystem).

Child Cognitive Development

Children's cognitive development was measured with one subtest from the fourth edition of Wechsler Intelligence Scale for Children (WISC-IV; Wechsler, 2003) and three subtests from the Cognitive Assessment System (CAS; Das & Naglieri, 2001). Subtests were selected to ensure comprehensive representation of child cognitive development (i.e., language, metacognition, perception, and memory). Expressive language was assessed with the WISC-IV vocabulary subtest (child provides verbal definitions to orally presented words). WISC-IV subtest scoring criteria was maintained; norms were not required because all comparisons occurred within the group of 128 children. With respect to the CAS, the matching numbers subtest measured metacognitive planning (find the two numbers that are the same in a series of numbers), the nonverbal matrices subtest assessed visual perception (select an option that best completes a visual matrix), and the word series subtest measured short-term auditory memory (repeat a string of words presented orally); CAS scoring criteria was maintained.

Each of the 128 children was individually administered the four cognitive subtests by one of two examiners (an educational psychologist and a trained research assistant). Rapport was initiated by in-class introduction of the examiners, explanation of testing procedures, and response to class questions. Rapport was further established by individual child-examiner interaction walking from the classroom to the testing room and, as required, upon entry into the testing room. Each individual assessment was complete in approximately 15 minutes. Table 1 presents a summary of cognitive developmental measures and description of children's cognitive scores.

Table 1. Summary of Cognitive Developmental Measures and Description of Children's Scores

			Children's Raw Scores		
Cognitive Skill	Test and Subtest	*N*	*Range*	*Mean*	*SD*
Expressive Language	WISC Vocabulary	127	14 – 54	30.7	8.38
Metacognitive Planning	CAS Matching Numbers	126	1 – 11	7.1	1.84
Visual Perception	CAS Nonverbal Matrices	128	4 – 33	14.3	4.64
Auditory Memory	CAS Word Series	128	2 – 16	9.7	2.51

Indices of Child Home Internet Use

The parent questionnaire included two yes/no response items: Do you have the Internet in your home? Does your child use the Internet at home? Approximately 83% of families (106/128) reported home Internet access; 71.9 % (92/128) indicated that their child used the Internet at home. For purposes of the current investigation, five indices of child home Internet use were obtained from parental response to questionnaire items. First, parents reported the number of years of home Internet access (range 0.2 to 12 years, mean 5.2 years, standard deviation 2.96 years). Additionally, parents who reported that their child used the Internet at home were asked to respond to the open-ended questionnaire item, what does your child do when he/she uses the Internet at home? Thematic analysis of 90 parental responses to the open-ended item revealed four categories or types of child home online behavior: learn (e.g., schoolwork, math practice, research for assignments), play (e.g., play games, have fun with friends), browse (e.g., visit websites, find things of interest), and communicate (e.g., email, chat). Approximately 17% of parents responded to the open-ended questionnaire item with description that suggested one type of child online behavior, 35.9% described two, 14.1% described three, and 3.1% described four types of online behavior. Using the Internet at home to learn was reported in 65 cases, to play was reported in 57 cases, to browse in 35 cases, and to communicate in 27 cases. Thus, the five indices of child home Internet use included: 1) the continuous variable years of home Internet access and the dichotomous (reported-unreported) variables of child home Internet use to 2) learn, 3) play, 4) browse, and 5) communicate.

Family Socioeconomic Characteristics

The parent questionnaire assessed five family characteristics commonly used to determine socioeconomic status (Bradley & Corwyn, 2002; Sirin, 2005). Two items queried father's and mother's employment status. Approximately 70% of mothers and 96% of fathers were employed, full-time or part-time. Two questionnaire items requested father's and mother's level of education, coded as: elementary = 1, junior high school = 2, high school incomplete = 3, high school complete = 4, technical school/college (complete or incomplete) = 5 and university (complete or incomplete) = 6. The mean educational level of mothers was 4.79 (SD = 0.95) suggesting that many mothers had post-secondary education; the mean educational level of fathers was 4.45 (SD = 1.02) suggesting that some fathers had post-secondary education. The final socioeconomic item on the questionnaire asked parents to indicate annual family income by selecting one of the following options: < \$20 000 = 1, \$20 000 to \$40 000 = 2, \$40 000 to \$60 000 = 3, \$60 000 to \$80 000 = 4, \$80 000 to \$100 000 = 5, > \$100 000 = 6. Annual income for participating families was approximately \$60,000 CD (M = 4.07, SD = 1.48).

Table 2 presents a summary of measured constructs which includes: four tests of children's cognitive development, five indices of children's home Internet use, and five family socioeconomic characteristics. Which are the better predictors of cognitive development during childhood, -- elements of the microsystem or elements of the techno-subsystem? Two series of stepwise regression analysis were conducted with the four cognitive development scores as the dependant variables. In the first regression analyses, family socioeconomic characteristics (elements of the microsystem) were the independent variables. In the second analyses, indices of home Internet use (elements of the techno-subsystem) were the independent variables.

Table 2. Description of Constructs and Measures

Ecological System	System Elements	Specific Measures
Bioecology	Cognitive Development	Expressive Language
		Metacognitive Planning
		Visual Perception
		Auditory Memory
Techno-Subsystem	Home Internet Use	Years of Internet Access
		Online Learning
		Online Playing
		Online Browsing
		Online Communication
Microsystem	Family Characteristics	Father Employment
		Mother Employment
		Father Education
		Mother Education
		Annual Family Income

Results

Results of analyses revealed that family socioeconomic characteristics (elements of the microsystem) explained a modest (but significant) amount of the variation in children's cognitive development scores. As presented in Table 3, adjusted R^2 values indicated that father's level of education accounted for approximately 7% of the variation in children's level of expressive language (as measured by the WISC-IV vocabulary subtest), 5% of the variation in children's visual perception and auditory memory (as measured by the CAS nonverbal matrices subtest and CAS word series subtest, respectively). Whether or not mothers were employed, part-time or full-time, accounted for approximately 6% of the differences in children's capacity to execute metacognitive functions such as planning (as measured by the CAS matching numbers subtest). While the other measures of familial socioeconomic status (e.g., mother's education and family income) explained some of the variance in children's cognitive development, such measures did not improve upon the predictive utility of father's education or maternal employment; variation is prerequisite to prediction. Almost all fathers were employed and almost all mothers had finished high school. For participating middle-class families, father's education and mother's employment were more sensitive to children's cognitive development scores than were family income, father's employment, and mother's education.

Table 3. Stepwise Regression Analysis: Family Characteristics Predicting Child Cognitive Development

Cognitive Score	Predictor	Beta Weight	*t* value	$R^2_{(adj)}$	*F* value
Expressive Language	Father Education	.292	2.70**	.074	(1, 78) = 7.29**
Metacognitive Planning	Mother Employed	.270	2.46*	.061	(1, 77) = 6.05*
Visual Perception	Father Education	.244	2.22*	.047	(1, 78) = 4.93*
Auditory Memory	Father Education	.258	2.36*	.054	(1, 78) = 5.55*

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

Results of analyses further revealed that indices of home Internet use (elements of the techno-subsystem), in general, explained more of the variation in children's cognitive development than did family socioeconomic characteristics (elements of the microsystem). Summarized in Table 4, specific types on online behavior (i.e., learning, communicating, and playing) and years of home Internet access combined to predicted child cognitive developmental outcomes. Indicated by adjusted R^2, children's online communication, years of home Internet access, and online learning (as reported by parents) accounted for approximately 29% of the variation in children's level of expressive language as measured by the WISC-IV vocabulary subtest. Online learning and communicating (reported-unreported) combined to explain 13.5% of the variation in children's metacognitive planning. Online learning and playing (reported-unreported) combined to explain 10.9% of the variation in children's auditory memory. Years of home Internet access explained approximately 3% of the differences in children's visual perception scores. With the exception of visual perception, indices of home Internet use (elements of the techno-subsystem) were better predictors of children's cognitive development than were family socioeconomic characteristics (elements of the microsystem).

Table 4. Stepwise Regression Analysis: Home Internet Use Predicting Child Cognitive Development

Cognitive Score	Predictor/s	Beta Weight	*t* value	$R^2_{(adj)}$	*F* value
Expressive Language	Online Communication	.344	4.00***		
	Years of Internet Access	.263	3.12 **		
	Online Learning	.256	2.99**	.287	(3, 101) = 14.97***
Metacognitive Planning	Online Learning	.287	3.03**		
	Online Communication	.201	2.12*	.135	(2, 101) = 9.06***
Visual Perception	Years of Internet Access	.192	1.99*	.028	(1, 104) = 3.98*
Auditory Memory	Online Learning	.242	2.60*		
	Online Playing	.228	2.46*	.109	(3, 101) = 14.97***

*p < .05; **p < .01; ***p < .001

Discussion

A variety of mechanisms linking family socioeconomic status to child cognitive development have been proposed including parenting (Petrill, Pike, Price, & Plomin, 2004; Mistry, Biesanz, Chien, Howes, & Benner, 2008) and resources (Bradley & Corwyn, 2002). For the current sample of middle class children, paternal education and maternal employment were associated with measures of child cognitive development. More educated fathers tended to have offspring who scored high on three of the four cognitive measures (expressive language, visual perception, and auditory memory). Mothers who were employed tended to have children who scored high on the measure of metacognitive planning. Educated fathers and employed mothers may genetically transmit to their offspring some neurological processing advantage (bioecology). Simultaneously, educated fathers may provide enhanced language models and stimulating environments that facilitate the cognitive development of their children (microsystemic influence). Employed mothers may provide models of organization and place increased demands on children to self-regulate thereby enhancing the metacognitive planning abilities of their offspring (microsystemic influence).

Family socioeconomic status (as measured and for the current sample) accounted for 5% to 7% of differences in child cognitive development scores. In contrast, indices of home Internet use (as measured and for the current sample) accounted for 3% to 29% of differences in child cognitive development scores. Meta-analysis confirms that the impact of socioeconomic status on academic achievement is eroding over time (Sirin, 2005). Increasingly effective structures of social equalization (e.g., public education, quality daycare, preschool intervention, and prenatal programs) and the expanding middle class create the need for more precise description of home environments. Current results suggest that indices of home Internet use (i.e., elements of the ecological techno-subsystem) provide more useful information regarding cognitive development than do family socioeconomic characteristics (elements of the microsystem).

Only two of five family socioeconomic characteristics added to the regression equation, suggesting that some measures (i.e., family income, father employment, and mother education) did not differ in relation to children's cognitive development. In contrast, four of the five indices of home Internet use during childhood added to the regression equation, suggesting that these measures differed in relation to children's cognitive development. In the context of the current investigation, socioeconomic status is a crude construct relative to home Internet use. Internet use includes both organized (e.g., search) and disorganized (e.g., browse) interactions with both human (e.g., chat) and nonhuman (e.g., database) elements in online environments (Johnson & Kulpa, 2007). Internet use is a complex set of behaviors that vary widely across individuals and that is influenced by cognitive and personality characteristics (Joinson, 2003). For the current sample of children, patterns of home Internet use explained more of the variation in cognitive development than did family socioeconomic characteristics.

In the context of middle class families, elements in the techno-subsystem (e.g., Internet access) may not necessarily facilitate child cognitive development; effective use of those elements, highly dependent upon parent behavior, may promote development. For example, Cho and Cheon (2005) surveyed families and found that parents' perceived control, obtained through shared web activities and family cohesion, reduced children's exposure to negative Internet content. Lee and Chae (2007) reported a positive relationship between parental mediation techniques (website recommendation and Internet co-use) and children's educational attainment. In the current investigation, the cognitive experiences provided to children by employed mothers may include Internet skills instruction (e.g., sending email) and models of information management (e.g., accessing websites for information). Such experiences, over time, may provide children with enhanced opportunities to direct their own cognitive development via increasingly sophisticated uses of the Internet. According to Livingston and Bober (2005), "a new divide is opening up between those for whom the internet is an increasingly rich, diverse, engaging and stimulating resource and those for whom it remains a narrow, unengaging, if occasionally useful, resource of rather less significance" (p. 2).

Bruner (2005) recently reiterated that "our minds appropriate ways of representing the world from using and relating to the codes or rules of available technology" (p. x). Cognitive abilities prerequisite to utilization of Internet applications constitute an implicit component of contemporary notions of intelligence (Maynard, Subrahmanyam, & Greenfield, 2005). The ecological techno-subsystem furthers our understanding of environmental influences on child development by emphasizing the impact of digital technologies on cognitive growth during childhood. The techno-subsystem provides precise description of microsystemic mechanisms of developmental influence which lead to intervention strategies. According to Livingston and Bober (2005), many parents lack the skills to guide and support their children's Internet use and Internet-literate parents have Internet-literate children. Subsequent research may evaluate the effectiveness of techno-subsystem interventions for elementary school children at-risk, for example, the provision of home Internet access and parent Internet literacy training. As stated elsewhere, "current anxiety surrounding children's Internet use should be for those whose cognitive processes are not influenced by the cultural tool" (Johnson, 2006, p. 570).

Limitations and Future Research

In the current investigation, children's use of the Internet at home was determined by parent-report (common in the literature, e.g., Livingston & Bober, 2005; Rideout, Vandewater, & Wartella, 2003). The validity of such approaches, however, has been questioned and alternatives suggested including asking the child directly (Media Awareness Network, 2006; Roberts, Foehr, & Rideout, 2005) and standardized measures such as the *Internet Vocabulary Test for Children* (Johnson, 2007b). In the current investigation, indices of children's use of the Internet at home were obtained with objective (i.e., years of home Internet access) and subjective (*what does your child do when he/she uses the Internet at home*) parental response to questionnaire items. Alternative indices of children's use of the Internet may not replicate current findings.

Type of child online behavior (learn, play, browse, and communicate) emerged from thematic analysis of parent response to an open-ended questionnaire item. Alternative abstraction is apparent. For example, parental response to the open-ended item, *what does your child do when he/she uses the Internet at home*, may be dichotomized into directed versus undirected or focused versus unfocused use of the Internet. Responses such as *schoolwork*, *math practice*, *research for assignments*, *email* and *chat* may be interpreted as reflecting goal-directed and focused behavior; responses such as *play games*, *have fun with friends*, *visit websites*, and *find things of interest* refer to behavior that is unfocused and undirected. As opposed to online learning and communication, it may be that focused and goal-directed Internet use contributes to cognitive development during childhood.

Childhood use of the Internet occurs in three contexts: home, school, and community. From an ecological perspective, Internet use in one environment influences Internet use in other environments. Because all children in the sample attended the same elementary school, school-based Internet experience was assumed equivalent. However, Gibson and Oberg (2004) noted that the quality of school-based Internet experience varies widely across classrooms. Subsequent theoretical and empirical research may expand techno-subsystem description to include child-peer interactions during home, school, and community Internet use.

References

Anderson, C. A., Gentile, D. A., & Buckley, K. E. (2007). *Violent video game effects on children and adolescents*, New York; Oxford University Press.

Australian Communications and Media Authority. (2007). *Media and communications in Australian families 2007,* Retrieved February 22, 2009, from http://www.acma.gov.au/WEB/STANDARD/pc=PC_310893

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*, 371-399.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*, Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1989). Ecological systems theory. *Annals of Child Development, 6*, 187-24.

Bronfenbrenner, U. (2005). *Making human beings human: Bioecological perspectives of human development*, Thousand Oaks, CA: Sage.

Bruner, J. (2005). Forward. In R. J. Sternberg & D. D. Preiss (Eds.), *Intelligence and technology: The impact of tools on the nature and development of human abilities (*pp. ix-xi), Mahwah, NJ: Lawrence Erlbaum.

Burnett, C., & Wilkinson, J. (2005). Holy lemons! Learning from children's uses of the Internet in out-of-school contexts. *Literacy, 39*, 158-164.

Cho, C. H., & Cheon, H. J., (2005) Children's exposure to negative Internet content: Effects of family context. *Journal of Broadcasting and Electronic Media, 49*, 488-509.

Corporation for Public Broadcasting. (2002). *Connected to the future: A report on children's Internet use*. Washington, DC. Retrieved February 22, 2009, from http://www.cpb.org/stations/reports/connected/.

Das, J. P., & Naglieri, J. A. (2001). The Das-Naglieri cognitive assessment system in theory and practice. In J. J. W. Andrews, D. H. Sakolfske, & H. L. Janzen (Eds.), *Handbook of psychoeducational assessment: Ability, achievement, and behavior in children* (pp. 34-64). San Diego, CA: Academic Press.

DeBell, M., & Chapman, C. (2006). *Computer and Internet use by students in 2003*. National Center for Educational Statistics. U.S. Department of Education, Washington, DC. Retrieved February 22, 2009, from http://nces.ed.gov/pubs2006/2006065.pdf.

Fish, A. M., Li, X., McCarrick, K., Butler, S. T., Stanton, B., Brumitt, G. A., et al. (2008). Early childhood computer experience and cognitive development among urban low-income preschoolers. *Journal of Educational Computing Research, 38*, 97-113.

Fuchs, T., & Wößmann, L. (2005). Computers and student learning: Bivariate and multivariate evidence on the availability and use of computers at home and school. *Brussels Economic Review, 47*, 359-385.

Funk, J. B., Chan, M., Brouwer, J., & Curtiss, K. (2006). A biopsychosocial analysis of the video-game-playing experience of children and adults in the United States. *Studies in Media and Information Literacy education, 6*(3), 1-15.

Gibson, S., & Oberg, D. (2004). Visions and realities of Internet use in schools: Canadian perspectives. *British Journal of Educational Technology, 35*, 569-585.

Greenfield, P., & Yan, Z. (2006). Children, adolescents, and the Internet: A new field of inquiry in developmental psychology. *Developmental Psychology, 42*, 391-394.

Jackson, L. A., Samona, R., Moomaw, J., Ramsay, L., Murray, C., Smith, A., & Murray, L. (2007). What children do on the Internet: Domains visited and their relationship to socio-demographic characteristics and academic performance? *CyberPsychology and Behavior, 10*, 182-190.

Jackson, L. A., Von Eye, A., Biocca, F. A., Barbatsis, G., Zhao, Y., & Fitzgerald, H. E. (2006). Does home Internet use Influence the academic performance of low income children? *Developmental Psychology, 42*, 429-435.

Johnson, G. M. (2006). Internet use and cognitive development: A theoretical framework. *E-Learning, 4, 565-573*.

Johnson, G. M. (2007a). Functional Internet literacy: Required cognitive skills with implications for instruction. *E-Learning, 4*, 433-441.

Johnson, G. M. (2007b). The Internet Vocabulary Test for Children: Preliminary development. *Internet Research, 17*, 235-248.

Johnson, G. M. (2008). Cognitive processing differences between frequent and infrequent Internet users. *Computers and Human Behavior, 24*, 2094-2106.

Johnson, G. M., & Kulpa, A. (2007). Dimensions of online behavior: Toward a user typology. *CyberPsychology and Behavior, 10*, 773-780.

Johnson, G. M., & Puplampu, P. (2008). A conceptual framework for understanding the effect of the Internet on child development: The ecological techno-subsystem. *Canadian Journal of Learning and Technology, 34*, 19-28

Joinson, A. N. (2003). *Understanding the psychology of Internet behaviour: Virtual worlds, real lives*, New York: Palgrave MacMillan.

Lee, S. J., & Chae, Y. G. (2007). Children's Internet use in a family context: Influence on family relationships and parental mediation. *CyberPsychology and Behavior, 10*, 640-644.

Li, X., & Atkins, M. S. (2004). Early childhood computer experience and cognitive and motor development. *Pediatrics, 113*, 1715-1722.

Livingstone, S., & Bober, M. (2005). *UK children go online: Emerging opportunities and dangers*. London, UK: London School of Economics. Retrieved February 22, 2009, from http://www.lse.ac.uk/collections/children-go-online/UKCGO_Final_report.pdf.

Livingstone, S., & Helpsper, E. (2007). Gradations in digital inclusion: Children, young people and the digital divide. *New Media & Society, 9*, 671-696.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations*, Cambridge MA: Harvard University Press.

Maynard, A. E., Subrahmanyam, K., & Greenfield, P. M. (2005). Technology and the development of intelligence: From the loom to the computer. In R. J. Sternberg & D. D. Preiss (Eds.), *Intelligence and technology: The impact of tools on the nature and development of human abilities* (pp. 29-54), Mahwah, NJ: Lawrence Erlbaum.

Media Awareness Network. (2006). *Young Canadians in a wired world*. Ottawa, ON. Media and Internet Education Resources. Retrieved February 22, 2009, from http://www.media-awareness.ca/english/research/YCWW/phaseII/key_findings.cfm.

Media Awareness Network. (2008). *Web awareness Canada – An Overview*. Ottawa, ON. Media and Internet Education Resources. Retrieved February 22, 2009, from http://www.media-awareness.ca/english/special_initiatives/web_awareness/.

Mistry, R., Biesanz, J., Chien, N., Howes, C., & Benner, A. (2008). Socioeconomic status, parental investments, and the cognitive and behavioral outcomes of low-income children from immigrant and native households. *Early Childhood Research Quarterly*, *23*, 193-212.

Nickerson, R. S. (2005). Technology and cognitive amplification. In R. J. Sternberg & D. D. Preiss (Eds.), *Intelligence and technology: The impact of tools on the nature and development of human abilities* (pp. 3-27). Mahwah, NJ: Lawrence Erlbaum.

Office of Communications. (2007). *The Communications Market, 2007*. London, UK. Retrieved February 22, 2009, from http://www.ofcom.org.uk/research/cm/cmr07/.

Petrill, S. A., Pike, A., Price, T., and Plomin, R. (2004). Chaos in the home and socioeconomic status are associated with cognitive development in early childhood: Environmental mediators identified in a genetic design. *Intelligence, 32*, 445-460.

Quigley, M., & Blashki, K. (2003). Beyond the boundaries of the sacred garden: Children and the Internet. *Educational Technology Review, 11*, 70-77.

Rideout, V. J., Vandewater, E. A., & Wartella, E. A. (2003). *Zero to six: Electronic media in the lives of infants, toddlers and preschoolers*. Menlo Park, CA: The Henry J. Kaiser Family Foundation. Retrieved February 22, 2009, from http://www.kaisernetwork.org/health_cast/uploaded_files/102803_kff_kids_report.pdf.

Roberts, D. F., Foehr, U. G., & Rideout, V. (2005). *Generation M: Media in the lives of 8 – 18 year olds*. Menlo Park, CA: The Henry J. Kaiser Family Foundation. Retrieved February 22, 2009, from http://www.kff.org/entmedia/7251.cfm.

Sandvig, C. (2003). Public Internet access for young children in the inner city: Evidence to inform access subsidies and content regulation, *The Information Society, 19*, 171-183.

Sirin, S. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417-453.

Subrahmanyam, K., Kraut, R., Greenfield, P., & Gross, E. (2000). The impact of home computer use on children's activities and development. *Future of Children, 10*, 123-144.

Takahira, M., Ando, R., & Sakamoto, A. (2007). Effect of Internet use on development of information literacy: A panel study with Japanese elementary school children. *Computers in the Schools, 24*, 65-82.

Tarpley, T. (2001). Children, the Internet, and other new technologies. In D. G. Singer & J. L. Singer (Eds.), *Handbook of children and the media* (pp. 547-556). Thousand Oaks, CA: Sage.

Van Deventer, S. S., & White, J. A. (2002). Expert behavior in children's video game playing. *Simulation and Games, 33*, 28-48.

Vygotsky, L. (1986). *Thought and language*, Cambridge, MA: The MIT Press.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children* (4th Ed.), San Antonio, TX: Harcourt Assessment.

Young, K. (2007). Toward a model for the study of children's informal Internet use. *Computers in Human Behavior, 24*, 173-184.

Uzunboylu, H., & Tuncay, N. (2010). Divergence of Digital World of Teachers. *Educational Technology & Society*, *13* (1), 186–194.

Divergence of Digital World of Teachers

Huseyin Uzunboylu and Nazime Tuncay
Department of Instructional Technology, Near East University, Nicosia, North Cyprus // huzunboylu@neu.edu.tr, nazime.tuncay@gmail.com

ABSTRACT

There exists great diversity in the teachers' digital world. Teachers are being discriminated based on numerous educational gaps. This paper seeks to assess the extent of the digital divide among the North Cyprus vocational teachers along the four axes: age, Internet access, computer access, and performance (computing knowledge/experience). A research was carried out through a questionnaire, which was then analyzed statistically. According to the experts' views, the questionnaire was divided into three factors: technology-based e-learning applications; web-based e-learning applications, and administrative e-learning applications. There was a significant digital divergence among the teachers surveyed, which may adversely affect their ability to prepare the students to become a part of the knowledge society. To bridge these gaps in the world, action plans should be prepared, collaboratively with the instructional technologist and ICT (Information and Communication Technology) experts. This unique research study is the first to investigate the divergence of the digital world of teachers.

Keywords

Digital world, gaps, teachers, e-learning

Introduction

Gaps in the Digital World

Gaps, which are defined as, have and have-nots; know and know-nots, are widening [United Nations Development Programme (UNDP), 1999; Malloch, 2000; James, 2000; Main, 2001; Dalsgaard, 2001; Cobb, 2002]. The obstacles that are defined as gaps significantly affect education. However, day-by-day, the effects of gaps in education has become clearer, and the easiness of e-learning applications in education has become more common among teachers. Unequal opportunities among countries in access to technology and Internet host have also been of interest for researchers for many decades. Correspondingly, UNDP (1999), OECD (2001, 2007), Hargittai (2003), Piskurich (2003), and Papastergiou and Solomonidou (2005) studied the existence of *access gaps*. In fact, all the divides and gaps have been observed to be interrelated in one way or the other. Using different approaches toward teaching and learning, both pedagogical (e.g., literacy teaching) and organizational (e.g., class size, ratio of teaching assistant to pupils) can help achieve positive outcomes and narrow the gap (Demiralay & Karadeniz, 2008; DCSF, 2007a; Cassen and Kingdon, 2007; Younger et al., 2005). Effective literacy interventions are an important element to narrowing the gap in the outcomes, as poor literacy at primary-school age is strongly and significantly associated with future low achievement (Cassen and Kingdon, 2007). With respect to these factors, earlier studies (Clarke et al, 2008; Ahmed, 2007; Souter, 2007; Bhanji, 2008) cited that the most-accepted gaps are Internet gaps, age gaps, digital gaps, knowledge gaps, access gaps, economic gaps, and performance gaps. Although there have been several studies (Tezer, & Bicen, 2008; Gunga & Ricketts, 2007; Cole, 2005, Manette, 2004) in the literature about gaps among students, there has been no study on the gaps among the teachers. Thus, this study attempts to examine the digital divide among the teachers along the four axes: age, Internet access, computer access, and performance (computing knowledge/experience) (see Figure 1).

Divides in the Digital World

Digital divide is used to describe the increasing gap between computer users and non-users (Becker, 2000). Digital divide is the gap between individuals, households, business, and geographic areas at different socioeconomic levels with regard to both their opportunities to access Information and Communication Technology (ICT) and the use of Internet for a wide variety of purposes (OECD, 2001). It is the division of the world between those who have access to new ICT and those who do not (Asian Development Bank, 2002). Digital divide is a term increasingly used to describe the social implications of unequal access to ICT by some sectors of the community, and to the acquisition of necessary skills (National Office for the Information Economy, 2001). It is the concrete and symbolic distance between those who enjoy the access and familiarity of the immense potential of technology, and those who do not (Munoz, 2002). Furthermore, it is the separation between those who have access to and can effectively use technology and those who could not (Pearson, 2002). Occurrence of a digital divide may be owing to several reasons, such as income, educational level, class, gender, race, and geographical location (Norris, 2001). As a result of access gaps, there also exist digital divides between (and within) rich and poor countries [Chinn & Fairlie, 2004; Hargittai, 2003]. Does this divide exist among teachers? Does age gap, Internet-access gap, computer-access gap, and performance gap result in divide in the digital world of teachers? Is there a convergence or a divergence? These are the questions that arise during the literature reviews, which must be addressed.

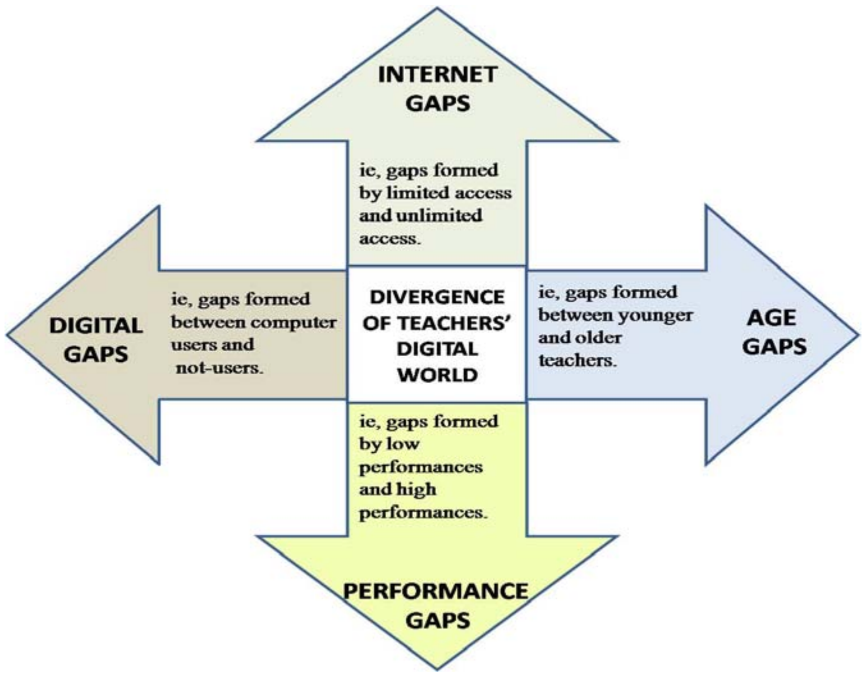

Figure 1. Divergence of teachers' digital world

Researches about Gaps in Education

There have been numerous researches on gaps and divides in education. However, most of them focused on digital divide among students. The European Centre for Development of Vocational Training (2001) carried out a survey among 446 individuals from several European countries, where one of the principle observations of this survey was that the trainers and vocational teachers undertake inadequate ICT skills development to improve the expertise in pedagogy and management issues, resulting in divergence in education. Uzunboylu (2006) observed technological, pedagogical, social, economic, and cultural barriers in education. According to Park and Ertmer (2008), barriers in education exist owing to lack of knowledge and skills, unclear expectations, and insufficient feedbacks. Although most of the researchers explained these factors in different terms, their research results revealed the existence of gaps in education. There exists a research gap with regard to studies focusing on digital world of teachers. Is there a divergence in the digital world of teachers which affects their ability to prepare students for the knowledge society?

Purpose of the Study

The purpose of this study is to determine the divergence of teachers in the digital world. In addition, the four sub-questions of this research are as follows:

1. Is there an Internet gap among teachers?
2. Is there a digital gap among teachers?
3. Is there an age gap among teachers?
4. Is there a performance gap among teachers?

Method

Population

The population of this research consisted of all the teaching staffs in 12 vocational schools in North Cyprus: teacher assistants, headmasters, headmaster assistants, area heads, curriculum managers, electric/electronics teachers, ICT teachers, motor teachers, machine technology teachers, accounting teachers, mathematics teachers, science teachers, history teachers, geography teachers, Turkish teachers, social science teachers, English teachers, metal teachers, etc. (throughout this paper, teaching staffs are simply referred as teachers). There were about 490 teachers in the vocational high schools in North Cyprus, and the questionnaires were given to all the teachers in the vocational schools. The response rate was 81.2%, with 396 valid and 2 incompletely answered questionnaires.

Instrument

A questionnaire was developed to examine the e-learning training needs of the vocational high-school teachers. To evaluate the items in the questionnaire, experts' evaluation (n = 17) was employed. An expert group of instructional technologists evaluated the data-gathering scale, both individually and collaboratively. Under the suggestions of the experts, necessary corrections were made to the draft form of the questionnaire. Thus, the content validity was maintained with the help of the educational technologist experts. After making the necessary corrections on the draft form of the questionnaire based on the experts' suggestions, the questionnaire was given to the teachers to fill, using a five-point Likert scale. In total, the questionnaire consisted of two parts: Part 1 contained the questions like:

- How long have you been working in the school?
- Where can you access to internet?
- What is your specification?

Part 2 consisted of 32 items, where the teachers were asked to choose the suitable scales for themselves, in each item of the questionnaire The scales were arranged as: needs to be improved, basic, good, very good, and excellent.. Among the main subjects measured in the Part 2 of the questionnaire were:

- Skill of publishing/visual tools and technology;
- Skill of using wireless technologies (phone, laptop);
- Ability to join video-conferences through internet.

After considering the experts' suggestions, the questionnaire was divided into three factors: technology-based e-learning applications (r=0.96), web-based e-learning applications (r=0.97), and administrative e-learning applications (r=0.97).

Process of Data Collection

The researchers visited all the vocational schools in North Cyprus and the questionnaires were given to all the teachers working in these schools, which was not an easy job, and demanded lots of traveling and explanations. As there were many part-time teachers in the vocational schools who teach on different days, each school was visited more than once, on different days. There were also other obstacles in the beginning of the research—some teachers did not want to fill the questionnaire as they were afraid of revealing their lack of necessary e-learning skill/knowledge for teaching, while some did not know the meanings of the terms, such as e-learning, e-TV, e-board, and e-discussion. The researchers explained each of the questionnaire items to the teachers and provided the necessary feedback needed for analyzing their own skills or knowledge as excellent, very good, good, basic, and needs to be improved. Thus, the maximum confidential atmosphere was provided to the participants.

Data Analysis

For data analysis, SPSS 16.0 was used. The questionnaire items were arranged according to the Internet gap, performance gap, digital gap, and age gap. The access to the Internet by teachers formed the basis of Internet gap. The total score of the questionnaire items about e-learning skills and ability was used as the performance level. The total score of the questions that were about ICT knowledge were used as the digital level. On the other hand, to reveal the age gap, the teachers who have more than 20 years of teaching experience were classified as older and those having less than 20 years of teaching experience were classified as younger. The teachers' accesses were classified as: "1" indicating limited and "2" indicating unlimited. Subsequently, the Internet score was calculated as the sum of the total score of the web-based applications and access score. Thus, the scores above this sum were considered to indicate narrow Internet gap and those below this total were considered to signify wide Internet gap. Furthermore, the gap score was considered to be the total of the four gap levels (Internet gap, digital gap, performance gap, and age gap). The score below the gap score mean was considered as "narrow gap" and the gap score above the mean was considered as "wide gap." The marks above the mean of the total score obtained from the questionnaire were regarded as high, and these teachers were accepted to be "e-literate." On the other hand, the marks below the mean of the total score obtained from the questionnaire were considered as low, and these teachers were accepted to be "not e-literate." The independent sample test was used to analyze all the variables, with 0.05 as the significant level. However, in all the tests, the resulting significant level was < 0.01.

Results and Discussions

The results and discussions will be explained in five sub-sections. The results obtained show the existence of digital gap, access gap, performance gap, and age gap among the teachers. In addition, they also demonstrate the effects of these gaps on the e-learning applications of the teachers, and are really important results showing the divergence of teachers' digital world.

Table 1. Descriptive statistics results of gap's effect on e-learning applications

	Gap Levels	***N***	***M***	**SD**	**Std. Error Mean**
Tech-based	Wide	269	22.96	8.61	0.53
	Narrow	129	40.66	10.44	0.92
Web-based	Wide	269	25.87	9.03	0.55
	Narrow	129	48.10	12.67	1.12
Admin-based	Wide	269	11.29	4.75	0.29
	Narrow	129	22.49	7.41	0.65

Furthermore, the effects of the gap score levels were also investigated (see Table 1**.** In technology-based e-learning applications, the number of teachers who were observed to have wide gaps (*M* = 22.96, SD = 8.61) were 269, and the number of teachers who were observed to have narrow gaps (*M* =40.66, SD = 10.44) were 129. Although the number of teachers who had wide gaps were more than those who had narrow gaps, the difference in the means was really significant. The narrow the gap's total was, the higher was the technology-based e-learning application scores. With regard to the web-based e-learning applications, the number of teachers who had wide gaps (*M* = 25.87, SD = 9.03) were 269 and those demonstrating narrow gaps (*M* =48.10, SD = 11.29) were 129. Although the number of teachers who had wide gaps were more than those with narrow gaps, the difference in the means was really significant. The narrow the gap's total was, the higher were the web-based e-learning application scores. On the other hand, with respect to the administrative e-learning applications, the number of teachers who had wide gaps (*M* = 11.29, SD = 4.75) were 269 and those who had narrow gaps (*M* =22.49, SD = 7.41) were 129. Although the numbers of teachers who had wide gaps were more than those who showed narrow gaps, the difference in the means was again high. The result demonstrates that if the gap's total is narrow, the administrative e-learning application scores are higher. Above all, the number of wide gaps and narrow gaps were found to be the same in all the three e-learning applications. These also prove the consistent existence of the effect of these gaps in e-learning application scores. Furthermore, the divergence of the e-learning applications according to the gaps can be seen in Figure 2. Here, the teachers who have high performance, unlimited Internet access, who are computer users, and are younger are observed to have an upper hand with regard to e-learning applications. Furthermore, an independent sample *t*-test also demonstrated a significant difference between the groups ($p<0.01$). These results show that the gaps between the teachers are really an important problem in the education world. It is an urgent problem that needs the attention of the educators.

According to the research results (see Table 2), the number of teachers who are: both "e-literate" and have "low performance" is 14; both "e-literate" and have "high performance" is 286; both "e-literate" and have "limited

Internet access” is 99; both “e-literate” and have “unlimited Internet” access is 22; both “e-literate” and are “non-computer users” is 259; both “e-literate” and “older” is 100; and both “e-literate” and “younger” is 52. On the other hand; the number of teachers who are: both “not e-literate” and have “low performance” is 0; both “ not e-literate” and have “high performance” is 98; both “ not e-literate” and have “limited Internet access” is 165; both “not e-literate” and have “unlimited Internet” access is 29; both “ not-literate” and are “non-computer users” is 99; both “ not e-literate” and “older” is 27; and both “ not e-literate” and “younger” is 10.

Table 2. Descriptive statistics of Axes of Divergence

	Low performance	**High performance**	**Limited Internet Access**	**Unlimited Internet Access**	**Non-Computer Users**	**Computer Users**	**Older**	**Younger**
e-Literate	14	286	70	99	22	259	100	52
Not e-literate	0	98	165	29	99	27	222	10

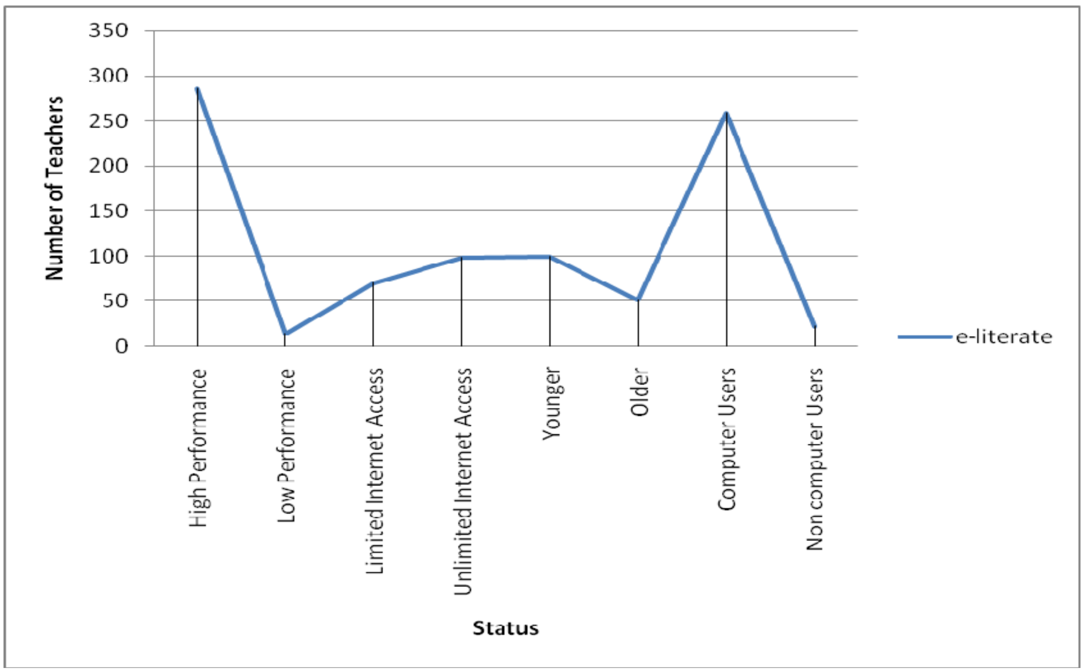

Figure 2. Discriminating Teachers

The teachers who have “high performance,” who are “computer users,” who have “limited Internet access,” and who are “older,” are those “discriminating digital world.” This case is illustrated in Figure 2.

It can be seen from Figure 2 that the low e-learning application score line is higher on the points of low performance, limited Internet access, and not users of computer, and older. However, the high e-learning application score line is higher on the points of high performance, unlimited Internet access, computer users, and younger. This proves the divergence of the digital world of teachers.

Internet Gaps of Teachers

The digital gap between low limited Internet access (M = 66.51, SD = 32.55) and unlimited Internet access (M = 87.50, SD = 31.25) is very much obvious in Table 3. Furthermore, an independent sample t-test also demonstrated a significant difference between the groups (p<0.01). This shows the divergence among the teachers, with respect to the limited and unlimited Internet access. Access to knowledge is clearly a fundamental requirement for development

(Cockerill & Knols, 2008). Hence, these results must be taken into consideration if a country wants to keep up with the speed of the digital world.

Table 3. Descriptive statistics of teachers' Internet access

Internet access	*N*	*M*	SD	Std. Error Mean	Std. Error Mean
Limited	205	66.51	66.51	32.55	2.27
Unlimited	193	87.50	87.5	31.25	2.25

Digital Gaps of Teachers

The digital gap between low ICT knowledge (*M* = 60.74, SD = 21.69) and high ICT knowledge (*M* = 115.47, SD = 24.52) is evident in Table 4. Furthermore, an independent sample *t*-test also showed a significant difference between the groups ($p<0.01$). This result demonstrates that the digital gap between the teachers is really an important problem in e-learning applications.

Table 4. Descriptive statistics of ICT-knowledge of teachers

ICT Knowledge Level	*N*	*M*	SD	Std. Error Mean
Low	282	60.74	21.69	1.29
High	116	115.47	24.52	2.28

Age Gaps of Teachers

The digital gap between the younger teachers (*M* = 80.60, SD = 33.37) and high ICT knowledge (*M* = 59.29, SD = 28.78) is evident in Table 5. Furthermore, an independent sample *t*-test also demonstrated a significant difference between the groups ($p<0.01$). These results show that if the teachers are more experienced, then their digital skills are limited. This may be because the younger teachers are more motivated, more in favor of ICT, or using more new technologies in their life.

Table 5. Descriptive statistics of teachers' ages in teaching

Age	*N*	*M*	SD	Std. Error Mean
Younger	325	80.60	33.37	1.85
Older	73	59.29	38.78	3.37

Performance Gaps of Teachers

The digital gap between low (*M* = 61.29, SD = 20.58) and high ICT knowledge (*M* = 123.84, SD = 18.13) is evident in Table 6. In addition, an independent sample *t*-test also demonstrated a significant difference between the groups ($p<0.01$), showing that a divergence exists in the teachers' world with respect to their low or high performance. Solutions to problems in the developing world depend on complete and effective collaboration between those working in the developed and the developing worlds.

Table 6. Descriptive statistics of performance levels of teachers

Performance Level	*N*	*M*	SD	Std. Error Mean
Low	300	61.29	20.58	1.19
High	98	123.84	18.13	1.83

Percentages of the Effects of gaps on Teachers' e-Learning Application Results

When the overall effect of these gaps on the e-learning application scores was calculated (see Figure 3), it was observed that the access gap (31%) had a greater effect on the e-learning scores. On the other hand, the age gap (26%) had the second greater effect on the e-learning scores. The third was the digital gap (19%) that had the lowest effect. Furthermore, several factors at the teachers' level were observed to influence the implementation of innovative ICT-use in education (Drent & Meelissen, 2008), and the existence of performance gaps, access gaps, digital gaps, and age gaps cannot be ignored in the digital world.

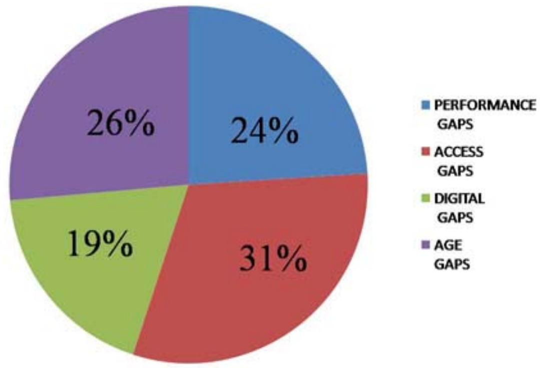

Figure 3. Percentages of gaps

Conclusions and Suggestions

The aim of this study was to find the divergence of teachers' digital world in North Cyprus. As the key point of education is teachers, their divergence in the digital world was explored. For this purpose, the four pathways (Internet gaps, age gaps, digital gaps, and performance gaps) were explored to determine their occurrence. It was observed that teachers who have low ICT skills also have low e-learning skills, which were proven to cause low teacher performance in digital technologies, in which teachers failures in that divergence in the digital world may be the most possible result (See Figure 5). The teachers must be able to prepare young people for the knowledge society in which the competency to use ICT to acquire and process information is very important (Ministry of Education, Culture and Science (MECS, 1999; Plomp et al., 1996).

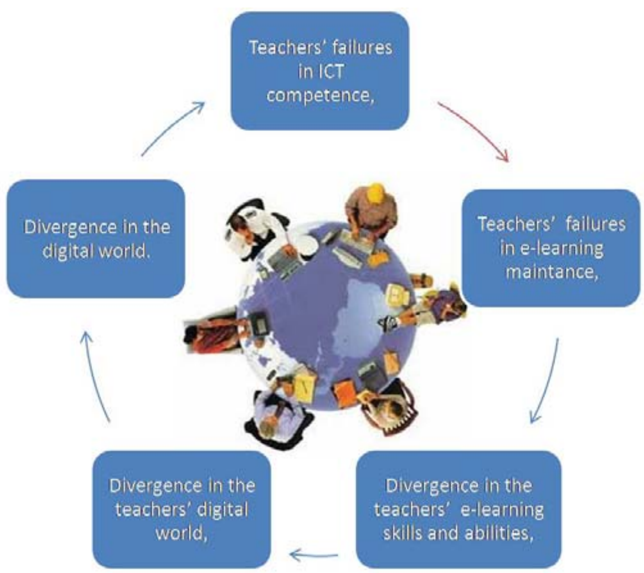

Figure 4. Circuit diagram of divergence in the world

Teachers are the key personnel in the integration of computers in instructional situations and in the adoption of all other innovations in schools. They are the key point in narrowing the gaps in education. There was a significant digital divergence observed among the teachers surveyed, which would adversely affect their ability to prepare students for the knowledge society. This study not only produces significant practical implications on the vocational high-school teachers' education in North Cyprus, but also contributes to the current literature related to ICT, e-learning, and digital divide. Teacher's digital world gaps should be taken seriously, not only in North Cyprus, but all over the world. As there has been no research on the divergence of the digital world of teachers, the current study is a unique attempt to address its limitations. It is found that there exists divergence among e-learning competences of teachers. It is found that if a teacher has high ICT performance, than he is e-literate. If a person is e-literate, his ICT performance level has not been investigated: This is another limitation of this study. Furthermore, an action plan could be devised for taking precautions in the digital education, comprising strategies for narrowing the Internet gap, age gap, digital gap, and performance gap. In addition, further investigations may be needed to determine the impact of gender gaps and economic gaps on the teachers' digital world.

References

Ahmed, A. (2007). Open access towards bridging the digital divide–policies and strategies for developing countries. *Information Technology for Development, 13*(4), 337-361.

Asian Development Bank. (2002). *Digital Divide: Determinants and policies with special reference to Asia,* ERD Working Paper Series No 27, Economic and Research Dept., Asian Development Bank, Philippines.

Attewell, P. (2001). The first and second digital divides. *Sociology of Education*, *74*, 252–259.

Becker, H. J., & Ravitz, J. L. (2001). Computer use by teachers: Are Cuban's predictions correct? *Paper presented at the 2001 annual meeting of the American Educational Research Association*, Seattle, WA, USA.

Becker, H. J. (2000). Who's wired and who's not: Children's access to and use of computer technology. *The Future of Children, 10*, 44-75.

Bhanji, Z. (2008). Transnational corporations in education: filling the governance gap through new social norms and market multilateralism? *Globalisation, Societies and Education, 6*(1), 55–73.

Chinn, M. & Fairlie, R. (2004). The determinants of the global digital divide: A cross- country analysis of computer and internet penetration. *National Bureau of Economic Research Working Paper,* No. 10686, Cambridge: Cambridge University Press.

Clarke, A., Milner, H., Killer, T. and Dixon, G. (2008). Bridging the digital divide. *Adults Learning, 20*(3), 20-22.

Cobb, L. (2002). Ethics and Reconciliation to Scarcity. *The 26th Annual Congress on Metropolisation in a Global Economy,* 23-26 June 2002, The Hague,The Netherlands.

Cockerill,M.J., & Knols,B.G.J.(2008). Open Access to research for the developing world. *Science and Technology, 24*(2), 65-69.

Cole, G. (2005). Bridge over the digital divide. *Times Educational Supplement*, 9/16/2005, 4652, 4-4.

Dalsgaard, S. (2001). Digital Discourses: An Essay on ICT in Development. Retrieved January 15, 2010, from http://www.allfreeessays.com/essays/Ict-Essay/4166.html.

Demiralay, R., & Karadeniz, S. (2008). Developing Information Literacy Skills for Lifelong Learning in Elementary Education. *Cypriot Journal of Educational Sciences, 2*(6), 89-119

Drent, M., & Meelissen, M.(2008). Which factors obstruct or stimulate teacher educators to use ICT innovatively? *Computers & Education, 51*, 187–199

Engelbrecht, J. H.(2008) Internet-based 'social sharing' as a new form of global production: The case of ETI@home. *Telematics and Informatics*, *25*(3), 156-168.

European Centre for Development of Vocational Training (2001). *E-Learning and the professional development of Trainers and Vocational Teachers*, Thessaloniki, Greece, Eric Documents: ED473658.

Gunga, S.O. & Ricketts, I.W. (2007). Facing the challenges of e-learning initiatives in African universities. *British Journal of Educational Technology, 38*(5), 896-906.

Hargittai, E. (2003). The digital divide and what to do about it. In Jones D. (Ed.), *New Economy Handbook* (pp. 821-839), San Diego: Elsevier/Academic Press.

Hawkins, B. L., & Oblinger, D. G. (2006). The myth about the digital divide. *Educause Review*, *41*(4), 12–13.

James, J. (2000). Pro-Poor Modes of Technical Integration into the Global Economy. *Development and Change*, *31*, 765-783.

Kim, M. C., & Kim, J. K. (2001). Digital divide: Conceptual discussions and prospect. Human Society and the Internet. *Lecture Notes in Computer Science, 2105*, 78–91.

Kong, S.C. (2008). A curriculum framework for implementing information technology in school education to foster information literacy. *Computers and Education, 51*(1), 129-141.

Korat, O., Bachar, E., & Snapir, M. (2003), Functional, social and cognitive aspects in emergent literacy: Relations to SES and to reading-writing acquisition in first grade, *Megamoth, 42*, 195–218.

Kozma, R., McGhee, R., Quellmalz, E., & Zalles, D. (2004). Closing the digital divide: Evaluation of the World Links Program. *International Journal of Educational Development 24*, 361–381

Lim, C. (2005). Development and effects of a Learning Management System for supporting self-regulated learning. *Journal of Educational Technology, 21*(4), 77-100.

Main, L. (2001). The Global Information Infrastructure: Empowerment or Imperialism? Through new social norms and market multilateralism? *Globalisation, Societies and Education,* 6(1), p.55–73.

Mairtin, O.F. (2001) *E-learning and Access: Some Issues and Implications.* Retrieved 15 January, 2010, from http://www.bath.ac.uk/iohm/fathaighr.rft.

Malloch, B. (2000).Commentary: The Internet and Development Choices*. The Third World Quarterly*, *22*(1), 83-97.

Manette, G. (2004). Leadership, money could carry college across the digitaldivide. *Tribal College Journal, 15*(4), p.26-27.

Ministry of Education & Culture and Science (1999). Maatwerk voor Morgen: het perspectief van een open arbeidsmarkt [Adjustments for tomorrow, perspectives on an open job market]. Retrieved 15 January, 2010, from http://www.minocw.nl/werkinonderwijs/publicaties.html.

Munoz, J. S. (2002). [Dis] Integrating Multiculturalism with Technology. *Multicultural Education, 10*, 19-24.

National Office for the Information Economy (NOIE) (2000) *Access and Equity*, Commonwealth of Australia. Retrieved January 20, 2010 from http://www.noie.gov.au.

Noce, A.A., & McKeown, L. (2008). A new benchmark for Internet use: A logistic modelling of factors influencing Internet use in Canada, 2005. *Government Information Quarterly*, *25*(3), 462-476.

Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the Internet worldwide.* New York: Cambridge University Press.

OECD (2007). *Broadband and ICT access and use by households and individuals*. Paris: OECD Publications.

OECD (2001). *Understanding the digital divide,* Paris: OECD Publications.

Papastergiou, M., & Solomonidou, C. (2005). Gender issues in Internet access and favorite Internet activities among Greek high school pupils inside and outside school. *Computers and Education*, *44*, 377–393.

Park, S.H., & Ertmer, P.A. (2008) Examining barriers in technology-enhanced problem-based learning: Using a performance support systems approach. *British Journal of Educational Technology*, 39(4), p.631-643.

Pearson, T. (2002). Falling behind: A technology crisis facing minority students. *Tech Trends, 46*, 15-20.

Peterson, M. (2008). Maps and the Internet. What a mess it is and how to fix it. *Cartographic Perspectives 59*, 4-11.

Piskurich, G. M. (2003). *Preparing Learners for e-LearnING*, Danvers, CA: Pfeiffer.

Plomp, Tj., ten Brummelhuis, A. C. A., & Rapmund, R. (1996). *Teaching and learning for the future. Report of the Committee on MultiMedia in Teacher Training (COMMITT)*, Den Haag: SDU.

Prevost, A.K., & Schaffner, B.F. (2008). Digital Divide or just another absentee ballot?: Evaluating internet voting in the 2004 Michigan democratic primary. *American Politics, 36*(4), 510-529.

Redd, T.M. (2003). Tryin to make a dolla out a fifteen cent: Teaching Composition with the Internet at an HBCU. *Computers and Composition*, *20*, 359–373.

Rice, R.E., & Katz, J.E (2008). Assessing new cell phone text and video services. *Telecommunications Policy, 32*(7), 455-467.

Singh, A.K., & Sahu, R.(2008). Integrating Internet, telephones, and all centres fro delivering better quality e-governance for all citizens*. Government Information Quarterly, 25*(3), 477-490.

Souter, D. (2007). Internet governance and development: Another digital divide? *Information Policy: The International Journal of Government & Democracy in the Information Age, 12*(1/2), 29-38.

Tezer, M., & Bicen, H. (2008). The Preparations University Teachers towards E-Education Systems. *Cypriot Journal of Educational Sciences, 3*(5), 16-27.

UNDP (1999). New technologies and the global race for knowledge. In *Human Development Report 1999*, UNDP, 57-76.

Uzunboylu, H. (2006). A Review of Two Mainline E-Learning Projects in the European. *Educational Technology Research and Development*, *54*(2), 201-209.

Vogelwiesche, U., Grob, A., &Winkler, B. (2006). Improving computer skills of socially disadvantaged adolescents: Same-age versus cross-age tutoring. *Learning and Instruction, 16*(3), 241-255.

Youssef, A.B. (2004). *Four dimensions of the digital divide* (in French), *Réseaux, 128*, 181-209.

Yang, J. C., & Lin, Y. L. (2010). Development and Evaluation of an Interactive Mobile Learning Environment with Shared Display Groupware. *Educational Technology & Society*, *13* (1), 195–207.

Development and Evaluation of an Interactive Mobile Learning Environment with Shared Display Groupware

Jie Chi Yang and Yi Lung Lin
Graduate Institute of Network Learning Technology, National Central University, Jhongli City 320, Taiwan // yang@cl.ncu.edu.tw // nigel@cl.ncu.edu.tw

ABSTRACT

When using mobile devices in support of learning activities, students gain mobility, but problems arise when group members share information. The small size of the mobile device screen becomes problematic when it is being used by two or more students to share and exchange information. This problem affects interactions among group members. To overcome the information sharing problem, the concept of Shared Display Groupware (SDG) has been proposed to support face-to-face collaboration using a shared display. However, little attention had been paid on the integration of the shared display with mobile devices in order to design a learning activity in the mobile learning environment. In this study, a learning activity was designed and a mobile learning environment was developed with the integration of the SDG to permit students to share information from individual and public spaces. During the learning activity, each student performed individual tasks using a PDA. Group tasks, following the individual tasks, were performed using a shared display, thus facilitating the sharing of information and group discussions. Each group was given their own display to share. To evaluate students' perceptions and learning effectiveness regarding the use of the SDG in supporting mobile learning, an empirical study was conducted. The study included a survey questionnaire as well as a learning achievement test. The participants in the experiment included thirty-four fourth-grade students and followed a one-group pretest-posttest design. The results show that the participants evaluated high scores in every category of the questionnaire. Significant differences were found between pretest and posttest in most aspects of the learning achievement test on the creation of conditions for classifying plants.

Keywords

Shared Display Groupware, Single Display Groupware, Mobile learning, One-to-one technology enhanced learning, Mobile devices, Science learning, Plant classification, Interactive learning environments

Introduction

Various mobile devices have been used in mobile learning, such as wrist-worn devices, mobile phones, handheld computers, web pads, pen tablet computers and laptop computers (Sharples & Beale, 2003). Many studies have reported achievements in the investigation of learning interests and the effectiveness of mobile learning (Rieger & Gay, 1997; Roschelle, 2003; Tatar, Roschelle, Vahey & Penuel, 2003; Zurita & Nussbaum, 2004). For an effective integration of mobile learning into a digital classroom environment, it is important for all students in a group to have their own computing device equipped with wireless communication capability to conduct learning tasks (Chan *et al*., 2006; Liang *et al*., 2005; Soloway *et al*., 2001). However, through the observation of learning activities with students using mobile devices for collaborative learning, some problems exist. For example, the experience of using Personal Digital Assistants (PDAs) as learning tools shows that it was difficult while sharing information with group members using mobile devices during the learning activity. When learners discuss and share information in a collaborative learning setting using mobile devices such as PDAs, screen size may have a negative influence on the learning activity due to the limitation of the screen size of the PDA (Magerkurth & Tandler, 2002). It is difficult to let all group members look at the same screen on a PDA simultaneously. To help other group members get access to the information simultaneously during the discussion, the learner may interrupt the task which still in progress, wait for other group members until they finished watching, and then continue the learner's own original task. Therefore, a shared display is needed, which group members can access, thus creating a common focus, facilitating group discussion and sharing of information without interruption.

The concept for supporting collaborative work between people via a shared computer with a single shared display was proposed by Stewart, Bederson, & Druin (1999). Subsequently, many studies followed this concept to develop environments or tools which support face-to-face collaboration, working with a shared display, in which all participants have their own input device (Ryall, Forlines, Shen, & Morris, 2004; Tse & Greenberg, 2004; Zanella & Greenberg, 2001). Most of the above-mentioned studies focused on the issues of the system interface design, such as the design of the screen size, transparent interface components, or support of multiple input devices. However, little attention focused on the integration of the shared display with mobile devices in order to design a learning activity in the mobile learning environment.

The aim of this study, based on the above considerations, was three-fold. Firstly, a learning activity was designed with the integration of the concept of Shared Display Groupware (SDG) to support mobile learning. Secondly, a mobile learning environment was developed under the design of a learning activity. The learning environment permits students to share and exchange information from individual and public spaces and allows each group member to continue with individual tasks in this environment without interruption of other simultaneous group tasks. The learning activity was integrated into a course based on the observation of plants which includes descriptions of plants and the creation of conditions for the classification of plants through the support of the SDG together with mobile devices. Finally, an empirical study was conducted to evaluate students' overall perceptions and the learning effectiveness on the use of the SDG in the support of mobile learning.

Literature review

Mobile learning

With the rapid development of network communication technologies, more and more wireless and mobile technology applications are integrated into classrooms to support teaching and learning. A study integrated wireless network technology with mobile devices including an e-whiteboard to build a Wireless Technology Enhanced Classroom (WiTEC) which provides various features for supporting teaching and learning in the classroom (Liu *et al*., 2003). For example, the feature *reducing the time for tedious work* allows the teacher to readily select materials and present or broadcast them to students, as well as mark and revise students' tasks through the use of e-whiteboard and mobile devices that enable numerous tedious tasks to be completed instantly. Another feature is *engaging students in learning activities* which indicates students engaging in learning activities such as exploring and organizing online course-related resources, as well as answering quizzes by means of group discussion using their mobile devices. In addition, students are not only able to discuss course work with each other face-to-face, but are also able to exchange personal materials through the mobile devices and the process of interaction can be recorded, thus *facilitating group collaborative learning*. Another feature is *empowering the teacher to monitor students' learning progress* which means a number-signal is provided to each mobile device that represents different statuses in using a mobile device such as disconnected to the server or request for help, therefore the teacher can monitor students' learning progress and determine how to implement the subsequent activities. Moreover, the *recording teaching and learning processes as portfolios* feature means the establishment of teaching records and learning portfolios and then promotes teachers' reflection on teaching as well as learners' portfolio assessment. Finally, the user-friendly interface by providing a handwriting function and the interactive classroom server effectively coordinate all system works to allow the teacher and students benefit from technology easily for *implementing technology-supported activities smoothly*.

In a Mobile Computer Supported Collaborative Learning activity, students engaged in collaborative learning through face-to-face communication on a social network with the support of handheld devices by a wireless network (Cortez *et al*., 2004; Zurita & Nussbaum, 2004). Many studies have demonstrated successful experiments which help students exchange information through PDAs as well as providing opportunities to interact with each other by using the PDAs as handheld devices for supporting learning. For example, *ad hoc* networks and mobile classrooms using PDAs in a wireless environment (Chang, Sheu & Chan, 2003), mobile learning systems for supporting outdoor learning about bird and butterfly watching (Chen, Kao & Sheu, 2003; Chen *et al*., 2004), and improving knowledge creation during experiential learning by mobile technologies (Lai, Yang, Chen, Ho & Chan, 2007).

Although the above-mentioned studies demonstrate positive results on the integration of mobile devices in support of collaborative learning, it is difficult to share information through mobile devices for group discussion. When two learners discuss and share information in a collaborative learning setting, they may not encounter any problems as they can transmit data using the built-in infrared light. However, this does not apply to multi-user transmissions among groups with three or more members (Danesh, Inkpen, Lau, Shu & Booth, 2001). In addition, students usually adopt an attitude of watching information displayed on the horizontal PDA which held by sharing the screen with other students who stand shoulder to shoulder or stand facing from the opposite direction (Chang, Sheu & Chan, 2003; Chen *et al*., 2004; Danesh *et al*., 2001). Students can also reverse the PDA screen to point the screen toward to the people who wants to see. Although the above-mentioned solutions can solve the problem of difficulties in sharing information on mobile devices, the action of information sharing may interfere with the learner's original task. While the learner is sharing information via his mobile device, showing data to group members for discussion for example, he is unable to use his own mobile device to continue the task and has to wait until the process has been completed. Limitations due to the small screen size of mobile devices, make it necessary to provide group members with a shared display to facilitate group discussion with a common focus.

Shared Display Groupware

In the development of SDG, originally SDG meant Single Display Groupware, where the purpose was to design a computer system that can support face-to-face collaborative interactions on a single display (Stewart, Bederson & Druin, 1999). Afterwards, the definition of SDG was extended to denote Shared Display Groupware (Ryall *et al*., 2004). Because of the popularization of the hardware, a class can be divided into several groups, each of which has its own SDG. The SDG can be seen as a groupware system, where one screen receives input from multiple devices, such as the mouse and keyboard, or adopts a touch screen environment to permit multi-user operation concurrently (Nicol & MacLeod, 2005; Scott, Mandryk & Inkpen, 2003). Another extended application is a multi-screen environment, which is called Distributed Display Environments (DDEs) (Inkpen & Mandryk, 2005). DDEs can be defined as one computer system which can have more than one physical display, such as applications of multiple-monitor desktop systems, rooms with networked projectors and displays, or campus-wide connected information display systems (Hutchings, Stasko, & Czerwinski, 2005).

In previous studies on the SDG, task arrangement was only available for the group's task, thus that all group members are working at the same time through the SDG to complete the common task of the group (Myers, Stiel & Gargiulo, 1998; Scott, Grant & Mandryk, 2003). Even though both the individual task and the group task were taken into consideration, the design was that after the individual task was completed in the private space of the SDG, the learner then worked in the public space of the SDG on the group task (Baudisch, Good, Bellotti, & Schraedley, 2002). This design is limited because the individual task affects the group task, as they are working on the same display simultaneously. Thus, the individual task was not genuinely performed on a personal device. The mobile device is better complemented with the SDG, thus the individual tasks should be completed on the mobile device, and the related group tasks should be completed on the SDG. Solutions to this problem were investigated to find out how people move from individual to group work through the use of both mobile devices and a shared public display (Greenberg, Boyle, & LaBerge, 1999; Kitamura, Osawa, Yamaguchi, Takemura, & Kishino, 2005; Liu & Kao, 2007). It indicated that users not only use shared display space as public space, but also they have their own personal display space as individual space. Therefore, such a system supports both group tasks such as discussions and individual tasks such as personal operations. Consequently, the individual tasks should not interfere with the group tasks and then could co-exist.

Table 1. Learning activity design

Phase	Step	Learning activity
1	1	Grouping and login
	2	Get clues
	3	Investigate and collect
	4	Compare and integrate clues
2	5	Second phase grouping
	6	Create conditions for classification
	7	Demonstrate results of classification

Learning activity design and system implementation

Design of learning activity

This study adopted a two-phase learning activity design. The first phase focuses on the observation of plants, in which students have to investigate various plants through characteristics of plants that assigned as clues to students. The second phase focuses on the classification of plants, in which students have to create specific conditions and then use those conditions to classify the plants. There are seven steps in the two phases, as shown in Table 1. The group members in the second phase differ from those in the first phase, which were formed by dividing original group members into newly formed groups.

Figure 1. Screenshots of the login on the PDA (left), and the login messages on the SDG (right)

During the learning activity, there are two types of tasks – individual tasks and group tasks. The individual task is mainly performed on each student's PDA. Students work on their own PDA to investigate plants that matched the assigned clues. The group task is performed on the SDG. Group members work together sharing information or discussing assigned tasks. The individual task is also performed on the SDG because the SDG is designed as a tool to promote information sharing and discussion but not limited as to interrupt individual tasks. Detailed explanations for each step of the learning activity are described as follows:

- Step 1: Grouping and login
  Individual task: Students login to the system using their PDAs (left side of Figure 1).
  Group task: Students find out their login messages on the SDG to confirm which group they belong to (right side of Figure 1), and meet all members in their group.

Figure 2. Screenshots of received clues (left), and the taking of photos and notes on the PDA (right)

- Step 2: Get clues
  Individual task: All students get individual clues on their PDAs (left side of Figure 2). Clues were created from different parts of plants for representing the characteristics of plants, such as "spiral shaped leaves" or "white flowers".
  Group task: Every clue is different for all group members. The information of the clues of the group is shown on the SDG.
- Step 3: Investigate and collect

Individual task: Students use their clues to investigate and observe different characteristics of plants. During the process, students take photos of the plants that fit the description found in the clues. They also take notes on their PDAs about the plants, as well as the locations and names of the plants observed (right side of Figure 2).
Group task: Group members collect information about various plants according to their clues. The task is achieved by either individually or collaboratively working with group members.

- Steps 4: Compare and integrate clues
Individual task: Students use their PDAs to interact with the group's SDG to share individual collected data on the shared display through the wireless network.
Group task: The left side of Figure 3 shows a screenshot of the SDG for comparing clues. The six blocks on the two sides represent individual spaces of six students in a group, whereas the two blocks in the middle represent public spaces of the group. Students can look all clues of their group members on the SDG. They can compare these clues with their collected data (annotation 1 and 2), and select one answer for uploading to the public space (annotation 3).Group members can explain their views on their answers to other group members for further discussion. They can also refer to materials on various plants (annotation 4). Group members should agree upon the final answer of the group by using the "vote function" which is provided in the system (right side of Figure 3, annotation 5).

Figure 3. Screenshots of clue comparisons (left), and voting (right) on the SDG
(1) The individual space displays collected data with enlarged photo
(2) The individual space displays collected data with photo and notation on the plant
(3) The public space displays students' answers with clues and photos
(4) Referred data of plants
(5) The vote function for group decisions

- Steps 5: Second phase grouping
Individual task: During the second phase, new groups were formed, different from the group of the first phase. Thus, each group member has the all clues of the plants from the first phase.
Group task: The task in the second phase is prompted in the SDG environment on completion of the second phase grouping.
- Step 6: Create conditions for classification
Individual task: Group members explain details of the plants that they are familiar with to their new group members. On completion of the tasks in the first phase, the students have a certain degree of understanding of all clues that they possess. Therefore, students can filter and choose clues that help the classification of plants which are worth sharing to all group members in creating conditions for classifying plants.
Group task: Group members should use the characteristics of plants to collaboratively create conditions for the classification of plants (left side of Figure 4), and use these created conditions to classify plants by using the method of Linnaeus' Binominal Nomenclature (Linnaeus, 1753) (middle and right sides of Figure 4, annotation 1 and 2). These plants are the target plants from the first phase. Group members should also vote on the created conditions, and determine which condition is better for classifying plants.
- Step 7: Demonstrate results of classification
Individual task: All students explain why they created the conditions for classifying the plants.

Group task: Students view the achievements made by other groups (Figure 5), as well as listening to teacher comments.

Figure 4. Screenshots of creating conditions for the classification of plants on the PDA (left), classifying plants on the PDA (middle), and classifying plants on the SDG (right)
(1) Individual space displays created conditions for classifying plants
(2) Public space displays the progress of classifying plants using the created conditions

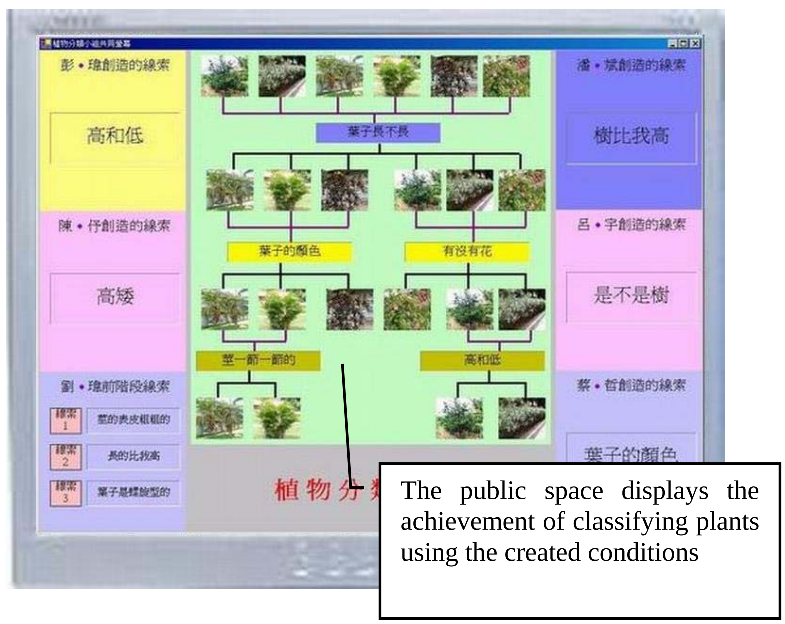

Figure 5. Completion of classifying plants

System implementation

The system can be divided into three parts from the viewpoint of the different hardware used: the server, the SDG, and the client (PDA). A system diagram is shown in Figure 6. The system functional modules are developed according to the three parts of the system. The server consists of a login module and a learning portfolio module. The SDG consists of a command receiving module and a transfer module. The client consists of a transfer module and a learning flow control module.

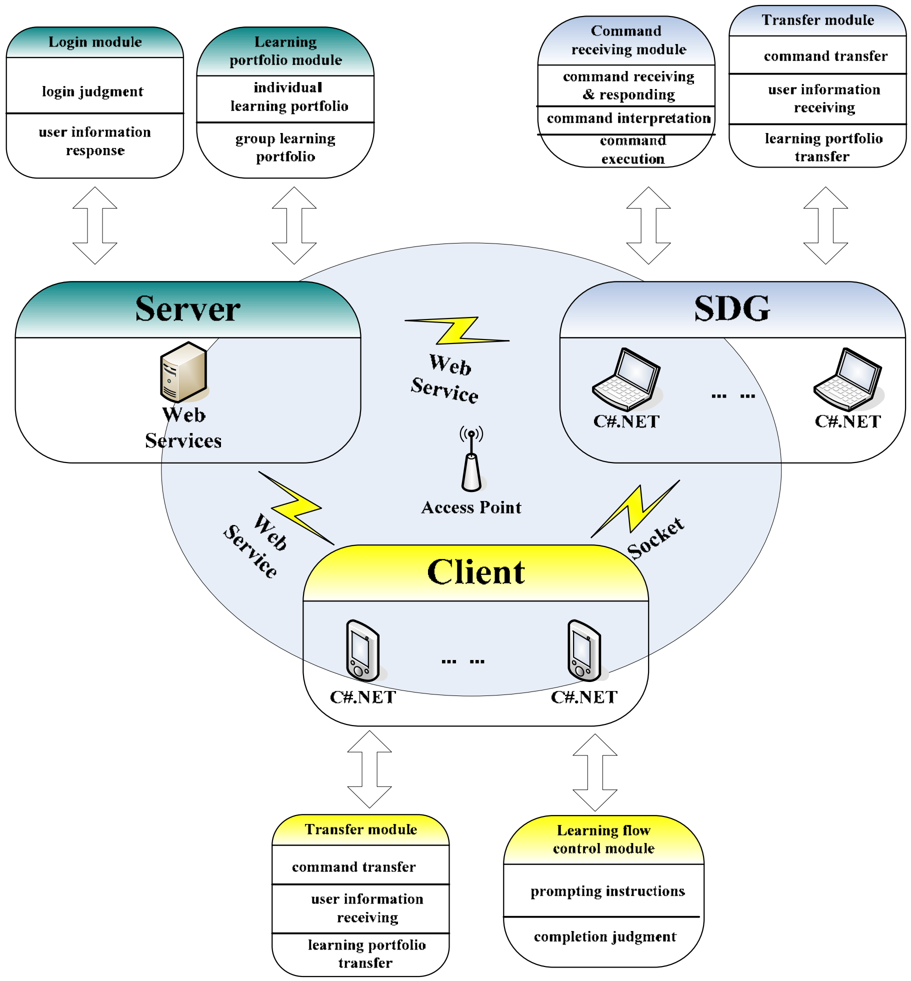

Figure 6. System diagram

Login module

The login module includes two functions: the login judgment and user information response functions. When the server receives a login request from a client, it checks the authorization for login to the system. If the authorization is approved, then user information containing the students' clues, group information, and IP address of the shared display will be transferred to the client. The user information can be used for indentifying groups for transmitting data among the server, clients and SDGs.

Learning portfolio module

The learning portfolio module includes two functions: an individual learning portfolio, and a group learning portfolio. For example, the students' notes in the first phase and created conditions for classifying plants in the second phase are recorded as an individual learning portfolio, whereas the answers of group decision in the first phase and the results of plant classifications in the second phase are recorded as a group learning portfolio.

Command receiving module

The command receiving module includes three functions: command receiving and responding, command interpretation, and command execution. The command receiving and responding function is responsible for receiving commands from clients, such as commands for enlarging photos, displaying notes, and voting for group decisions.

The types of commands are judged by the command interpretation function. Corresponding functions are executed on the SDG after receiving commands from clients by the command execution function.

Transfer module

The transfer module belongs to both the SDG and the client, and includes three functions: command transfer, user information receiving, and learning portfolio transfer functions. The command transfer function is responsible for the transmission of commands to the SDG from clients, such as uploading answers to the SDG. The user information receiving functions are responsible for receiving information from the server to the client, and from the client to the SDG, such as students' clues and group information. The learning portfolio transfer function is responsible for transferring the learning portfolio from SDGs to the server which contains individual and group learning portfolios.

Learning flow control module

The Learning flow control module includes two functions: prompting instructions for the learning activity, and judging the completeness of tasks. Clients are prompted with instructions on the tasks for each step of the learning activity. Students can refer to these instructions to complete the tasks. When a student has completed a task, this module is responsible for judging whether the student has completed the task well enough to begin the next step. If the task has not been satisfactorily completed, the student may not proceed to the next step.

The study

Based on the learning activity design with the support of the SDG and mobile devices, the objective of the empirical study was conducted to examine students' learning effectiveness on plant observations in terms of descriptions of plants and the creation of conditions for the classification of plants. Students' perceptions on the use of the SDG in the learning activity were also evaluated through a questionnaire survey.

Methods

The subjects of this study were fourth-grade students from an elementary school located in northern Taiwan. There were 34 subjects – 18 boys and 16 girls who participated in the study. Equipment used in the study included the PDAs, embedded mobile cameras and laptop computers as SDGs. The elementary school classroom was not equipped with any large screens for the participant groups to use, thus the screen of a laptop computer was used as the SDG.

Taking the subjects' former experience on PDA, mobile camera, and SDG environment usage into account, some training courses were given to the participants before the main experiment took place to avoid the "novelty effect" while facing a new form of media (Clark, 1983, 1994). The subjects were trained on the basic system functions for operating the PDA, mobile camera, as well as how to operate the PDA within the SDG environment. They were also familiarized with collaboration skills in the SDG environment. After the training courses, the main experiment took place and was comprised of four sessions, in total 180 minutes of the learning activity.

The instruments used in this study included a questionnaire and a learning achievement test. The questionnaire was designed to comprise nine categories, which included SDG function, reference materials, system message response, SDG equipment, collaboration, clue design, plant classification, learning attitude, and impression of the learning activity. The questionnaire consisted of 33 questions. Each question was evaluated on a five-point Likert scale (5 indicating strong agreement and 1 indicating strong disagreement).

The learning achievement test was conducted to examine the student's understanding of the description of plants and the creation of conditions for classifying plants through the support of the SDG with mobile devices. The test questions were designed by referring to the teacher's manual for the "Natural and Living Technology" course. The type of test questions consisted of two parts, being the descriptions of the plants and the creation of conditions for classifying the plants. The test on descriptions of the plants asked students to describe plants which cannot be found on the campus. Students had five minutes to answer the test. The test on the creation of conditions for classifying plants asked students to create conditions for classifying plants and classify the plants by Linnaean Binomial Nomenclature. Students had ten minutes to answer the test. The learning achievement test was designed as pretest and posttest, and the test questions used in the posttest were the same as in the pretest.

In the test on the descriptions of plants, the students answers were evaluated from four aspects, which included the use of the plant's six parts (roots, stems, leaves, flowers, fruits, and seeds), adjective words (for describing the appearance of plants such as color, size, quantity, shape), perceptual words (depending on personal affinity, for example, beautiful, ugly, good and bad), and the number of answered words (the total number of words in the answers). In the test on the creation of conditions for classifying plants, students answers were evaluated on five aspects, which included the use of the plant's parts, adjective words, perceptual words, successful classification (clearly classifying the plant into the A class and the non-A class successfully), and the number of created conditions for classifying plants.

Results of the questionnaire

Thirty-one valid copies of the questionnaire were collected after the experiment. Table 2 shows the results of the questionnaire with the mean scores for all questions in the nine categories. The results show that students evaluated high scores in each category of the questionnaire.

Table 2. Results of the questionnaire

Questions	M
SDG function	*4.32*
Sharing information with group members through the SDG was convenient.	4.32
Sharing photos and notes on the SDG facilitated group discussion.	4.29
The enlarged photos were clearly seen on the SDG.	4.16
Group discussions were aided through the individual and public spaces on the SDG.	4.29
The vote function on the SDG was useful for forming group decisions and agreements.	4.52
Reference materials	*4.23*
It was helpful to find the target plant by referring the materials on the SDG.	4.65
I referred to materials on the SDG to look up plants information during group discussion.	4.42
I referred to materials on the SDG for proof of my explanation to group members.	4.26
I referred to materials on the SDG to contradict group members' ideas.	3.58
System message response	*4.37*
I knew how to get to the next step by looking the system messages.	4.10
I understood the meaning of the system messages.	4.48
I carefully read the system messages.	4.52
SDG equipment	*3.80*
I think the screen size of the SDG was large enough for viewing information.	4.23
I can clearly watch information on the SDG.	3.36
Collaboration	*4.19*
I understood the group members' explanation during the process of group discussion.	4.16
I could work on my individual task without interrupting the process of the group task.	4.10
Group discussions were facilitated with a common focus on the SDG.	4.42
I collaborated with group members to observe plants using different clues.	4.03
I collaborated with group members to create and discuss conditions for classifying plants.	4.26
Clue design	*4.20*
I found many plants by using the assigned clue.	4.42
I collected plant information which matched the assigned clue.	4.32
The clue was clearly described for finding plants that matched the clue.	3.87
Plant classification	*4.08*
The clue used in the first phase was helpful for creating conditions for the classification of plants.	4.48
Group members working together to create conditions for classifying plants was meaningful.	4.00

Classifying plants with group members was an easy task.	3.77
Learning attitude	*4.38*
Group discussions were interesting through the use of the SDG.	4.52
Controlling both the PDA and the SDG were convenient.	4.25
The learning activity was fun like a game with clues to find plants.	4.51
I carefully read group members' data on the SDG.	4.25
Impression of the learning activity	*4.29*
I am willing to participate in the learning activity again.	4.28
I am willing to use the SDG again to share information with group members.	4.23
I am willing to use the SDG again to discuss information with group members.	4.42
I am willing to use the SDG again to create more conditions for classifying plants.	4.23

The results of the questionnaire show that SDG functions, such as information sharing, photos enlargement, individual and public spaces, and vote function were highly rated by the students. Students also positively rated the use of reference materials as well as the system message responses. Although the screen size of the laptop computer was small, most of the students stated that the screen size of the laptop computer was large enough for viewing information. In regards to collaboration supported by the SDG, students agreed that the SDG supported information sharing and group discussion, which they shared on the SDG including items such as photos and notes on the description of plants, as well as conditions for classifying plants from other group members. During the discussion, students did not interrupt the process of sharing information with group members while they completed individual tasks. In addition, students also highly rated the design of the clues and plant classifications, even though some students did not match plants with the clues or correctly classify plants with their group members. Moreover, the findings on students' learning attitude and impressions of the learning activity show that most students agreed that the learning activity is designed with the support of the SDG, and were willing to participate in the learning activity again.

Results of learning achievement test

Description of plants

In comparing the difference between the pretest and posttest, only a valid sample of twenty-seven students, who participated in both the pretest and posttest were used in the final analysis. A paired-samples *t* test was the statistical method adopted for use in the study. Table 3 shows the results of the test for the descriptions of plants. The results show that there was no significant difference between the pretest and posttest in all aspects of the test on the descriptions of plants.

Table 3. Results of the test on the descriptions of plants

Aspect	Average		p-value
	pretest	posttest	
Number of answered words	22.46	23.44	.721
Use of plant's parts	2.83	2.74	.183
Adjective words	3.42	3.56	.185
Perceptual words	0.70	0.52	.449

Creation of conditions for classifying plants

To compare the difference between the pretest and the posttest, only the valid sample of twenty-five students whose answers appeared in both the pretest and posttest were used in the analysis. A paired-samples *t* test was the statistical method adopted for use in the study. Table 4 shows the results of the test on the creation of conditions for classifying plants.

The results show that there were significant differences between pretest and posttest in most aspects of the test on the creation of conditions for classifying plants, except for the successful classification aspect. The number of created conditions improved from 4.20 (84%) in the pretest to 4.92 (98%) in the posttest, showing a significant difference (p=0.023). The number of plant's parts used in creating conditions had also greatly improved, from 1.80 (43%) in the pretest to 3.12 (63%) in the posttest, showing a significant difference (p=0.000). Regarding the number of adjective words used to create conditions for classifying plants, the number had increased from 1.76 (42%) in the pretest to 2.60 (53%) in the posttest, showing a significant difference (p=0.002). In contrast, the number of perceptual words used decreased from 1.04 (25%) in the pretest to 0.72 (13%) in the posttest, showing a significant difference (p=0.043).

Table 4. Results of the test on creation of conditions for classifying plants

Aspect	**Average**		**p-value**
	pretest	posttest	
Created conditions	4.20	4.92	0.023*
Use of plant's parts	1.80	3.12	0.000***
Adjective words	1.76	2.60	0.002**
Perceptual words	1.04	0.72	0.043*
Successful classification	1.12	1.48	0.095

*: $p<0.05$; **: $p<0.01$; ***: $p<0.001$

Discussions

The results of the questionnaire show that most students rated high scores in each category. Although some questions of the questionnaire were rated a little lower than the rest of the questions, they were still higher than the average score. The question that got the lowest score is the question asking students if they can clearly look at the information on the SDG in the category of "SDG equipment". A reasonable explanation for this result could be found from the observation of students' behaviors during the learning activity. Some students answered that they cannot clearly see information from certain viewpoints due to light reflections. However, it was found that students could easily adjust the position of the laptop to give them a better view of the screen. This indicates that the screen of laptop computer could be adapted as the SDG in the absence of a large-scale screen. This also demonstrates the feasibility of integrating the SDG into the classroom.

From the results of the test on the descriptions of plants, there was no significant difference between pretest and posttest in all aspects of the test. One of the possible reasons may be that the students only observed three kinds of parts of plants, such as the stems, leaves, and flowers. The other three parts of the plants (roots, fruits and seeds) were not observed due to limited time in the experiment. Therefore, students answered with few words in the test because they did not observe many parts of plants, causing no significant difference between pretest and posttest.

From the results of the test on the creation of conditions for classifying plants, there were significant differences between pretest and posttest in most aspects of the test. This indicates that students had improved their abilities to create conditions for classifying plants, and they became familiar with the plant's parts during the learning activity for observing plants and creating conditions for classifying plants. From a comparison of the pretest and the posttest, the increase in the amount of adjective words used and the reduction of subjective perceptual words used, reveal that students have learned a better way of classifying plants.

Conclusion

In this study, the authors integrated the Shared Display Groupware (SDG) concept with the use of mobile devices to design a learning activity and develop a mobile learning environment. The two-phase seven-step learning activity was designed to examine the characteristics of plants by means of observation and classification of plants. The SDG supports the simultaneous sharing and discussing of information among group members within the mobile learning environment. The system also supports the concurrent use of all individual and group tasks. Each task may be carried out without the interruption to any other ongoing tasks.

The results of the empirical study demonstrated that students positively evaluated on the questionnaire. For example, students reacted positively when surveyed on the convenience of the SDG functions for sharing information and creating a common focus during group discussions. Most students answered that they were willing to use the SDG again. In addition, the results of the learning achievement test demonstrated that students had improved their abilities to create conditions for classifying plants through the support of the SDG together with mobile devices under the learning activity design.

The empirical study described in this paper was only a small-scale study with sample consisting of one class and a one-group pretest-posttest design. This limitation may have influenced the validity of the results. Therefore, it is suggested that further studies be undertaken with a large sample, under a two-group experimental design to verify the results described in this paper. Another limitation is that the period of the study was short, which caused a lack time to observe plants and create sufficient conditions for the classification of plants. A significant amount of time was needed for students to familiarize themselves with operations of the SDG in group discussions. Therefore, a long-term experiment is required for examining students' learning effectiveness by using the SDG for the support of the learning activity. In addition, students' comments may also be gathered and analyzed for the improvement of the interface design of the SDG creating a more flexible design for use.

Acknowledgements

The authors would like to thank Mr. Aubrey Neil Leveridge for proofreading on an earlier version of this paper. The authors would also like to thank Mr. Ting Yen Lin for assisting in the system development. In addition, the authors wish to thank all the subjects who participated and cooperated in the experiment. This study was partially supported by grants (NSC 97-2628-S-008-001-MY3, NSC 97-2631-S-008-002) from the National Science Council of Taiwan.

References

Baudisch, P., Good, N., Bellotti, V., & Schraedley, P. (2002). Keeping things in context: A comparative evaluation of focus plus context screens, overviews, and zooming. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Minneapolis, Minnesota, USA. 259-266.

Chan, T. W., Roschelle, J., His, S., Kinshuk, Sharples, M., Brown, T., Patton, C., Cherniavsky, J., Pea, R., Norris, C., Soloway, E., Balacheff, N., Scardamalia, M., Dillenbourg, P., Looi, C. K., Milrad, M., & Hoope, U. (2006). One-to-one technology enhanced learning: an opportunity for global research collaboration. *Research and Practice in Technology Enhanced Learning, 1*(1), 3-29.

Chang, C. Y., Sheu, J. P., & Chan, T. W. (2003). Concept and design of Ad Hoc and Mobile classrooms. *Journal of Computer Assisted Learning, 19*(3), 336-346.

Chen Y. S., Kao T. C., & Sheu J. P. (2003). A mobile learning system for scaffolding bird watching learning. *Journal of Computer Assisted Learning, 19*(3), 347-359.

Chen, Y. S., Kao, T. C., Yu, G. J., & Sheu, J. P. (2004). A Mobile Butterfly-Watching Learning System for Supporting Independent Learning. *Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE 2004)*. Jhongli, Taiwan. 11-18.

Clark, R. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445-459.

Clark, R. (1994). Media will never influence learning. *Educational Technology Research and Development, 42*(2), 21-29.

Cortez, C., Nussbaum, M., Santelices, R. A., Rodríguez, P., Zurita, G., Correa, M., Cautivo, R. (2004). Teaching Science with Mobile Computer Supported Collaborative Learning (MCSCL). *Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE 2004)*. Jhongli, Taiwan. 67-74.

Danesh, A., Inkpen, K., Lau, F., Shu, K., & Booth. K. S. (2001). Geney: Designing a collaborative activity for the Palm handheld computer. *Proceedings of the ACM CHI 2001 Conference on Human Factors in Computing Systems*. Seattle, Washington, USA. 388-395.

Greenberg, S., Boyle, M., & LaBerge, J. (1999). PDAs and Shared Public Displays: Making Personal Information Public, and Public Information Personal. *Personal Technologies, 3*(1), 54-64.

Hutchings, D. R., Stasko, J., & Czerwinski, M. (2005). Distributed display environments. *Interactions, 12*(6), 50-53.

Inkpen, K., & Mandryk, R. L. (2005). Multi-Display Environment for Co-located Collaboration. *Proceedings of the ACM CHI 2005 Workshop on Distributed Display Environments*. Portland, Oregon, USA.

Kitamura, Y., Osawa, W., Yamaguchi, T., Takemura, H., & Kishino, F. (2005). A display table for strategic collaboration preserving private and public information. *Lecture Notes in Computer Science: Entertainment Computing, 3711*, 167-179.

Lai, C. H., Yang, J. C., Chen, F. C., Ho, C. W., & Chan, T. W. (2007). Affordances of Mobile Technologies for Experiential Learning: The Interplay of Technology and Pedagogical Practices. *Journal of Computer Assisted Learning, 23*(4), 326-337.

Liang, J. K., Liu, T. C., Wang, H. Y., Chang, L. J., Deng, Y. C., Yang, J. C., Chou, C. Y., Ko, H. W., Yang, S., & Chan, T. W. (2005). A Few Design Perspectives on One-on-one Digital Classroom Environment. *Journal of Computer Assisted Learning, 21*(3), 181-189.

Linnaeus, C. (1753). *Species Plantarum*. Stockholm, Sweden.

Liu, C. C., & Kao, L. C. (2007). Do handheld devices facilitate face-to-face collaboration? Handheld devices with large shared display groupware to facilitate group interactions. *Journal of Computer Assisted Learning, 23*(4), 285-299.

Liu, T. C., Wang, H. Y., Liang, J. K., Chan, T. W., Ko, H. W., & Yang, J. C. (2003). Wireless and mobile technologies to enhance teaching and learning. *Journal of Computer Assisted Learning, 19*(3), 371-382.

Magerkurth, C., & Tandler, P. (2002). Interactive Walls and Handheld Devices - Applications for a Smart Environment. *Proceedings of the UbiComp'02 Workshop on Collaboration with Interactive Walls and Tables*. Göteborg, Sweden.

Myers, B. A., Stiel, H., & Gargiulo, R. (1998). Collaboration Using Multiple PDAs Connected to a PC. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW'98)*. Seattle, Washington, USA. 285-294.

Nicol, D. J., & MacLeod, I. A. (2005).Using a Shared Workspace and Wireless Laptops to Improve Collaborative Project Learning in an Engineering Design Class. *Computers & Education, 44*(4), 459-475.

Rieger, R., & Gay, G. (1997). Using Mobile Computing to Enhance Field Study. *Proceedings of the Computer-Supported Collaborative Learning Conference (CSCL'97)*. Toronto, Canada, 215-223.

Roschelle J. (2003) Unlocking the learning value of wireless mobile devices. *Journal of Computer Assisted Learning, 19*(3), 260-272.

Ryall, K., Forlines, C., Shen, C., Morris, M. R. (2004). Exploring the Effects of Group Size and Table Size on Interactions with Tabletop Shared-Display Groupware. *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. Chicago, Illinois, USA. 284-293.

Scott, S. D., Mandryk, R. L., & Inkpen, K. (2003). Understanding Children's Collaborative Interactions in Shared Environments. *Journal of Computer Assisted Learning, 19*(2), 220-228.

Scott, S. D., Grant, K., & Mandryk, R. L. (2003). System Guidelines for Co-located Collaborative Work on a Tabletop Display. *Proceedings of Eighth Conference on European Conference on Computer Supported Cooperative Work (ECSCW 2003)*. Helsinki, Finland. 159-178.

Sharples, M., & Beale, R. (2003). A Technical Review of Mobile Computational Devices. *Journal of Computer Assisted Learning, 19*(3), 392-395.

Soloway, E., Norris, C., Blumenfeld, P., Fishman, B., Krajcik, J., & Marx, R. (2001). Log on Education: Handheld devices are ready-at-hand. *Communication of the ACM, 44*(6), 15-20.

Stewart, J., Bederson, B. B., & Druin, A. (1999). Single Display Groupware: A model for co-present collaboration. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the CHI is the limit*. Pittsburgh, Pennsylvania, USA. 286-293.

Tatar, D., Roschelle, J., Vahey, P., & Penuel, W. R. (2003). Handhelds Go to School. *IEEE Computer, 36*(9), 30-37.

Tse, E., & Greenberg, E. (2004). Rapidly prototyping Single Display Groupware through the SDGToolkit. *Proceedings of the fifth conference on Australasian user interface*. Dunedin, New Zealand. 101-110.

Zanella, A., & Greenberg, S. (2001). Reducing interference in single display groupware through transparency. *Proceedings of the seventh conference on European Conference on Computer Supported Cooperative Work*. Bonn, Germany. 339-358.

Zurita, G., & Nussbaum, M. (2004). Computer Supported Collaborative Learning Using Wirelessly Interconnected Handheld Computers. *Computers & Education, 42*(3), 289-314.

Vernadakis, N., Giannousi, M., Derri, V., Kellis, I., & Kioumourtzoglou, E. (2010). Athens 2004 Team Leaders' Attitudes toward the Educational Multimedia Application "Leonidas". *Educational Technology & Society, 13* (1), 208–219.

Athens 2004 Team Leaders' Attitudes toward the Educational Multimedia Application "Leonidas"

Nikolaos Vernadakis, Maria Giannousi, Vassiliki Derri, Iraklis Kellis and Efthimis Kioumourtzoglou
Department of Physical Education and Sport Science, Democritus University of Thrace, Greece // nvps@otenet.gr // mgiannou@phyed.duth.gr // vaderri@phyed.duth.gr // opkellis@yahoo.com // kioumour@phyed.duth.gr

ABSTRACT

The purpose of this study was to adapt the questionnaire Multimedia Attitude Survey (MAS; Garcia, 2001) to the Greek population in order to evaluate the educational multimedia application "Leonidas" considering the attitudes of ATHENS 2004 team leaders. In addition, the differences among the sex were also investigated. Participants were 232 team leaders, between the ages from 33-44 years old. One hundred twenty two (52.6%) of the participants were men and one hundred ten were women (47.4%). Data was collected using an on-line survey at the end of this study. Results from the factor analysis yielded eight factors accounting for 89.98% of the variance. Reliability analysis indicated a satisfactory internal consistency estimate of reliability for the attitude questionnaire. Independent-samples t test analysis revealed significant differences between the two sex groups, in the case of one factor: "general experience". In the factor above the women reported better results. In conclusion the team leaders' feedback from the questionnaires indicated a general level of satisfaction and contentment with this particular multimedia application. The scale adapted in the present study can be a useful tool for the evaluation of other relative multimedia applications by multimedia developers. Nevertheless, further examination is warranted in order to obtain additional information concerning the difficulties of multimedia experience on employees' attitudes toward multimedia applications.

Keywords

Multimedia application, Attitude, Olympic Games, Gender, Technology

Introduction

Technology offers a promising resource (via computer networks, distance learning systems, multimedia software, and video materials) for training staff and volunteers, sharing information about promising practices, and reducing the isolation of many programs. The new technologies offer ways of individualizing instruction to meet the needs of types of learners and potentially to reach all types of learners in ways they learn best.

So, is there any value added by using technology in adult education? As with so many other innovations, the value is not intrinsic, but rather depends on how and for what purposes one uses the innovation. Simply adding technology without challenging ourselves to do things we could not do before, or to do them differently, is meaningless at best, and very expensive at worst. On the other hand, technology applications and activities that lead to expanded opportunities for learning can only help adult learners acquire the skills and mastery of tools to support independent, lifelong learning.

Multimedia computer-assisted instruction (MCAI) is increasingly being used as a means of delivering educational content in organizations training. Efficiency, portability, consistency, and effectiveness have all been cited as reasons for employing this technology in the company's educational environment. These visual learning symbols, pictures, and other representative techniques allow students to go deeper into ideas and concepts (Chandler, 2003).

However, the rapid growth of multimedia implementation in learning settings does not guarantee participation and acceptance on the part of employees. Negative attitudes towards multimedia-based instruction could be a deterrent to using multimedia technology as a learning tool. Therefore, the thoughts, tendencies and attitudes of the learners' towards these tools are needed to be determined (Becker and Maunsaiyat, 2002; Christensen and Knezek, 2000; İşman and Dabaj, 2004; Selwyn, 1997).

Awareness of employees' attitudes toward MCAI is a critical criterion in the evaluation of multimedia courses and in the development of multimedia computer-assisted curricula. Attitudes toward multimedia-enhanced instruction are considered to influence not only the acceptance of this medium of instruction, but also future behaviors in the learning process. For this reason, the promotion and maintenance of positive attitudes toward MCAI is of paramount importance. Negative attitudes must not be allowed to limit the knowledge and creativity of learners, nor anxiety to interfere with the learning process. If the utilization of multimedia teaching/learning environments is to be maximized, attitudes toward these learning settings must be continuously monitored. Fast, effective instruments to assess attitudes toward multimedia instruction are crucial to this process.

There is a wealth of computer attitude scales available in the literature. Many instruments have been developed with the purpose of measuring computer anxiety, computer usage, computer appreciation, and other computer-related attitudes (Jones and Clarke, 1994; Kay, 1993; Selwyn, 1997). There are a number of studies which provide useful empirical comparisons of available computer attitude scales (Gardner, Discenza & Dukes, 1993; Woodrow, 1991). All the previous references to existing surveys focus on general attitudinal parameters rather than on in-depth attitude- related dimensions. Besides, all the surveys reviewed explore learners' attitudes towards computers, and none of them elicit students' perceptions toward multimedia instruction as such.

Garcia (2001) reported a practical, multi-dimensional, easy-to-administer research tool specifically intended to assess the attitude of learners towards multimedia-enhanced instruction. The specificity of this 25-item instrument constituted a powerful tool for the assessment of student attitudes towards multimedia technology when this was used for educational purposes. This instrument was tested by Garcia on 40 subjects. The internal reliability coefficient for each of the attitude sub-scales making up the survey - student attitudes on individualized instruction; student attitudes toward self-paced instruction; student attitudes toward the user-friendliness of the learning-environment; student levels of anxiety when working with multimedia; and the general opinion of the students toward their experience with the instructional material - showed a high degree of internal consistency. The independence of these subscales has allowed practitioners, evaluators and researchers to make their own selection of factors in order to adapt the survey to meet their own needs with an eye toward evaluating and predicting the performance of learners in a multimedia-enhanced learning setting.

Cardoso, Peralta & Costa (2005) examined the attitudes and the perceptions of the Portuguese students about educational multimedia software’s criteria of quality. Their study was part of an international project supported by EU (PEDACTICE - Educational Multimedia in Compulsory School: From Pedagogical Assessment to Product Assessment), and the sample of interviewed pupils can be considered as representative of the Lisbon schools, attended by teachers and pupils very much interested in multimedia materials. The results indicated the confirmation of the success of computers and multimedia among the young Portuguese student population, being manifest either in their attitudes or in the diversity of their experiences, including the technical mastery of informatics.

Teoh and Neo (2007) investigated students’ learning impact and attitudes towards independent learning and self-paced discovery. A set of multimedia tools were employed to create the student–centred learning environment and were designed using Gagne’s Nine Events of Instructions which provides a proper theoretical framework of a good instructional lesson plan. In general, this study has found that interactive learning using this Multimedia-based environment is feasible and is a viable alternative to the traditional classroom which has proved to be limited in achieving the necessary needs of the students in the modern learning context. Students were positive towards active learning and were confident in enforcing self-paced strategy. This is a viable learning strategy and should be encouraged by educationists.

Gandole, Khandewale & Mishra (2006) examined the effect of multimedia software support on the attitude towards electronics subject of students while working in laboratory of electronics science. The investigator developed an attitude scale having 25 items by covering various aspects related to electronics experiments and laboratory communication. There were 21 positive and four negative items on five-point scale (Likert type). The difference of points in pre test and posttest decided the change in Attitude. The findings showed that the multimedia software support used for laboratory communication was much effective in bringing an attitudinal change among the students. There was a remarkable enhancement in attitude for all items.

Numerous studies have indicated sex differences in computer attitudes whereby males hold more positive and less negative attitudes and females vice versa (e.g. Bebetsos, Kouli & Antoniou, 2007; Ho & Lee, 2001; Schumacher and Moharan-Martin, 2001). Equally there is evidence that denies a difference (e.g. Antoniou, Patsi, Bebetsos & Ifantidou, 2006) or conversely finding that females liked computers more than males (Keasar, Baruch & Grobgeld-Dahan, 2005). North and Noyes (2002) found that the impact of psychological gender (sex and sex-role), does not influence significantly attitudes or cognitions towards computers. This does not support the notion that a technological gender gap is developing, nor the literature that suggests males hold more positive attitudes and cognitions than females. However, in this instance it appears that the computer is viewed positively.

In general, the relationship between computer attitudes and gender is not straightforward. Although research implies that males hold more global positive computer attitudes, there is evidence that on certain aspects females view computers more favourably than males. This is in relation to elements of computer culture, on certain new technology items or when computers are presented as useful tools. In addition, it has been suggested that whilst males view computers as more appropriate for them, females regard computers as more gender neutral than males and do not regard mathematics ability as a prerequisite. This supports the finding that males hold more gender-stereotypes of computers than females (Sanders, 2006; Whitely, 1997).

Learners' attitudes have contributed to our understanding of why MCAI have enhanced achievement and performance and motivation. Multimedia applications are profit tools for individual and student-centered learning, so, in order to be informative, effective and attractive on their use, they should have a clear view of the students' attitudes on the use of multimedia. Learners are no doubt the most important stakeholders in what concerns the use of educational multimedia software. However they are seldom questioned about their interests, difficulties or suggestions on this matter (Cardoso, Peralta & Costa, 2005). That's why this study takes the learners' point of view about quality of educational multimedia software as its main concern.

The leading aim of this study was to adapt the Multimedia Attitude Survey (MAS; Garcia, 2001) to the Greek population in order to evaluate the educational multimedia application "Leonidas" considering the attitudes of ATHENS 2004 team leaders. In addition, the differences among the sex were also investigated. Some more specific objectives come out from the following three research questions:

1. Is there a single dimension or are multiple dimensions underlying the 25 attitude items toward the multimedia application?
2. How reliable is our 25-item measure of attitude control?
3. Does the average amount of students' attitude differ between males and females?

Methods

Participants

Participants in this study consisted of two hundred thirty two (*n*=232) employees who were enrolled in "Team Leaders" Training Program at Organizing Committee for the Olympic Games ATHENS 2004 during Spring 2004. One hundred twenty two (52.6%) of the participants were male and one hundred ten were female (47.4%), between the ages from 33-44 years old.

Instrumentation

Software instrument

The multimedia application "Leonidas" was developed by the education and training department to support the "Team Leader" Training Program of the Organizing Committee for the Olympic Games ATHENS 2004. The material was constructed using the following programs: a) Macromedia Flash MX 2004, b) Adobe in Design 2.1, c) Adobe Photoshop 7.0 and d) Adobe Premiere Pro. The multimedia application run under Windows and Mac personal computer systems and was divided into five theme groupings: 1. Introduction, 2. Leadership, 3. Leader & Team, 4. Team Leader Skills and 5. Audit & Evaluation (see figure 1).

In order to help the team leader implementing the required policies, providing members of his/her challenge of pressing Games-time conditions and contribute to their success, "Leonidas" included the use of a simple language, a host of interactive applications such as audio flash movies and video, a wealth of photographic and other high quality illustrative material.

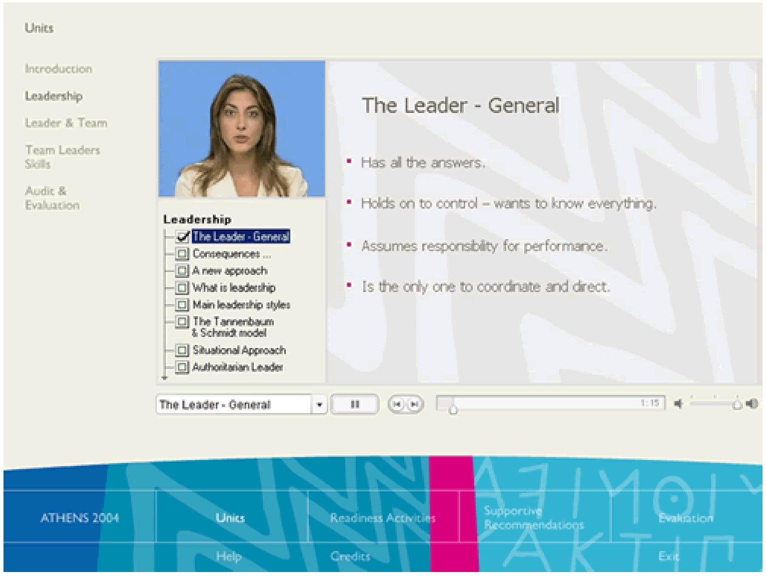

Figure 1. Example screen from the multimedia application "Leonidas"

The multimedia application consisted of 905 pages; 10 pages were introductory, 7 were main menus, 486 were information, 124 were practice, 243 were feedback and 35 were help. At the end of each topic and sometimes in certain sub-topics, a quiz was provided which contained 10 multiple-choice questions on the material. In this quiz learners were asked to resolve various readiness situations and giving responses to scenarios of potential problems. The users had the ability to navigate through the path structured by the programmer via the site map or from the menu appearing on each page.

Attitude instrument

The MAS questionnaire (Garcia, 2001) consisting of 25 items and one open-response item was adapted in order to elicit relevant information on the participants' attitude towards using the multimedia application "Leonidas".

The repeated forward–backward translation procedure was adopted as it is most commonly quoted in the adaptation and translation process (Carlson, 2001) and was considered to be the best within the strategies which were pragmatically possible. In this procedure a forward translation is made from the source original language to the target new language. The target language version is then translated back into the source language and compared to the original version. Errors in the target language version are identified through changes in meaning that arise in the back translation.

The procedure was broadly divided into four phases. Phase 1 was to make four Greek translated versions of the original survey and unify these four. Phase 2 was to produce a back-translated version. Phase 3 was to check the equivalence between the original survey and the back-translated version. Phase 4 was to continue forward and backward translation until satisfactory equivalence was agreed.

In Phase 1, two pairs of two bilingual and bicultural colleagues of the ATHENS 2004 translation department were separately asked to translate the original scale into Greek while discussing among in pairs about content, semantic and conceptual equivalence between the original and their translation. All the translators were fully informed of the objectives of their role in the whole procedure.

One of the authors (whose first language was Greek) unified the two Greek translations created by this process into a single translated version. Selection among alternative Greek translations was based upon perceived ‘‘naturalness’’ of the linguistic expression in the Greek language version.

In Phase 2, a further pair with a native English speaker and a native Greek speaker, both unaware of the original scale, was identified. They were asked to back-translate the Greek version produced in Phase 1. Again, they were professional translators, and were required to discuss content, semantic and conceptual equivalence and to emphasise meaning rather than word-to-word translation.

In Phase 3, a panel of two language experts, two software industry professionals, and two educational leadership content experts compared the original scale and the back-translation brought about by Phase 2, and checked for semantic discrepancies. In Phase 4, the author altered the Greek expression of the parts found to be problematic in Phase 3 with reference to any alternatives rejected in Phase 1. The pair used in Phase 2 re-translated them into English. The panel of experts used in Phase 3 checked discrepancies between the original scale and the re-translation. Detailed discussion of cultural difference and nuance ensured semantic equivalence and aimed to overcome conceptual differences by identifying parallel concepts that might be perceived as stressful. More specifically, the language experts helped eliminate unintended complexity and imprecision in wording. They remarks also helped ensure cultural neutrality and detect wording that might bias responses. Software industry professionals and educational leadership content experts suggested ancillary constructs and operationalization techniques suitable to the goals of the study, in addition to critiquing the instrument for clarity. These experts reviewed questions in such detail that in some cases they identified individual words that 'didn't feel right'. This process was repeated until problems were resolved.

According to expert's recommendation, a few of the items were modified to meet the team leaders Olympic Games milieu. Moreover, the open-ended question was also excluded from the selected parts, since there were only six participants that wrote some comments about the multimedia application.

The modified questionnaire contained twenty five (25) items on a format that used a 5-point Likert-type scale (1=strongly disagree, 2=disagree, 3=neither disagree nor agree, 4=agree, 5=strongly agree). This format allowed participants to select a response from “1,” to “5,” representing their disagreement or agreement on the particular item respectively where “3” stood for a neutral response (see Table 2).

Procedure

Students enrolled in “Team Leaders” Training Program of the Organizing Committee for the Olympic Games ATHENS 2004, were invited to participate in a study designed to understand the user attitude levels of a multimedia application used as a supportive tool for a course in a traditional classroom. The researchers administered the questionnaire during the last class session on the ninth week of the training program due to the nature of the questionnaire, since the questionnaire is typically offered to users after they have completed a session of work with the particular multimedia application. Students were informed verbally and briefly on the research topic and the questionnaire. Participation of the students was voluntary since confidentiality was guaranteed (i.e., students did not place their name on any of the materials in the study). Participants were presented with a letter of informed consent and provided the URL to the online survey. No technical errors were encountered during the completion of the online survey. Data were analyzed using SPSS 13 statistical software.

Design

Due to practical limitations, a field experiment, instead of a laboratory experiment was conducted to test the hypotheses. The experiment was a factorial design with sex groups (males and females) as independent variables, and attitude performance as dependent variable.

Factor analysis was conducted to identify underlying clusters or relationships concerning the learners' attitude towards the multimedia application “Leonidas”. In determining the internal consistency of the attitude scale, the alpha reliability method was used. Independent-samples t test analysis was conducted to investigate the differences of this attitude among the sex of the participants.

The hypotheses of this study were:
H1: There are multiple dimensions underlying the 25 attitude items toward the multimedia application.

H2: The 25-item measure of attitude control is reliable.
H3: The males will have more positive attitudes than females toward the multimedia application.

Results

Means and standard deviations for each factor in this study are presented on Table 1, while the means and standard deviations for the sex groups are presented on Table 3. The results of each analysis are given separately below.

Table 1. Means[1] and standard deviations for each factor

Factors	N	Mean	S.D.
Individualized instruction	232	4.25	.70
Self-paced instruction	232	4.22	.66
Involvement	232	3.48	.63
General experience	232	3.73	.88
Interaction	232	3.63	.97
Learner's control	232	3.64	.75
Anxiety	232	3.73	.43
User-friendliness	232	4.47	.63

[1] Scale: 1=strongly disagree, 2=disagree, 3=neither disagree nor agree, 4=agree, 5=strongly agree

Factor analysis

A principal component analysis of the 25-item scale was performed in order to investigate the underlying dimensions of the educational web site's evaluation, using the SPSS Factor Analysis program. Prior to performing principal component analysis the suitability of data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of many coefficients of .35 and above. The Kaiser-Meyer-Oklin values was .769, exceeding the recommended value of .6 and the Bartlett's Test of Sphericity =5304.243, reached statistical significance ($p<.001$), supporting the factorability of the correlation matrix (Tabachnick, & Fidell, 2001).

Results indicated that our initial hypothesis of multidimensionality was correct. The principal components analysis revealed the presence of eight components with eigenvalue exceeding 1. An inspection of the screen plot revealed a clear break after the eighth component. Based on screen plot and the eigenvalues, it was decided to retain eight components for further investigation. To aid in the interpretation of these eight components, Varimax rotation was performed (Stevens, 1996). The rotated solution (presented in Table 2) revealed the presence of simple structure, with eight components showing a number of strong loadings, and all variables loading substantially on only one component. The eight factors solution explained a total of 89.98 per cent of the variance, with component 1 contributing 15.98 per cent, component 2 contributing 12.45 per cent, component 3 contributing 12.26 per cent, component 4 contributing 11.92 per cent component 5 contributing 10.87 per cent, component 6 contributing 10.85 per cent, component 7 contributing 7.89 per cent and component 8 contributing 7.74 per cent. The interpretation of the eight components was defined as follows:
(1) Individualized instruction, (4 items)
(2) Self-paced instruction (3 items)
(3) Involvement (3 items)
(4) General experience (4 items)
(5) Interaction (3 items)
(6) Learner's control (3 items)
(7) Anxiety (3 items) and
(8) User-friendliness (2 items).

Table 2. The rotated loading matrix from the factor analysis[1]

Items	1	2	3	4	5	6	7	8	H^2
I enjoyed doing this exercise by myself.	.984								.995
I would have liked to have had a partner to work within these multimedia lessons.	.972								.976

The multimedia exercises turned out to be efficient thanks to the fact that there was only one student per session.	.937								.921
I liked working with the application without having to share it with other students.	.947								.929
To be able to work at my pace resulted in a more effective instruction.		.917							.957
I feel more motivated when I am allowed to work at my own pace.		.948							.953
I did not like to be left working at my own pace		.938							.943
The time flew while I was working with the multimedia lessons.			.939						.955
I had the feeling that the time to finish with the multimedia sessions never got to its end.			.922						.920
As soon as I start to work with the multimedia lessons I feel immersed in the activity.			.952						.932
The software I have worked with looks good to me.				.855					.879
The multimedia lessons are well designed.				.910					.900
The lessons have been planned out well.				.749					.872
In general. it has been a good experience to work with the interactive lessons.				.751					.762
The interaction with the instructional material through the computer was pleasant					.762				.817
The interactions with the computer were more positive.					.852				.799
The interactions with the computer made me be more attentive all the time					.718				.620
When exploring the program I was not happy when I found out that it was me who had to decide what needed to be done at every step.						.816			.862
I was grateful for the freedom I was given to explore the activity in my own way.						.892			.939
I did not like to be able to navigate freely throughout the program						.848			.932
The lessons with interactive multimedia make me feel tense.							.853		.964
I get nervous when I think that I am going to study lessons with multimedia technology.							.818		.856
When I examine material with interactive multimedia I feel comfortable.							.616		.856
I did not find the program confusing to use.								.932	.976
The program was user-friendly.								.888	.982
% of variance	15.98	12.45	12.26	11.92	10.87	10.85	7.89	7.74	
Total variance									89.98
Eigenvalue	3.997	3.112	3.066	2.980	2.719	2.715	1.973	1.935	

[1] H^2 = communalities

Reliability analysis

Coefficient alpha is the statistic mostly used to assess the internal consistency. The Cronbach-alpha coefficient was calculated for each of the sub-scales. The *"Individualized instruction"* factor had an a =.83, the *"Self-paced instruction"* had an a =.89, the *"Involvement"* factor had an α =.86, the *"General experience"* factor had an α =.86, the *"Interaction"* factor had an α =.75, the *"Learner's control"* factor had an α =.90, the *"Anxiety"* factor had an α =.82 and the *"User-friendliness"* factor had an a =.79. Although statistical texts (DeVellis, 1991) suggest that scale with reliabilities more than 0.70 should normally be considered as acceptable, in practice lower limits have been set up as acceptable by researchers.

Independent-Samples *t* Test analysis

An independent-samples *t* test was conducted to evaluate the hypothesis that males have more positive attitudes than females toward the multimedia application. There was significant difference in scores for males (M=3.46, SD=.77) and females (M=4.00, SD=.52) in the factor "*General experience*" $t(230)=9.452$, $p<.01$. As shown in Table 3, the females scored significantly higher in the above factor, counter to the research hypothesis. No significant difference was found between the two sexes groups in any case of the remaining seven factors of the MAS.

Table 3. Means[1] and standard deviations for the sex groups in each factor

Factors	Males			Females		
	N	Mean	S.D.	N	Mean	S.D.
Individualized instruction	122	4.28	.70	110	4.21	.66
Self-paced instruction	122	4.19	.63	110	4.23	.68
Involvement	122	3.49	.74	110	3.47	.74
General experience	122	3.46	.77	110	4.00	.52
Interaction	122	3.59	.87	110	3.67	.84
Learner's control	122	3.63	.78	110	3.66	.73
Anxiety	122	3.72	.59	110	3.74	.53
User-friendliness	122	4.44	.68	110	4.50	.62

[1] Scale: 1=strongly disagree, 2=disagree, 3=neither disagree nor agree, 4=agree, 5=strongly agree.

Discussion

This study adapted a questionnaire in order to evaluate the multimedia application "Leonidas" considering the attitudes of Athens 2004 employees. The study also sought to investigate differences among the sex of the participants. Results indicated that the evaluation on a pedagogic multimedia application was a multidimensional concept. This fact has been proved from other studies that have examined the role of the multimedia application as an educational tool (Garcia 2001; Selwyn, 1997). As a result of the factor analysis conducted in each of the pre-defined subscales, all items agree with the attitudinal dimensions of MAS proposed by Garcia (2001). The reaction of learners to the multimedia application "Leonidas" was encouraging. Analysis of the survey revealed a generally strong positive attitude towards this particular multimedia application.

The finding was not a surprise given the learners' positive attitude toward the multimedia application but the level of positive reaction was higher than expected. The explanation in this phenomenon could be that participants in this study already had increased interest in Olympic issues. Factors that could have contributed in this were the multimedia experience of the participants with Olympic applications and their Greek origin. If this were the cases, it was also likely that some other group of learners were less favourable toward the multimedia application "Leonidas" of Athens 2004 training department. Also, the use of volunteers clearly had predisposed the learners towards more positive attitudes.

Further analysis of the survey showed that the first factor of the questionnaire "Individualized instruction" had positive ranging from "agree" to "strongly agree" for the majority of learners (76%). This reveals that participants found the multimedia application "Leonidas" as an individual and self-paced learning tool that allows them to work privately, in an enjoyable environment on their own. The factor "Self-paced instruction" had positive ranging from “agree” to “strongly agree” in 78% of the learners. This indicates that the multimedia application “Leonidas” contained materials that allowed the learners to learn at their own pace, giving them a sense of control over learning. The third factor “Involvement” had the smallest positive impact on attitude of the multimedia application “Leonidas” ranging from “agree” to “strongly agree” in 53% of the learners. The explanation to this phenomenon could be that learners between the ages from 33-44 years old may need more sophisticated and complicated applications to have their work done. Another consideration could be that the learners were not satisfied with the amount and the clarity of information received. Also, participants found the learning experience in general worthwhile since the 90/% of the respondents rated the “General experience” questions by answering, from “agree” to “strongly agree”. The fifth factor “Interaction” had positive ranging from “agree” to “strongly agree” in 68% of the learners. This means that the particular multimedia application contained interactive features that would empower the learners to control the content and the flow of information and encouraged them to be responsible for their own learning. Moreover, the factor “Learner’s control” had positive ranging from “agree” to “strongly agree” in 73% of the learners. This indicates that the participants felt happy when they explored the multimedia application and they found out that they have to decide by themselves what needs to be done at every step, by exploring the activity in their own way. The seventh factor “Anxiety” had positive ranging from “agree” to “strongly agree” in 72% of the learners. This reveals that participants felt nerveless and comfortable when they studied lessons by browsing the material via multimedia application. Finally, the strong positive responses on the last factor “User-friendliness” made it the most dominant in increasing “Leonidas” attitude. This shows that participants found the multimedia application easy to use, all necessary special commands were clear and the user interface issues such as menu design and readability of screens had been addressed.

The research on how sex changes attitudes of the multimedia application “Leonidas” showed no significant differences. Males and females answered the questions of the survey the same way, indicating similar attitude. This suggests that using the multimedia application “Leonidas” has a positive effect for both sexes. Similar results have been reported by Antoniou, Patsi, Bebetsos & Ifantidou, (2006) and North and Noyes (2002), who found that the impact of psychological gender, does not influence significantly attitudes towards computers. Other researchers report that males have more positive attitudes than females (Bebetsos, Kouli & Antoniou, 2007; Ho and Lee, 2001; Schumacher and Moharan-Martin, 2001) or conversely finding that females liked computers more than males (Keasar, Baruch & Grobgeld-Dahan, 2005). Thus, the subsequent psychological gender theories of human–computer interaction (namely the socialization theory as applied by Whitely, 1997) are unsupported.

The fact that the gender differentiation has not occurred may be viewed on two levels. First, there may be a general cohort effect or second, there may be confounding factors exclusive to the sample group. In relation to the first point, the positive attitudes found in both males and females may be associated with changes in societal values and the socialisation processes in today’s computer generation (Whitely, 1997). These are perhaps mediated by the impact of increased use of multimedia applications in organizations, at home and software developments improving multimedia applications interfaces. In other words the MCAI may no longer cultivate gender differences in multimedia applications attitudes per se. The second possible explanation for an absence of gender differences is that there may be some factors intrinsic to this sample group that were responsible. This overlaps with the previous idea of a cohort effect, e.g. it may be that the managers and the trainers at ATHENS 2004 organization were particularly keen to ensure all employees viewed multimedia applications positively and did not convey a gender bias.

The more detailed component analysis found that the attitude towards the factor “General experience” had a significant sex effect. This in part supports the suggestion by Whitely (1997) that gender effect exist on some attitude components but not others, which is fundamentally based on the assumption that adults hold bi-directional views about computers and especially multimedia applications. The fact that the females in this study viewed multimedia computing ability more impartially than males, may explain why they displayed positive attitudes in general about multimedia experience and why a significant difference was not found on other attitude components. In other words females did not accept the belief that multimedia computing was related to gender, mathematics background or nationality and viewed multimedia ability perhaps as an open option (Sanders, 2006). According to the socialization hypothesis, a greater acceptance of the belief that multimedia computing is inappropriate should be associated with more negative attitudes (Whitely, 1997). Since the females in this sample did not endorse the views that multimedia computing ability was related to sex then the absence of a sex difference on attitudes is not surprising.

Implications for practice

The use of multimedia technology in traditional classrooms has been growing at a rapid pace. Though many instructors are using various modes of multimedia technology to communicate with and instruct their learners, it is important to understand that these various modes affect not only learner acceptance and performance but also future behaviors in the learning process. This is a major concern because the cost of multimedia technology infrastructures continues to absorb an increasing percentage of organizations' budgets. Therefore, this study is timely and has several practical implications.

First, the findings reveal that the use of multimedia resources provides complementary learning activities that aid the learning process. There is great interest and potential in MCAI flexible learning with many instructors incorporating some form of multimedia technology as a part of their instruction in organizations training (Chandler, 2003). While instructors are the focal point in most course settings, it should be noted that complementary learning activities are just as important for practice, if not more so.

Second, the use of multimedia technology in organization training is a matter that not only the Organizing Committee for the Olympic Games ATHENS 2004 should be interested in, but also all the training organizations should benefit from the results of this study for further corporate planning. Furthermore, they should not only provide the opportunities of multimedia technology and MCAI for this organization but also they should take into consideration the experience other training organizations had in this area of study. They should take into consideration the attitudes of the employees on the uses of multimedia technology in organization and should prepare the courses required multimedia using for their learners. Because as the employees' success increases the success of the organization increases.

Third, using multimedia applications is time-consuming and labor-intensive, if productive outcomes are to be derived. Learners and instructors may find valuable resources and increased opportunities in communication through the multimedia technology, but at the expense of continuous effort and time consumption. Establishing an interactive and dynamic MCAI course like the learning environment of this study can help overcome time consumption difficulty, while providing learners with quick and convenient ways to find useful information. Multimedia applications create a much more interactive learning environment thereby increasing the effectiveness of learning.

Finally, a well-designed MCAI course provides the balance of real and virtual classrooms and class sessions. This ideally makes the class a more continuous environment rather than an environment, which is done in one or two hours and then set aside for the remainder of the week. Continuing education of employees can offer competitive advantages; field experiments within real-world organizations would be very useful to organizations already using interactive multimedia training to identify if their programs are effective and acceptable. For others, a field experiment might suggest if migration to self-paced instruction via interactive multimedia would be relevant and how it should be designed.

Limitations

As with all investigations, this study is not without limitations. First, the data used in this study were drawn from a single corporation sample. The organization is best described as a large, organizing committee for the Olympic Games located in Greece. Thus, the findings should be interpreted with caution and generalizations may only be relevant to organizations similar in size, control status, and corporation emphasis. The present study used self-report data and this may be another possible limitation. To the extent that respondents did not know the information being requested or found survey questions to be ambiguous and unclear, the generalizability of these findings may be limited.

Perhaps another limitation relates to the dataset used in the study — the MAS questionnaire. Perhaps, the attitude instrument of this study was limited to factors that could be defined or operationalized using items drawn from the database. It is highly possible that the MAS did not measure all of the variables needed to explain the variance in student self-reported attitude toward multimedia applications. Likewise, it is plausible that the MAS items have a marginal relationship with the constructs (e.g., Interaction, Learner's control, Anxiety, etc.) that they are purported to measure (Garcia, 2001).

Despite these limitations, this study contributes to our understanding of the potential effect of various uses of multimedia technology on learners' attitudes in training organizations. Specifically, it provided information about the association between employees' use of multimedia technology and self-reported attitude toward multimedia application "Leonidas". In addition, this research provides a foray into group differences that exist between males and females of multimedia technology.

Conclusion

In conclusion the learner's feedback from the questionnaires indicated a general level of satisfaction and contentment with this particular multimedia application. Yet, in order to have the learners make constructive and flexible use of the educational multimedia technologies, the "Individualized instruction", the "Self-paced instruction", the "Involvement", the "General experience", the "Interaction", the "Learner's control", the "Anxiety" and the "User-friendliness" seem to be crucial considerations. Perhaps, adherence to these basic principles will not only improve overall multimedia impressions, but also will increase use frequency to the multimedia application concerned. The scale adapted in the present study can be a useful tool for the evaluation of other relative multimedia applications by multimedia developers. Nevertheless, further examination is warranted in order to obtain additional information concerning the difficulties of multimedia experience on employees' attitudes toward multimedia applications.

When using multimedia technology in organization, it is strongly recommended that trainers take some time in assessing employees' attitudes toward multimedia technology prior to the structuring of instruction and its implementation in the training sessions. This approach is appropriate in that it ensures that the learners will have maximum gains in utilizing multimedia applications as a tool for learning. Furthermore, education managers will be given the chance to create an environment that can be conducive to the learners.

Research and development in this area will be continued with the view to refining any kind of multimedia educational environment so that it meets and full fills all expectations for supporting and enhancing employees learning process. More studies should be conducted to investigate the effect of multimedia experience on learner's attitudes toward the multimedia applications, especially when its effect is linked to gender. Also, one can reasonably assume that most people – regardless of gender, age, or other demographic factors – access multimedia application credibility in similar ways. Although real differences do exist, it's more striking to see how many things were not different, suggesting that the various demographic groups shared similar approaches to evaluating multimedia applications.

References

Antoniou, P., Patsi, H., Bebetsos, E., & Ifantidou, G. (2006). Validity of scale and evaluation of students' attitudes toward computers. Compare with students' attitudes toward physical education and physical activity. *Inquiries in Sport & Physical Education, 4*(1), 114–124.

Bebetsos, E., Kouli, O., & Antoniou, P. (2007). Attitudes and Behaviors of University PE Students Towards the Use of Computers. *International Journal of Computer Science in Sport, 6*(1), 55–63.

Becker, K. H., & Maunsaiyat, S. (2002). Thai Students' Attitudes and Concepts of Technology. *Journal of Technology Education, 13*(2), 6–19.

Cardoso, A., Peralta, H., & Costa, F. (2005). The students' point of view about quality of educational multimedia software. *Interactive Educational Multimedia, 11,* 38–59.

Carlson, E. D. (2001). A Case Study in Translation Methodology Using the Health-Promotion Lifestyle Profile II. *Public Health Nursing, 17*(1), 61–70.

Chandler, H. (2003). Concept mapping: WebQuests in Social Studies. *Media & Methods, 39*(3), 38–39.

Christensen, R., & Knezek, G. (2000). Internal Consistency Reliabilities for 14 Computer Attitude Scales. *Journal of Technology and Teacher Education, 8*(4), 327–336.

DeVellis, R. F. (1991). *Scale Development: Theory and Applications.* Newbury Park, CA: Sage Publications.

Gandole, Y. B., Khandewale, S. S., & Mishra R. A. (2006). A Comparison of Students Attitudes between Computer Software Support and Traditional Laboratory Practical Learning Environments in Undergraduate Electronics Science. *E–Journal of Instructional Science and Technology, 9*(1). Retrieved July 10, 2008, from http://www.usq.edu.au/electpub/e-jist/docs/vol9_no1/papers/current_practice/gandole_khandewale_mishra.pdf.

Garcia, J. C. (2001). An instrument to help teachers assess learners' attitudes towards multimedia instruction. *Education, 122*(1), 94–101.

Gardner, D. G., Discenza, R., & Dukes, R. (1993). The measurement of computer attitudes: An empirical comparison of available scales. *Journal of Educational Computing Research, 9*(4), 487–507.

Ho, S. M. Y., & Lee, T. M. C. (2001). Computer usage and its relationship with adolescent lifestyle in Hong Kong. *Journal of Adolescent Health, 29*(4), 258–266.

İşman, A., & Dabaj, F. (2004). Attitudes of Students towards Internet. *Turkish Online Journal of Distance Education, 5*(4). Retrieved July 11, 2008, from http://tojde.anadolu.edu.tr/tojde16/pdf/dabaj.pdf.

Jones, T., & Clarke, V. A. (1994). A computer attitude scale for secondary students. *Computers and Education. 22*(4), 315–318.

Kay, R. H. (1993). A practical research tool for assessing ability to use computers: The computer ability survey (CAS). *Journal of Research on Computing in Education, 26*(1), 16–25.

Keasar, T., Baruch, R., & Grobgeld-Dahan, E. (2005). An evaluation of web enhanced instruction in college level biology Courses. *Australasian Journal of Educational Technology, 21*(4), 533–545.

North, A. S., & Noyes J. M. (2002). Gender influences on children's computer attitudes and cognitions. *Computers in Human Behavior, 18,* 135–150.

Sanders, J. (2006). Gender and technology: A research review. In Skelton, C., Francis, B. & Smulyan L. (Eds.). *Handbook of Gender and Education* (pp. 1–40). London: Sage Publications.

Schumacher, P., & Moharan-Martin, T. (2001). Gender, Internet and computer experiences. *Computers in Human Behavior, 17,* 95–110.

Selwyn, N. (1997). Students' Attitudes toward Computers: Validation of a Computer Attitude Scale for 16–19 Education. *Computers and Education, 28*(1), 35–41.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th Ed.). Needham Heights, MA: Allyn & Bacon.

Teoh, B. S. P., & Neo, T. K. (2007). Interactive Multimedia Learning: Students' Attitudes and Learning Impact in an Animation Course. *The Turkish Online Journal of Educational Technology, 6*(4). Retrieved July 12, 2008, from http://www.tojet.net/articles/643.doc.

Whitely, B. (1997). Gender differences in computer related attitudes and behavior; a meta-analysis. *Computers in Human Behavior, 13*(1), 1–22.

Woodrow, J.E. (1991). A comparison of four computer attitude scales. *Journal of Educational Computing Research, 7*(2), 165–187.

Lu, E. J.-L., Horng, G., Yu, C.-S., & Chou, L.Y. (2010). Extended Relation Metadata for SCORM-based Learning Content Management Systems. *Educational Technology & Society*, *13* (1), 220–235.

Extended Relation Metadata for SCORM-based Learning Content Management Systems

Eric Jui-Lin Lu[1], Gwoboa Horng[2], Chia-Ssu Yu[2] and Ling-Ying Chou[3]
[1]Department of Management Information Systems, National Chung Hsing University, Taichung, Taiwan // jllu@nchu.edu.tw
[2]Department of Computer Science and Engineering, National Chung Hsing University, Taichung, Taiwan // gbhorng@cs.nchu.edu.tw // s9456010@cs.nchu.edu.tw
[3]Department of Applied English, Diwan University, Tainan, Taiwan // annechou@cc.kuas.edu.tw

ABSTRACT

To increase the interoperability and reusability of learning objects, Advanced Distributed Learning Initiative developed a model called Content Aggregation Model (CAM) to describe learning objects and express relationships between learning objects. However, the suggested relations defined in the CAM can only describe structure-oriented relationships and cannot express semantic relationships between learning objects. Although extended relations were proposed in the past, some of the proposed relations are redundant and even inappropriate. In addition, the usefulness of these relations has never been formally studied. To solve the problems, we systematically studied these relations from authors' perspective and proposed an extension to CAM. The extension was tested by 30 authors using a web-based learning content management system that was developed by us.

Keywords

Learning content management systems, SCORM, Reusability, CAM, Metadata

Introduction

Due to the emergence and flourishing of the Internet, the development of e-Learning systems has become an important research topic in both academia and industries. As a result, many learning systems and learning objects (LOs) were developed. One major problem of these LOs is that they cannot be reused among different learning systems. To resolve the problem, Advanced Distributed Learning Initiative (ADL) developed a reference model called Sharable Content Object Reference Model (SCORM).

There are two kinds of LOs defined in SCORM. One is asset, and the other is sharable content object (SCO). Assets are digital media such as text, images, sound, assessment objects, or any other piece of data. Each SCO is composed of assets or other SCOs. To increase reusability and interoperability of LOs, metadata can be defined for each LO.

ADL developed a metadata model called Content Aggregation Model (CAM). CAM, adapted from IEEE Learning Object Metadata (LOM), classifies all metadata into nine categories, and one of the categories is "RELATION". A relation in the "RELATION" category is mainly used to describe a LO and express relationships between learning resources. When used skillfully, a relation is a very useful metadata that can enhance learning effectiveness as well as increase the reusability of LOs. For example, as shown in Figure 1, LO_A describes how bubble sort works. At the bottom of LO_A, there is a figure illustrating how bubble sort works in steps. Using the proposed relation metadata, we can define the figure as a learning object of type "Illustration". When the figure is stored in a repository, it can be easily searched and reused by other users. Additionally, the application of relations can be further extended. If the author of LO_A wishes to provide more illustrations to help learners, she can easily provide links to more illustrations such as LO_B and LO_C. LOs such as LO_B and LO_C can be created by the author or other authors as long as they can be accessed. Also, these LOs can be searched and retrieved from repositories.

As defined in CAM, there are twelve suggested values as shown in Table 1 for the "RELATION" category. However, these suggested values can only describe structure-oriented relationships and cannot express semantic relationships between learning resources (Steinacker et al., 1999). In the past, many relations were proposed (Karger et al., 2006; Ullrich, 2004, 2003, 2005; Loser et al., 2002; Steinacker et al., 1999, 2001; Sddik et al., 2001; Fischer, 2001; Fischer et al., 1999). These relations are mainly based on two major theories. One is instructional design theory (IDT), and the other is rhetorical structure theory (RST). Although these relations can express meaningful relationships between LOs, they are limited in the following ways.

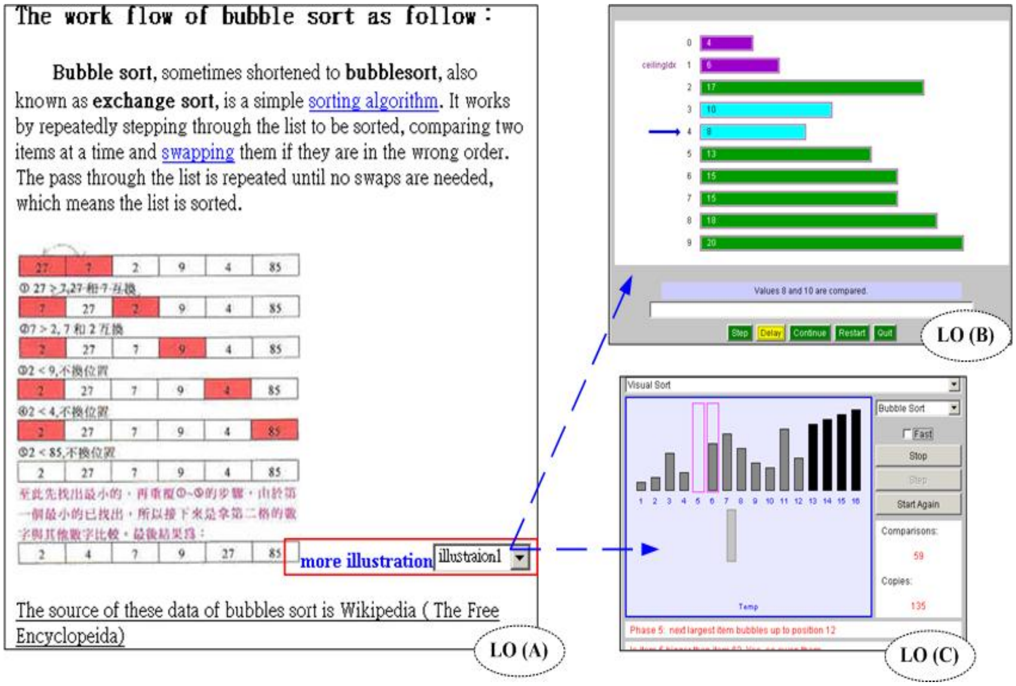

Figure 1: An Example Application of RELATION Category

Table 1: The Suggested Values for RELATION Category

ispartof	haspart	isversionof	hasversion
isformatof	hasformat	references	isreferencedby
isbasedon	isbasisfor	requires	isrequiredby

First of all, some of the relations are redundant. For example, both “Example” and “Illustration” relations are not only defined in IDT, but also defined in RST. Secondly, several relations have already been defined in CAM. For example, the “Interactivity” relation defined in IDT has been defined in the “EDUCATION” category of CAM. Thirdly, although the relations can express semantic relationships between LOs, the usefulness of the relations has never been formally studied. Fourthly, there is no learning content management system supporting new relations. Finally, after studying the relations, we found that some relations are confusing and may not be easy for authors to clearly specify relation for each LO.

In this paper, we studied all 32 relations proposed in IDT and RST, removed 8 duplicated relations, and analyzed the rest 24 relations. A web-based learning content management system supporting these 24 relations was designed and implemented. By using the system, authors are allowed to assign a specific relation to a LO. Also, the relationships between LOs can be easily defined. When learners read LOs (for example, LO_A as shown in Figure 1), “more illustration” will be shown along with LO_A.

The developed system was also used as the platform to survey 30 authors. The survey was divided into two parts: one was to study whether or not they would use these relations in LOs to increase learning effectiveness; the other was to investigate, when specifying a relation for a LO, whether or not they might possibly use the other 23 relations for the same LO. Based on the results of the survey, extended relation metadata was proposed.

The rest of the paper is organized as follows. First, we briefly discussed the relationships between LOs and described all 32 relations in Section 2. The design and implementation of the prototype is described in Section 3. In Section 4, we described how the survey was conducted and showed the survey results. Additionally, the discussions of the results and suggestions were presented. Finally, we drew our conclusions and future works in Section 5.

Literature Review

Currently, existing relations are mainly based on two theories: one is instruction design theory (IDT) (Reigeluth, 1999), and the other is rhetorical structure theory (RST) (Mann and Thompson, 1988).

Instruction Design Theory

IDT encourages teachers to search for related learning resources and exploit them to satisfy all possible learning needs. Based on IDT, Ullrich (Ullrich, 2003, 2004; Karger et al., 2006) proposed twenty-three relations, and they are summarized in Table 2. The descriptions and examples for these relations are provided in Appendix A for reference.

Table 2: The Relations Defined in IDT

Definition	Fact	Law	Law of Nature	Theorem
Process	Policy	Procedure	Interactivity	Illustration
Example	Counterexample	Evidence	Proof	Demonstration
Explanation	Introduction	Conclusion	Remark	

The Rhetorical Structure Theory

In the past, RST is an approach to analyze the rhetorical structure of article contents. Recently, researchers (Steinacker et al., 1999; Seeberg et al., 1999; Sddik et al., 2001; Steinacker et al., 2001; Fischer, 2001; Loser et al., 2002) extended the concept to express the relationships between LOs and proposed rhetorical-didactic relations. The rhetorical-didactic relations include 9 relations, and they are summarized in Table 3. The descriptions and examples for these relations are provided in Appendix B for reference.

Table 3: The Relations Defined in RST

Example	Illustration	Instance	Restriction	Amplify
Continues	Deepen	Opposition	Alternative	

Discussions

After careful study of these relations, it is found that both “Example” and “Illustration” were redundantly defined in both IDT and RST. “Interactivity” and its 4 subclasses have already been defined in the “EDUCATION” category of CAM. Additionally, “Alternative” can be replaced with “hasformat” which is defined in the “RELATION” category of CAM. As a result, out of 32 relations defined in IDT and RST, only 24 relations need to be further investigated.

The Design and Implementation of a Learning Content Management System

Design

Like IEEE LOM, SCORM CAM classifies metadata into nine categories. “RELATION” is one of the nine categories. To define that LO_A references LO_B, one must define a <relation> element in the <lom> of LO_A as shown in Figure 2. Additionally, one has to define a <kind> element in <relation> and define both <source> and <value> elements in <kind>. The values for the <source> and <value> elements are *LOMv1.0* and *references*, respectively, to express that it is a references relation based on LOM version 1.0. LO_B is defined in a <resource> element which is a subelement of <relation>. In the <resource> element, <identifier> is used to identify the location of LO_B, and <description> is used to describe LO_B.

```
<lom>
  <relation>
    <kind>
      <source>LOMv1.0</source>
      <value>references</value>
    </kind>
    <resource>
      <identifier>
        <catalog>URI</catalog>
        <entry>http://www.example.org/courses/B</entry>
      </identifier>
      <description>
        <string language="en">An example of LMa</string>
      </description>
    </resource>
  </relation>
</lom>
```

Figure 2: An Example SCORM CAM Relation

When designing the learning content management system, we attempted to revise CAM as less as possible. We also used the concept of RDF triple to design our metadata. First of all, the name of the extended relation metadata is called NCHUMISv0.1. In other words, if the extended relation metadata were used, the value for the <source> element should be NCHUMISv0.1. The suggested values for the <value> element in NCHUMISv0.1 are the 24 relations which include “Conclusion”, “Continues”, “Counterexample”, “Deepen”, “Definition”, “Demonstration”, “Evidence”, “Example”, “Explanation”, “Extension”, “Fact”, “Guideline”, “Illustration”, “Instance”, “Introduction”, “Law”, “Law of Nature”, “Opposition”, “Procedure”, “Process”, “Proof”, “Remark”, “Restriction”, and “Theorem”.

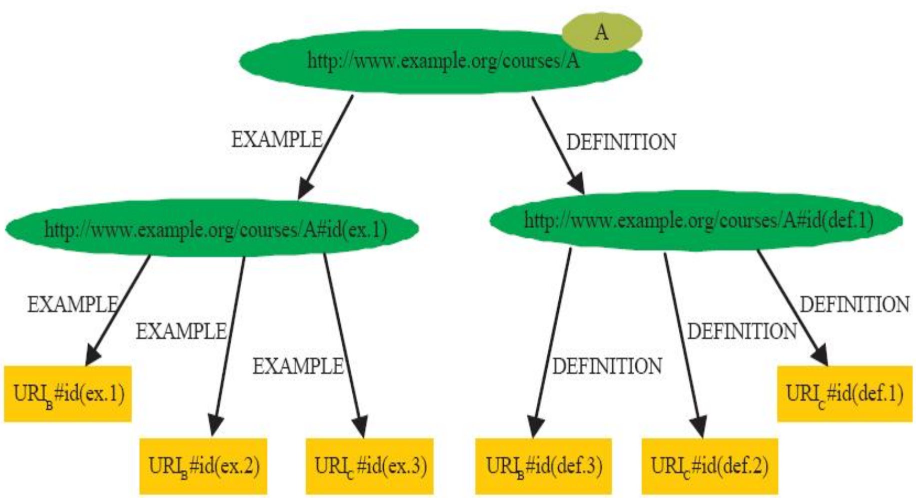

Figure 3: An example Relation Architecture

To describe how to use the relations in applications such as the one shown in Figure 1, Figure 3 is used as an example. In Figure 3, it shows that LO_A, identified as http://www.example.org/courses/A, has an example (identified as http://www.example.org/courses/A#id(ex.1)) and a definition (identified as http://www.example.org/courses/A#id(def.1)). The example was described by a <relation> element as shown in lines 02-13 of Figure 4, while the definition was defined in lines 15-26. Note that, as shown in lines 04 and 17, the value for both <source> elements is NCHUMISv0.1.

The author of LO_A may also provide more examples for A#id(ex.1) or more definitions for A#id(def.1). These examples or definitions are generally stored in repositories and can be authored by her or other authors as long as they are accessible. To be able to express such relationships, we defined extra attributes *id* and *idref* for <relation> elements. With XPointer, we can clearly define the relations between LOs by using *id* and *idref*. For example, to define one extra example link for A#id(ex.1), we first defined an id="ex.1" for the <relation> in line 02. Then, we added another <relation> element in lines 27-38. As shown in line 27, an idref="ex.1" attribute has to be defined to indicate that this relation is for the relation identified as id="ex.1". The newly created <relation> element represents an example located at URI_B#id(ex.1). With the proposed design, one can easily define a resource (which can be all or part of a resource) and link other resources to the resource.

```
01 <lom>
02   <relation id="ex.1">
03     <kind>
04       <source>NCHUMISv0.1</source>
05       <value>example</value>
06     </kind>
07     <resource>
08       <identifier>
09         <catalog>URI</catalog>
10         <entry>http://www.example.org/courses/A#id(ex.1)</entry>
11       </identifier>
12     </resource>
13   </relation>
14
15   <relation id="def.1">
16     <kind>
17       <source>NCHUMISv0.1</source>
18       <value>definition</value>
19     </kind>
20     <resource>
21       <identifier>
22         <catalog>URI</catalog>
23         <entry>http://www.example.org/courses/A#id(def.1)</entry>
24       </identifier>
25     </resource>
26   </relation>
27   <relation idref="ex.1">
28     <kind>
29       <source>NCHUMISv0.1</source>
30       <value>example</value>
31     </kind>
32     <resource>
33       <identifier>
34         <catalog>URI</catalog>
35         <entry>URIB#id(ex.1)</entry>
36       </identifier>
37     </resource>
38   </relation>
39 </lom>
```

Figure 4: The Definition for The Example Relation Architecture

```
<lom>
  <relation id="multi.1">
    <kind>
      <source>NCHUMISv0.1</source>
      <value>example restriction</value>
    </kind>
    <resource>
      <identifier>
        <catalog>URI</catalog>
        <entry>http://www.example.org/courses/A#id(multi.1)</entry>
      </identifier>
    </resource>
  </relation>
</lom>
```

Figure 5: An Example Metadata That Supports Two Relations

Sometimes, a LO can be of more than two relation types. For example, as shown in Figure 5, A#id(multi.1) is an example in LO_A. However, A#id(multi.1) may be a counterexample for other LOs. To support this, multiple relations are allowed to define in <value> element as long as each relation is separated by a blank space (“ ”). Therefore, for the example shown in Figure 5, <value> example restriction </value> is defined in line 5. It is noted that we think that the relation “restriction” defined in RST is more appropriate to be used in this example than the relation “counterexample” defined in IDT. The possible confusion caused by these two relations would be further investigated in Section 4.

Implementation

Sun’s Java Standard Edition 1.5 and Apache’s Tomcat were used to implement the prototype of the learning content management system. Although all metadata defined in CAM are optional, there are some restrictions imposed when using the extended relation metadata. For example, to be able to reuse existing LOs, it is required for authors to enter values for <keyword> element which is in the “GENERAL” category of CAM.

Figure 6 is a screenshot that demonstrates the prototype allows authors to create or edit a LO (called LO_A). Similar to many personal Blog systems, the prototype allows authors to click on buttons above the text area to insert a relation. Each button represents a relation in the metadata. It is also possible to upload a file which can be an image, a sound clip, or any other multimedia file. The *IMG* and *A* buttons can be used to create images or hyper-links in LOs. As shown in Figure 6, two for loop examples were defined. Also, authors are allowed to enter other relation types of LOs.

After finishing editing LO_A, an author can click on the *OK* button. LO_A will be saved in XML format, an *id* attribute will be inserted for each relation, and the value for each *id* attribute is automatically generated to ensure uniqueness. The resulting page is shown in Figure 7(a). Note that, because two EXAMPLE relations were defined in Figure 6, a drop down menu with these two relations was created as shown in Figure 7(a).

For each relation in the drop down menu, authors are allowed to create more links for the selected relation. For example, if an author wished to provide more example links for the first example, she first selects the first example from the drop down menu and clicks on the *GO* button. The system will then search for all possible relations in repositories. Currently, the search mechanism used is quite simple, although it can be further enhanced in the future. The search engine embedded in the prototype will first look up LOs with the same relation name. From the matched LOs, the search engine will then check their keyword values. LOs with more matched keywords will be ranked higher than those with less matched keywords. Using LO_A as an example, all LOs in repositories with the relation “Example” and the keyword *loop* will be matched and displayed which are shown in Figure 7(b). Authors can select zero or more links for each relation in the drop down menu. As shown in Figure 7(b), the author selected two example links for the first example of LO_A. When learners study LO_A, they will see a page similar to Figure 1. After reading LO_A, a learner may choose to read more examples by selecting an entry in the drop down menu.

Analysis and Discussions

As mentioned earlier, the usefulness of relations based on IDT and RST has never been formally studied. In addition, it is sometimes really difficult for authors to clearly specify one relation for a LO without thinking maybe another relation is more appropriate for the LO. For example, it can be very difficult for authors to choose either “Process”, “Procedure”, or “Continue” for a “sorting algorithm”.

To investigate the above issues, we used questionnaire to survey 30 graduate students at the Department of Computer Science and Engineering, National Chung Hsing University, Taiwan. Out of 30 students, 17 of them have teaching experience, while the other 13 have no teaching experience. The questionnaire was divided into two parts. In the first part, authors were asked whether or not the 24 relations will be helpful for their learners. There are 24 questions (one for each relation) in the questionnaire. Each question was rated on a discrete scale from 5 to 1, with corresponding verbal descriptions ranging from ‘highly useful’ through ‘useful’, ‘no-opinion’, ‘useless’, to ‘highly useless’; respectively.

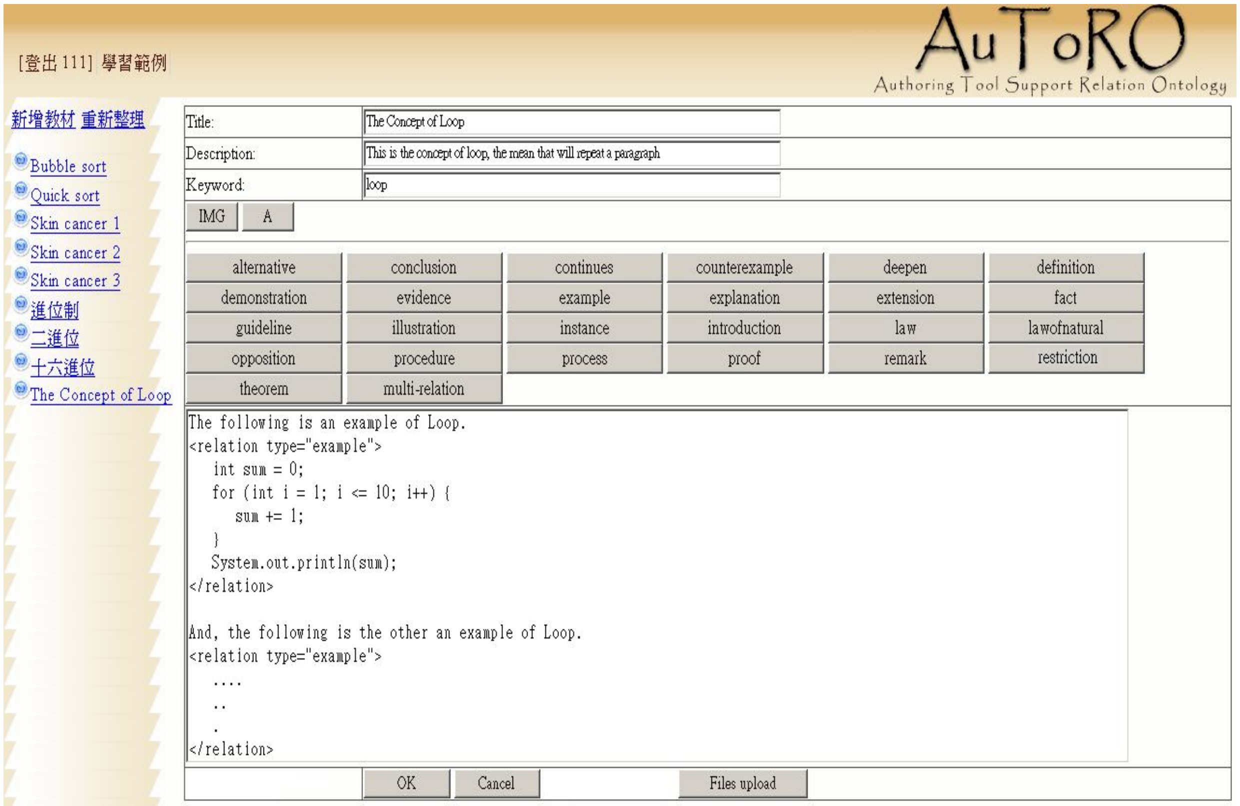

Figure 6: A Screenshot for Creating/Editing Los

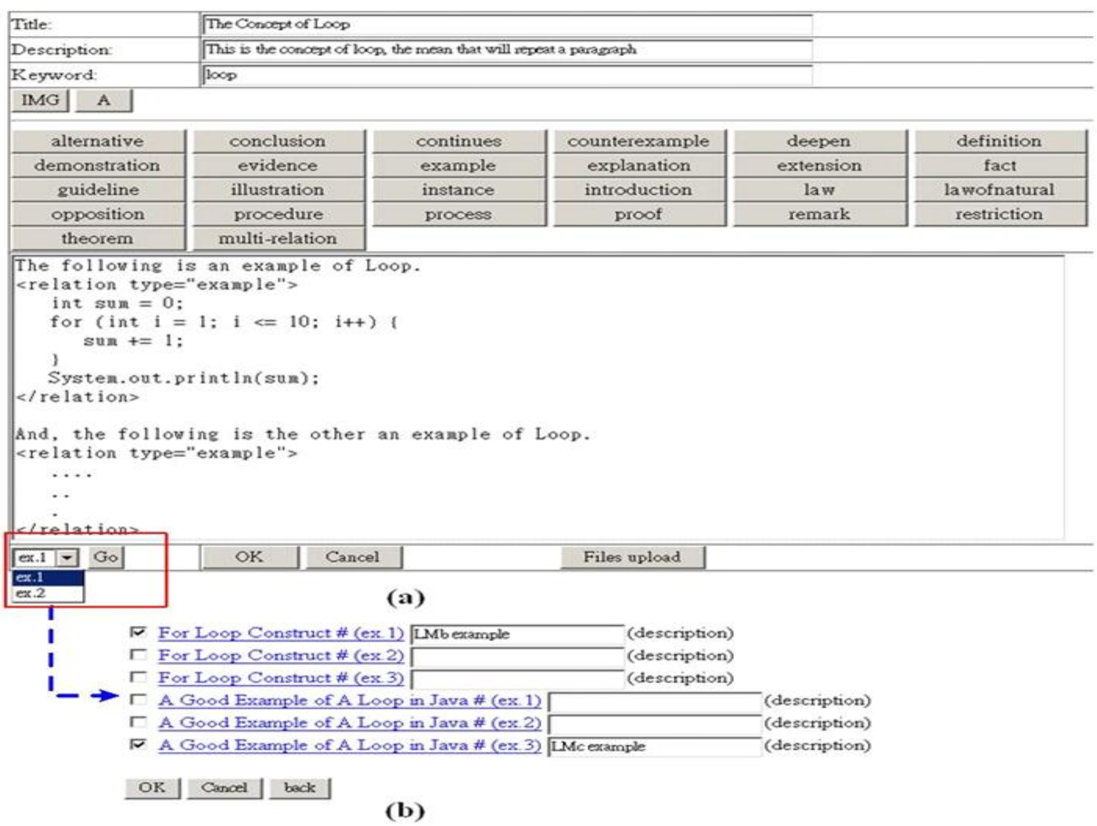

Figure 7: Example screenshots.

In the second part of the questionnaire, authors were asked, when specifying a relation for a LO, whether or not she thought other relations might also be adequate for the LO. There are also 24 questions because there are 24 relations. For each relation, authors were allowed to select zero or more than one relation out of the other 23 relations.

The survey was conducted as follows: Firstly, we explained the meanings of the 24 relations. Then, the prototype is used as the platform to show them how authors can use the prototype to assign a relation for a LO as well as to search and link other LOs to the LO. In the prototype, several LOs had been created using the prototype. Lastly, we let authors to use the prototype and answer the questions. To avoid possible confusions, every author was surveyed individually and each session took about 30 to 60 minutes.

Analysis: Part I

To study the usefulness of all 24 relations, we employed One-Sample T Test on each relation. The null hypothesis and the alternative hypothesis were defined as below where μ represents the population mean:

H0: $\mu 1 \leq \mu$
H1: $\mu 1 > \mu$

With the level of significance (α) set as 0.05, the confidence interval of the difference between μ and the sample mean for each relation is summarized in Table 4.

Table 4: The Results of One-Sample T Tests

		$\mu = 3$		$\mu = 4$	
		95% Confidence Interval of the Difference			
Variable	Mean	Lower	Upper	Lower	Upper
demonstration	4.43	1.18	1.69	0.18	0.69
example	4.43	1.14	1.72	0.14	0.72
illustration	4.33	1.00	1.66	0.00	0.66
conclusion	4.23	0.91	1.55	-0.09	0.55
definition	4.23	0.90	1.57	-0.10	0.57
process	4.23	0.98	1.49	-0.02	0.49
instance	4.10	0.83	1.37	-0.17	0.37
procedure	4.03	0.70	1.37	-0.30	0.37
introduction	3.87	0.55	1.19	-0.45	0.19
continues	3.80	0.40	1.20	-0.60	0.20
theorem	3.80	0.44	1.16	-0.56	0.16
counterexample	3.70	0.34	1.06	-0.66	0.06
restriction	3.67	0.30	1.04	-0.70	0.04
deepen	3.63	0.26	1.01	-0.74	0.01
remark	3.63	0.22	1.04	-0.78	0.04
fact	3.57	0.19	0.94	-0.81	-0.06
proof	3.57	0.18	0.96	-0.82	-0.04
extension	3.53	0.13	0.93	-0.87	-0.07
evidence	3.50	0.14	0.86	-0.86	-0.14
explanation	3.50	0.14	0.86	-0.86	-0.14
guideline	3.40	0.04	0.76	-0.96	-0.24
law	3.40	0.07	0.73	-0.93	-0.27
lawofnatural	3.30	-0.04	0.64	-1.04	-0.36
opposition	3.20	-0.26	0.66	-1.26	-0.34

In Table 4, the second column is the sample mean; the third and fourth columns are the lower and upper bounds, respectively, of the confidence interval when μ is 3; and the fifth and sixth columns are the lower and upper bounds, respectively, of the confidence interval when μ is 4.

Take "Demonstration" as an example. Because its confidence interval is [1.18, 1.69], H_0 should be rejected. In other words, the usefulness of the relation "Demonstration" is significant. We can conclude that authors believed that using the relation "Demonstration" in LOs will increase learning effectiveness. Therefore, as shown in Table 4, the usefulness of all relations, except for "Law of Nature" and "Opposition", is significant when μ is 3. Furthermore, when μ is 4, only "Demonstration", "Example", and "Illustration" are significant.

In addition, we like to study if there is a significant difference between authors who have teaching experience and those who have no teaching experience. Independent-Samples T Test was employed to study the difference between two groups. The means, *t*-values, and *p*-values are calculated for each relation and summarized in Table 5. In Table 5, the second and third columns represent the numbers of the surveyed graduate students who have teaching experience and those who have no teaching experience, respectively; the fourth and fifth columns are the sample means for the graduate students who have teaching experience and those who have no teaching experience, respectively; and the sixth and seventh columns are *t*-values and *p*-values, respectively, of two groups.

Table 5: The Results of Independent-Sample T Tests

Teaching Experience	Num.		Mean		t-value	p-value
Variable	Yes	No	Yes	No		
conclusion	17	13	4.06	4.46	-1.35	0.19
continues	17	13	4.12	3.38	1.95*	0.06
counterexample	17	13	3.71	3.69	0.04	0.97
deepen	17	13	3.59	3.69	-0.28	0.78
definition	17	13	4.29	4.15	0.42	0.68
demonstration	17	13	4.47	4.38	0.34	0.73
evidence	17	13	3.29	3.77	-1.38	0.18
example	17	13	4.41	4.46	-0.17	0.86
explanation	17	13	3.71	3.23	1.30	0.21
extension	17	13	3.65	3.38	0.64	0.53
fact	17	13	3.53	3.62	-0.23	0.82
guideline	17	13	3.65	3.08	1.77*	0.09
illustration	17	13	4.35	4.31	0.14	0.89
instance	17	13	4.12	4.08	0.16	0.88
introduction	17	13	3.71	4.08	-1.21	0.24
law	17	13	3.35	3.46	-0.35	0.73
lawofnatural	17	13	3.06	3.62	-1.76*	0.09
opposition	17	13	3.29	3.08	0.46	0.65
procedure	17	13	4.18	3.85	1.03	0.31
process	17	13	4.29	4.15	0.55	0.58
proof	17	13	3.53	3.62	-0.23	0.82
remark	17	13	3.65	3.62	0.07	0.94
restriction	17	13	3.76	3.54	0.62	0.54
theorem	17	13	3.53	4.15	-1.89*	0.07

* The mean difference is significant at the 0.1 level.

From the table, it shows that there is no significant difference between two groups if the level of significance is 0.05. However, if the level of significance is 0.1, two groups are significantly different on the relations “Continues”, “Guideline”, “Law of Nature”, and “Theorem”. Because “Law of Nature” is not considered as useful from the previous analysis, only the relations “Continues”, “Guideline”, and “Theorem” were further studied. The ratios shown in Figure 8 are computed by dividing the number of authors in a group who believed a relation is either ‘highly useful’ or ‘useful’ by the total number of authors in the group. Take the relation “Theorem” as an example. The number of authors who have teaching experience and believed the relation is either ‘highly useful’ or ‘useful’ is 9. When divided by the total number of authors in the group which is 17, the ratio is 0.53. From the figure, it is clear that, authors who have teaching experience preferred both “Continues” and “Guide-line”, while authors who have no teaching experience preferred “Theorem”.

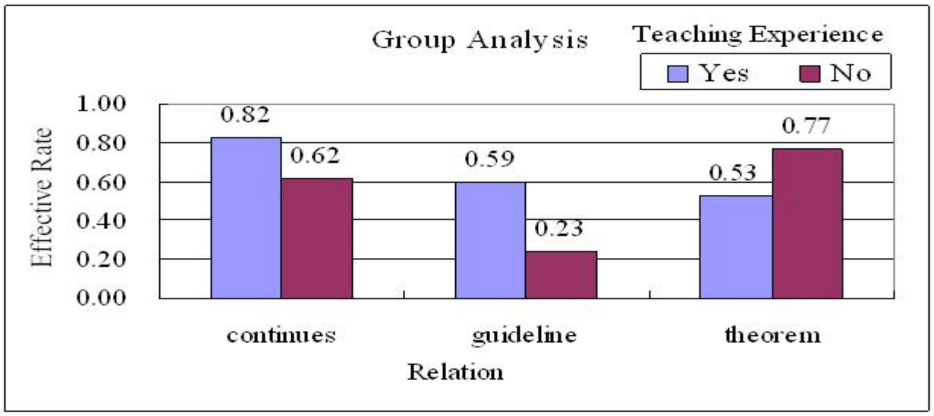

Figure 8: Group Analysis

Table 6: Association Analysis (The Minimum Support Value=3%)

	correlative relation	support %
(1)	example instance	8.5
	example illustration	4.9
	example illustration instance	4.9
	illustration instance	4.9
(2)	law lawofnatural	7.2
	law theorem	5.4
	lawofnatural theorem	4.7
	law lawofnatural theorem	4.7
(3)	procedure process	6.8
	continues process	4.2
	continues procedure	3.8
	continues procedure process	3.8
(4)	counterexample restriction	3.6
(5)	deepen extension	3.2
	total of transactions = 720	

Analysis: Part II

It is believed that, when specifying a relation for a LO, authors may have difficulty in choosing one relation over another. Therefore, the main purpose of the second part of the questionnaire was to examine whether or not authors would be confused when selecting an appropriate relation for LOs. Association rule mining (Agrawal and Srikant, 1994; Agrawal et al., 1993), a popular technique used in data mining, was used to examine associations or correlation relationships among the 24 relations. In particular, Apriori algorithm (Agrawal and Srikant, 1994) was employed.

When specifying a relation for a LO, authors were asked whether or not they might use other relations for the LO. Thus, the answer for one relation and the relation itself can be treated as a *transaction* (or called *itemset*). For example, if an author selected nothing for “Conclusion” and selected both “Procedure” and “Process” for “Continues”, there were two transactions which were shown as follows:

Conclusion
Continues Procedure Process

Because there were 24 questions for each author, there were 720 transactions in total. A program developed by Borgelt (Borgelt, 2007) was used to analyze transactions. With the *support* value as 3%, the results are shown in Table 6.

From the results, it is clear that relations which have high associations can be categorized into five groups, and they are (1) “Example”, “Illustration”, and “Instance”; (2) “Law”, “Law of Nature”, and “Theorem”; (3) “Continues”, “Procedure”, and “Process”; (4) “Counterexample” and “Restriction”; and (5) “Deepen” and “Extension”.

Discussions

Based on the previous analysis, the following suggestions were made. However, please bear in mind that both the surveyed authors and the studied learning materials are closely related to information technology. The suggestions below may not be applied to other domains.

First of all, both “Law of Nature” and “Opposition” were considered useless, while “Demonstration”, “Example”, and “Illustration” were considered the most important relations. It is, therefore, suggested that both “Law of Nature” and “Opposition” can be removed from the relation metadata.

Secondly, in the analysis of correlation relationships among relations, it is shown that some relations are highly correlated and are clustered into five groups. In the followings, each group will be discussed further.

Example, Illustration, and Instance

Based on its original definition in RST, an instance is an example of a learning object. In addition, because “Example” and “Instance” are highly correlated, either one can be removed from the relation metadata. It is suggested that “Instance” is removed from the metadata.

The meanings of “Illustration” and “Example” are somewhat overlapped and sometimes cannot be clearly distinguished one from the others. According to OxFord Advanced Learner’s Dictionary, an illustration is either “a drawing or picture in a book, magazine, etc. especially one that explains something”, “the process of illustrating something”, or “a story, an event or an example that clearly shows the truth about something”. For the second case, we can use the relation “Process” to describe LOs. For the third case, we can use the relation “Example” to describe LOs. Therefore, in the relation metadata, an illustration is re-defined as “a drawing or picture that explains something”.

Law, Law Of Nature, and Theorem

Because “Law of Nature” was removed from the metadata, it will not be discussed again. As shown in Table 4, the surveyed authors believed the usefulness of “Theorem” is higher than the usefulness of “Law”. As a result, it is suggested that “Law” can be removed from the metadata. However, it is noted that we believed the main reason why both “Law” and “Law of Nature” are considered less useful is simply because they are rarely used in IT domain. It may not be true in other domains.

Continues, Procedure, and Process

All three relations are highly correlated. From Table 4, it showed that “Process” is considered more useful than “Continues” and “Procedure”. Therefore, it is suggested that “Continues” and “Procedure” can be removed from the metadata.

Counterexample and Restriction

Although we clearly explained the definitions of “Counterexample” and “Restriction” before the surveyed authors answered questions, these two relations still confused many authors which was shown in Table 4. In order to avoid confusions, it is suggested that “Restriction” is removed from the metadata and “Counterexample” is re-defined as the unsuccessful or exceptional situation of a LO.

Deepen and Extension

A LO of type “Deepen” is to explain another LO in depth. An extension object is a LO that is extended from another LO. Because they were considered useful (with the average values higher than 3.53), and because their support value was only slightly higher than 3%, it is suggested that both relations should be kept in the metadata. Also, they can be considered as “optional”.

From the above discussions, it was concluded that 7 relations (including “Law of Nature”, “Law”, “Opposition”, “Instance”, “Continues”, “Procedure”, and “Restriction”) were removed and 2 relations (including “Illustration” and “Counterexample”) were re-defined. The proposed extended relation metadata consists of 17 relations and are summarized in Table 7. Out of 17 relations, it is highly recommended to define “Demonstration”, “Example”, and “Illustration” LOs. As shown in Table 7, redefined relations were italicized and highly recommended relations were bolded.
Finally, although the survey was carefully and systematically conducted, it was limited by several factors. For one, the sample size is small because the survey had to be conducted in one-to-one manners. Also, as mentioned earlier, the results were focus on IT domain. Other factors may include the proficiency of English of the surveyed authors.

Table 7: The Proposed Relation Metadata

Conclusion	*Counterexample*	Deepen	Definition	**Demonstration**
Evidence	**Example**	Explanation	Extension	Fact
Guideline	***Illustration***	Introduction	Process	Proof
Remark	Theorem			

Conclusions and Future Works

It is believed that the reusability of learning objects can be significantly increased if metadata is employed properly. In this paper, relation metadata were thoroughly studied and extended relation metadata were proposed. Additionally, a prototype of learning content management system was designed and implemented. Using the learning content system that support the proposed metadata, many authors believed the learning effectiveness can be increased.

In SCORM CAM, metadata are divided into 9 categories. Currently, only one category (i.e., relation) was studied. The utilization of other 8 categories has not been fully explored and may worth further investigation. In addition, the current search mechanism used in the prototype is primitive. It is worth further investigation in the design of search mechanisms by utilizing metadata (Lu and Jung, 2003) to increase both performance and search precision.

Acknowledgements

This research was partially supported by the National Science Council, Taiwan, R.O.C., under contract no.: NSC94-2213-E-005-037 and NSC95-2221-E-005-138.

References

Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *Paper presented at the ACM SIG-MOD Conference on Management of Data*, June 9-12, San Diego, California, USA.

Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. *Paper presented at the 20th International Conference on Very Large Databases*, September 12-15, Santiago, Chile.

Borgelt, C. (2007). *Apriori - Association rule induction / frequent item set mining*. Retrieved October 19, 2009, from http://www.borgelt.net//apriori.html.

Cover, R. (2006). *SGML/XML: Using element and attributes*. Retrieved October 19, 2009, from http://www.oasis-open.org/cover/elementsAndAttrs.html.

Fischer, S. (2001). Course and exercise sequencing using metadata in adaptive hypermedia learning systems. *ACM Journal of Education Resource in Computing*, *1*(1), 1-21.

Fischer, S., Steinacker, A., Seberg, C., & Steinmetz, R. (1999). Multibook: Metadata for web-based learning systems. *Paper presented at the 2nd International Conference on New Learning Technologies*, August 30-31, Bern, Switzerland.

Karger, P., Ullrich, C., & Melis, E. (2006). Integrating learning object repositories using a mediator architecture. *Paper presented at the First European Conference on Technology Enhanced Learning*, October 1-4, Crete, Greece.

Loser, A., Grune, C., & HoRmann, M. (2002). *A didactical model, definition of learning objects and selection of metadata for an online curriculum*, Retrieved January 2, 2010, from http://www.ibi.tu-berlin.de/diskurs/veranst/online_educa/oeb_02/Talk_Online_Educa__02_Loeser_TU_berlin.pdf.

Lu, E. J.-L. & Jung, Y.-M. (2003). XDSearch: An effcient search engine for XML document schemata. *Expert Systems with Applications*, *24*(2), 213-224.

Lu, E. J.-L., Wu, B.-C., & Chuang, P.-Y. (2006). An empirical study of XML data management in business information systems. *Journal of Systems and Software*, *79*(7), 984-1000.

Mann, W. C. & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, *8*(3), 243-281.

Reigeluth, C. (1999). *Instructional-design theories and models, Volume II: A new paradigm of instructional theory*. Mahwah, NJ: Lawrence Erlbaum.

Sddik, A. E., Fischer, S., & Steinmetz, R. (2001). Reusable multimedia content in web-based learning systems. *IEEE Multimedia*, *8*(3), 30-38.

Seeberg, C., Saddik, A. E., Steinacker, A., Reichenberger, K., Fischer, S., & Steinmetz, R. (1999). From the users' needs to adaptive documents. *Paper presented at the 4th World Conference on the Integrated Design and Process Technology*, June 27 - July 2, Kusadasi, Turkey.

Steinacker, A., Faatz, A., Seeberg, C., Rimac, L., Hormann, S., Saddik, A. E., & Steinmetz, R. (2001). Medibook: Combining semantic networks metadata for learning resources to build a web based learning system. *Paper presented at the Ed-Media Conference*, June 25-30, Tampere, Finland.

Steinacker, A., Seberg, C., Fischer, S., & Steinmetz, R. (1999). Multibook: Metadata for webbased learning systems. *Paper presented at the 2nd International Conference on New Learning Technologies*, August 30-31, Bern, Switzerland.

Ullrich, C. (2003). An instructional component for dynamic course generation and delivery. *Paper presented at the Berliner XML Tage*, October 13-15, Berlin, Germany.

Ullrich, C. (2004). Description of an instructional ontology and its application in web services for education. *Paper presented at the Workshop on Application of Semantic Web Technologies for E-learning*, November 8, Hiroshima, Japan.

Ullrich, C. (2005). The learning-resource-type is dead, long live the learning-resource-type. *Learning Objects and Learning Designs*, *1*(1), 7-15.

Ullrich, C. (2006). *Discuss forum for the ontology of instructional objects*, From http://forum.activemath.org/viewtopic.php, this discuss forum was closed.

Appendix A. The Relations Defined in Instruction Design Theory

The twenty-three relations defined in IDT are described as follows:

Definition
A definition is used to describe the meaning of a word, a phrase, a symbol, or a proper noun which appears in LOs. For example, e-learning by definition is to engage in learning activities by utilizing information technology and the Internet.

Fact
A fact is an event that happened. For instance, "Tim Berners-Lee is the creator of HTML" is a fact.

Law
A law describes a general principle that can be found in natural phenomena or statements that have been proven to be true. The "Law" class also consists of two sub-classes: "Law of Nature" and "Theorem". For example, Moore's Law stated that the number of transistors on an integrated circuit doubles approximately every eighteen months, but the price reduces by one half.

Law of Nature
A learning object of type "Law of Nature" describes a general rule observed in nature. For example, Newton's Laws of Motion is a law of nature in the physics.

Theorem
A theorem is a concept that has been shown to be true. Bayes theorem, for instance, is a theorem.

Process
The "Process" class includes two sub-classes: "Policy" and "Procedure". A process is a flow of events that describes how a task can be accomplished in steps. For example, software development life cycle (SDLC) describes the steps how information systems can be developed.

Policy
A policy describes a set of predefined principles of actions. It is usually composed of informal suggestions or guidelines for specific activities. For example, interview is a method for system analysts to obtain system requirements. A good policy for interview is to confirm time with the interviewer and then make sure the she comprehends the subject in advance. Based on the definition, a policy is similar to a guideline. Ullrich, the creator of the instructional ontology, also said that the definition of policy is not crystal clear and thus suggested one can make changes if necessary (Ullrich, 2006). Therefore, we renamed "Policy" to "Guideline".

Procedure
A procedure is a sequence of steps that can accomplish a goal. An algorithm of, for example, bubble sort is a procedure.

Interactivity
A learning object of type "Interactivity" is some kind of activities that allow learners to practise or develop a skill interactively. The "Interactivity" class consists of four sub-classes: "Exploration", "Real World Problem", "Invitation", and "Exercise". They have been defined in the "EDUCATION" category in SCORM.

Illustration
An illustration is to illustrate a concept or parts of a concept of a learning object. For example, LO_B and LO_C shown in Figure 1 are two illustration objects for the concept of bubble sort.

Example
An example is an auxiliary learning object that is used to further explain parts or the whole of a fundamental learning object.

Counterexample
In IDT, a counterexample is not an example of a fundamental learning object, but it is often mistakenly thought of as one. For example, a parallelogram is often mistakenly treated as a rectangle. Therefore, parallelogram can be used as a counterexample for a learning object that describes rectangle.

Evidence
An evidence is a learning object that supports the claims made for a law or any learning object of its subclasses. The "Evidence" class includes two sub-classes: "Proof" and "Demonstration". For instance, the time complexity for bubble sort is $O(n^2)$, and the time complexity for quick sort is $O(n \log n)$. Thus, this is an evidence that quick sort is faster than bubble sort.

Proof
A proof is an evidence that is derived formally or mathematically to support a law.

Demonstration
A demonstration is used to demonstrate, in general through experiments, that a law holds under a certain condition. For example, Galileo's experiment showed that falling objects of different weights landed at the same time.

Explanation
A learning object of type "Explanation" provides extra information for a fundamental learning object so as to highlight its important properties. The "Explanation" class includes three sub-classes: "Introduction", "Conclusion", and "Remark".

Introduction
A learning object of type "Introduction" provides the bird's-eye view of a fundamental learning object so that learners have a rough idea what will be covered in a learning resource.

Conclusion
A learning object of type "Conclusion" summarizes key points covered in a fundamental learning object.

Remark
A remark provides extra but inessential information for a fundamental learning object. For instance, when a learning topic is about entity relationship diagram (ERD), an example remark can be "ERD is similar to the class diagram in Unified Model Language (UML). ERD is mainly used in structured analysis and design, while UML is used in object-oriented analysis and design."

Appendix B. The Relations Defined in Instruction Design Theory

The nine relations defined in RST are described as follows:

Example and Illustration
The definitions of “Example” and “Illustration” are identical to the “Example” and “Illustration”, respectively, defined in the previous section.

Instance
If a learning topic LT_A is sorting, it is then said that bubble sort is an instance of LT_A.

Restriction
A restriction describes cases where a certain theory fails. For example, in 1640s, Fermat stated that all Fermat numbers are prime numbers. The formula of Fermat numbers is:

$$F_n = 2^{2n} + 1 \qquad (1)$$

However, Euler found that Fermat number is not a prime number when n is 5 in 1732.

Amplify/Extension
An amplify or extension object is a learning object that is extended from another learning object. For example, semantic web has grown out of the traditional web.

Continues
A learning object of type “Continues” describes the sequence relationship between two learning objects where one is performed after the other. For example, LO_A represents data before sorting, and LO_B represents the data after sorting. Then, LO_B continues LO_A.

Deepen/Intensification
A deepen object provides information for another learning object in depth. For example, LO_A describes how greatest common divisor (GCD) is obtained, and LO_B describes in details the reasons why the process described in LO_A can obtain GCD. Then, LO_B deepens LO_A.

Opposition
An opposition describes a statement proposed by a specialist that is contradicting another statement made by another specialist. For example, when designing XML documents, some experts suggested avoid using attributes can reduce processing time (Cover, 2006). Still, some experts stated that it can shorten processing time by using attributes (Lu et al., 2006).

Alternative
A learning object of type “Alternative” describes a thing that has been explained in another learning object but in different format. For example, LO_A describes bubble sort in text, and LO_B describes bubble sort in animation. Then, LO_B is an alternative to LO_A. However, as stated earlier, the relation “hasformat” is already defined in the “RELATION” category. Thus, the relationship “LO_A Alternative LO_B” can be replaced with the relationship “LO_A hasformat ANIMATION in LO_B”.

A Study of the Efficacy of Project-based Learning Integrated with Computer-based Simulation - STELLA

Rogheyeh Eskrootchi and G. Reza Oskrochi[1*]
School of Management and medical Information Sciences, Iran University of Medical Sciences & Health services, Tehran, Iran // Eskrootchi@iums.ac.ir
[1]School of Technology, Oxford Brookes University, Oxford, UK // roskrochi@brookes.ac.uk
[*]Contact author

ABSTRACT

Incorporating computer-simulation modelling into project-based learning may be effective but requires careful planning and implementation. Teachers, especially, need pedagogical content knowledge which refers to knowledge about how students learn from materials infused with technology. This study suggests that students learn best by actively constructing knowledge from a combination of experience, interpretation and structured interactions with peers and teachers when using technology. Simulations do not work on their own, there needs to be some structuring of the students' interactions with the simulation to increase effectiveness. The purpose of this study was to investigate the effectiveness of project-based learning in a technology-rich environment. A science project, Land-use in Watershed, that takes advantage of Internet facilities was developed and integrated with a simulation software package, Structural Thinking and Experiential Learning Laboratory, with Animation, (STELLA) developed to promote deeper understanding of Land-use by students. The Participants in the study were 72 students in a quasi-experimental research design. Statistical analyses showed that students who participated in the manipulation of the experimental model of the watershed experiment and the STELLA simulation performed best on understanding the watershed concept.

Keywords

STELLA, computer-assisted simulation, learning technology, watershed concepts and modelling, project based learning

Introduction

An often-stated belief is that transferring skills is the main job of education. However, an increasing body of research shows that the way knowledge is presented to students in school and the kinds of operations they are asked to perform often result in students knowing something but failing to use it when relevant. Brown, Collins and Duguid (1989) believed that classroom activities lack the contextual features of real-life problem-solving situations and therefore weaken the ability of students to transfer and apply their knowledge from the school setting to the outside world. The challenge as Santos-Trigo and Camacho-Machín, (2009) purposed is to question; “Can routine problems be transformed into problem solving activities that promote students' mathematical reflection?”

Studies suggest that in order to facilitate transfer, promote effective learning and encourage a high degree of ownership and personal relevance, educators should provide training on real tasks. Similarly, researchers believe that cases and examples must be studied as they really occur, in their natural contexts, not as stripped-down ‘textbook examples’ that conveniently illustrate some principle (Blumenfeld et al., 1991). The extent at which the process of solving textbook problems can help students develop a way of thinking to be consistent with mathematical practice is still under an investigation (Santos-Trigo & Camacho-Machín, 2009).

The National Research Council’s (NRC, 1999) standards for science education suggest long-term inquiry activities including argumentation and explanation, communicating ideas to others and using a wide range of manipulative, cognitive and procedural skills will promote learning. The standards suggest that to develop their understanding, students need to relate new information to existing knowledge and build connected networks of concepts. In addition, NRC (1999) called for the teaching of “fluency with information technology” and strongly recommended the use of technology to promote understanding of science and mathematics.

Theorists and educators are promoting reality-centred projects and other reality-centred activities as ways to engage students in meaningful learning. Experienced educators tend to agree that students learn best through a project-based approach in which they are able to discover things for themselves and take advantage of technological tools (Blumenfeld et al., 1991; Clinchy, 1989; Linn, et al., 2000; Lebow, & Wager, 1994).

While technology can be valuable in supporting students and teachers in projects requiring higher level thinking (Blumenfeld et al., 1991), it is not the kind of technology that matters most, but rather how it is used (Dyrli & Kinnaman, 1994; Ehrmann, 1995; Green & Gilbert, 1995). The intention of the researcher in this study was to examine and discuss ways in which learners engage in an intentional learning process and analyze the effectiveness of such approach in two areas: a) an authentic learning activity in the content of project-based learning; and b) educational simulations.

Project-based learning

Project-based learning (PBL) blends traditional subject-matter goals and objectives with authentic learning environments. The primary rationale for using authentic activity as the model for appropriate learning activities is the enhanced understanding that develops through application and manipulation of knowledge within context. Finding solutions to a problem whether posed by the teacher or a new social environment, more likely develop generic, as well as subject specific skills when using project-based curriculum. In other words, PBL provides productive environments for the development of meta-cognition (Downing et al., 2009). In addition, in another study conducted to identify dental students' self-reported sources of stress, the findings revealed that PBL when compared to traditional curricula was inversely associated with perceived stress and that in turn had a strong impact on learning (Polychronopoulou & Divans, 2009). Nevertheless, transformation of the conventional classroom into an authentic learning environment involves much more than incorporating features of real-life situations into school work. Furthermore, curriculum innovations are never easy to implement or to examine systematically. Balasooriya et al., (2009) carried out a study on the impact of a new integrated medical educational design on students' approaches to learning. Although, the program was based on curriculum features identified in the research literature to promote deeper approaches to learning, the results indicate shifting students towards deeper approaches to learning may be a more complex task than previously understood.

Krajcik et al., (1994) suggest that there are five features of PBL that help communicate the complexity of the innovation in terms that are familiar to teachers. These are driving questions, investigations, artefacts, collaboration and technological tools. This approach can be supported by multimedia and network technologies such as the Internet. The introduction of microcomputers into classrooms has generated innumerable instances of such innovations which involve considerable change in classroom management, lesson structure and student assessment. Powerful hardware and sophisticated software tools are enabling people to become more active learners about their environment (Jackson et al., 1997).

Research has shown that students can make significant gains when computers are incorporated in the learning process. Computer-based technologies integrated in project-based learning are particularly useful for constructive learning (Roschelle et al., 2000). Students instantaneously can see the results of their experiment. Computer technology supports learning; it is especially useful in developing the higher-order skills of critical thinking, analysis and scientific inquiry (Roschelle et al., 2000). But the mere presence of computers in the classroom does not ensure their effective use. Many factors influence how and who learns in the classroom: (1) active engagement, (2) participation in groups, (3) frequent interaction and feedback and (4) connections to real-world contexts. Omale et al., (2009) performed a study to investigate how the attributes of 3-D technology affect participants' social, cognitive and teaching presences in a PBL environment. The results indicated that although the attributes of 3-D technology promoted participants' social presence, additional technical and instructional features of the 3-D environment were required to further enhance cognitive and teaching presence leading to overall learning experience. Some of the pioneers in learning research believe computer simulation when used effectively has the potential to address these related factors (Blake & Scanlon, 2007).

Educational Simulations

Simulations as defined by Alessi and Trollip (2001) are a representation of some phenomenon or activity that users learn about through interaction with the simulation. Simulations offer an easy way of controlling experimental variables, opening up the possibility of exploration and hypothesizing. Simulation is also valuable in presenting many types of representational formats including diagrams, graphics, animations, sound and video that can facilitate understanding.

Pea's framework of distributed intelligence suggests that computer-assisted simulations have the potential to reorganize mental processes by closing the temporal gaps between thought and action and between hypothesis and experiment. Pea has proposed that by allowing the user to engage in "what-if thinking" through a partnership between user and technology, deep qualitative effects are made possible on how problem solving occurs (Lebow & Wager, 1994). From an instructional design perspective, educational simulations support predetermined learning outcomes by providing participants with opportunities to deal with the consequences of their actions and to respond to feedback. In other words, the simulation construction kit is a laboratory for scientific inquiry, for exploration, explanation and testing. STELLA, which stands for Structural Thinking and Experiential Learning Laboratory, with Animation, seems to facilitate this disciplined approach to inquiry (Steed, 1992). Understanding how a simulation construction kit, like STELLA, can be used to refine thinking is important. Simulation models are simplified representations of real-world systems over hypothetical time. Using simulation software, characteristics of selected variables can be altered and their effects on other variables and the entire system assessed (Steed, 1992). STELLA is a program designed to assist users in creating their own simulations using system dynamics. One needs to think in terms of dynamic processes, positive and negative causal loops, flows, accumulation and converters. STELLA is one technology that can enable individuals to enhance their understanding of and appreciation for the complex web of interrelationships that govern environmental behaviour (Peterson, 1985). The real value of the STELLA modelling package is the cognitive processing that goes on in the creation and development of its model. Good science is good questions. Through creating simulations one has to generate good questions and as the simulation evolves interesting inquires are naturally pursued.

The simulation modelling generally takes two forms. Depending on the courses, students are (a) required to develop their own models of scientific phenomena, or (b) given existing models and are asked to alter particular parameters to examine the subsequent effects on the entire system. These two distinct approaches to modelling are likely to produce different cognitive outcomes in terms of content knowledge and general problem-solving skills. In this paper, we are considering interactive simulations (Blake & Scanlon, 2007) that allow students to change some of the parameters in the program and observe what happens as a result. Although theory supports using technology to engage students in project-based learning, and the literature provides descriptions of suitable classroom technology to engage students, we have few case studies of middle-school teachers describing their development and effective use of simulation in educational learning (Blake & Scanlon, 2007 and Steed, 1992). In fact, with the re-emergence of experiential learning as a dominant model of learning in education and the recent research on infusing information technologies into classrooms, it is a good time to examine the effectiveness of project-based learning integrated with computer-based simulation.

The study reported here explores the effectiveness of the project-based learning which takes advantage of simulation experiment. The following research question was proposed: How effective is the reality-based project in engaging students in meaningful learning and enhancing their motivational attitude, especially when integrated with computer-based simulation?

A science project, Land-use in Watershed, funded by the National Science Foundation (NSF) was developed to investigate this research question. The project was integrated with STELLA simulation software to enhance deeper understanding for students. Land-use in Watershed was a collaborative science project which was developed by the researcher for the Kansas Collaborative Research Network (KanCRN).

Method & Aims

Procedure & Sample

The participants of this study were 72 sixth to eighth graders, (32 males and 40 females). All students attended Northwest Middle-school, Kansas city, Kansas at the time of the study. Three separate multi-age classrooms were included in the study. The multi-age classrooms were randomly allocated to one of the three treatment groups. Students from all groups read the project materials on-line through KanCRN and benefited from on-line learning features such as collaboration with peers, finding definitions of related terminologies and using hyperlinks to additional information. Teaching material was delivered by the same teacher. Students were pre-tested with respect to their understanding of the watershed concept and their content knowledge. The first treatment group (n_1=19) the Project-Based group (PB), were taught the subject by receiving a traditional lecture. The second treatment group (n_2=33), the Project-Based Experimental Simulation group (PBES), were taught the subject by performing an experimental model and a simulation model. The third group n_3=20, the Project-Based Simulation group (PBS), performed a simulation model but not an experimental model. Students were further divided into sub-groups of three or four within each classroom based on their pre-test score in order to improve learning while obtaining homogenous groups.

The researcher prepared a lesson on the "effect of land-use on the watershed" in which she designed an experimental model of the watershed using sponge and cardboard. Two simulation applications using STELLA software were also designed by the researcher to further emphasize the concept of watershed and in particular the effect of land-use on runoff. The first application represented the experimental model and was created using experimental data obtained from the sponge-cardboard experiment (STELLA1). The second application was more advanced and was created by real data from a watershed (STELLA2). The simulation models played a role as a supplemental tool for practicing "what if scenarios" by manipulating interacting factors and understanding the impact of important parameters on the watershed.

Hypotheses

The first null hypothesis (H_{01}) is: gain in students' content knowledge is the same between groups. The second null hypothesis (H_{02}) is: students' comprehension knowledge is the same between groups.

The third null hypothesis (H_{03}) is: students' attitude towards the project is the same between groups.

Instruments

A 58-question student survey was used to collect four types of information:

I. Eight true/false questions; used in statistical analysis to test gain in content knowledge.
II. Seven open-ended questions to test students' understanding of the watershed concept.
III. Twenty-three multiple-choice questions that tests the students' attitude toward the project.
IV. Five true/false questions that gathered information about students' computer background knowledge.

In addition, information about the amount of time spent on reading the project on-line, type and amount of small-group interaction and number of requests for help in different parts of the project were collected by observation.

Results

The design of the study included three independent factors: gender (male or female), grade (sixth, seventh or eighth), treatment (PB, PBES or PBS) and computer background as a categorical covariate on the 5-points Likert scale.

The dependent variables consisted of three measures: content knowledge; comprehension of the subject; students' attitudes towards the project.

The results of a one way ANOVA test failed to reject H_{01} (Pvalue >0.05) and therefore conclude that the gain in students' content knowledge is not significantly different between groups (Table 1). In contrast H_{02} was rejected at 5% significance level, hence students' comprehension knowledge is significantly different between groups (Pvalue = 0.002). A Post Hoc test result indicated that students in the PBES (mean = 8.61, SD = 2.42) treatment groups outperformed the other two groups (PBS; mean=5.45 SD=3.78 & PB; mean=4.16 SD=2.52) on subject comprehension.

Table 1: Mean change and significance level of comprehension knowledge between treatments

Treatment (I)	Treatment (J)	Mean (I-J)	Sig.
Project-Based Experimental Simulation	Project-Based Simulation	**3.162**	**.002**
	Traditional	**4.450**	**.000**
Project-Based Simulation	Traditional	1.294	.692

Based on observed means. The mean difference is significant at the .05 level.

Second, analysis of interaction between gender and treatment showed that there were no significant interaction effects between the PB and PBES groups within the male category while the females in the PBES group performed significantly higher than the females in the PB group. This interaction between gender and treatment variable indicated that the project had a stronger effect on females and led to the higher mean score for the PBES group in comparison to the PB group.

Further analysis suggested that the gain (value added) in students' comprehension of the watershed concept was independent of the students' grade-level.

Students' attitudes toward the project, especially toward STELLA simulation, were promising (85%). Students who experienced STELLA units and enhanced their understanding about the watershed through this activity found the reading part of the project that they encountered prior to the STELLA lesson less difficult.

The analysis of data on students' computer background indicated that only 2% of the students have never used a computer before, 61% of students use a computer at home, 76% of students had previous experience with computer simulation, but 80% of students had never worked with STELLA.

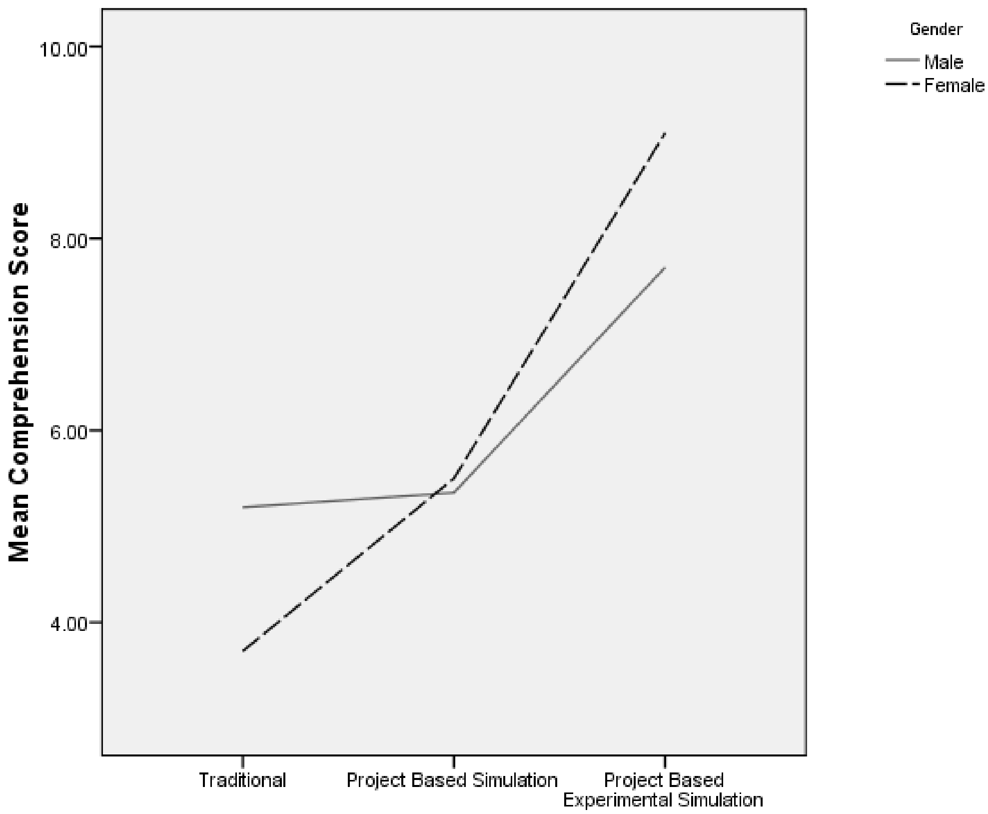

Figure 1: Students' comprehension gains by treatments for male and female

Discussion

This study emphasizes on two important and interesting findings. First result indicated that although the PBES performed significantly higher than the PB class, the PBS group which used computer-based simulation but did not perform the experimental part of the project (Sponge-Cardboard model) and consequently could not benefit from the project to its full potential did not score higher than the PB class. According to NCTM, (Linn et al., 2000, pp. 25) this would further emphasize that “technology should not be used as a replacement for basic understandings and intuitions; rather, it can and should be used to foster those understandings and intuitions”. Furthermore, the PBES groups outperformed the PBS as well. However there was no significant difference between the two classrooms treated with PBES.

Simulation and its Role

An explanation of the high performance of the PBES group could be due to the experience they gained first through the experimental model of the watershed which generated data for modelling the first STELLA application (STELLA1), and enhanced deeper appreciation of a more advanced and complicated model (STELLA2). Many studies have identified the benefits of combining technology and other materials, such as construction kits or hands-on experiments for projects to make projects successful (Linn et al., 2000). For example modelling environments such as Model-It combine the strengths of technology with opportunities for personal data collection. Technology can support and encourage those projects and activities that are more meaningful to students (Roschelle et al., 2000). Research indicates that simulations can play an important role in science learning, however their use is not as straight-forward as it seems. Many researchers emphasize the importance of a good instructional plan when using simulations. They argue that despite their popularity in instruction, the lack of support for learners in some simulations is the reason for finding no conclusive evidence for effectiveness and efficiency of simulations (Blake & Scanlon, 2007).

PBL and Interactions

In the PBES treatment group there is more opportunity for collaboration during the experiment and simulation, hence more interaction between students, and also with the teacher and the model compared to the other two methods. This finding also supports Schutte’s (1997) suggestion that the enhanced levels of interaction with other students and the teacher results in greater efficacy of computer-mediated communications. Another study demonstrates that all learning activities such as constructive, self-directed and collaborative learning occur as a result of verbal interactions through PBL environment (Yew & Schmidt, 2009).

Two factors explain reasons for improvement of problem-solving skills in a group work (Teaching Professor, 2009). One is the knowledge that students acquire from other students' explanations and another is the students' involvement to solve the problem causes them to think more deeply about the problem and its solution. Other findings report that successful innovative programs improve instruction for all students and have the greatest impact on those who are at greatest risk for learning less (Linn et al., 2000). Chu et al., (2009) indicate that in spite of most e-learning platforms that offer theoretical knowledge content, a problem-based e-learning (PBeL) model which incorporates the PBL theory, social constructivism, and situated learning theories assist teachers in effectively developing practical knowledge for mathematics teaching for students with mild disabilities.

On the other hand, according to Linn et al., (2000), adding technology to science instruction has the danger of increasing biased stereotypes and promoting the idea that these are male domains; however, if used effectively, technology can connect science to problems that interest individuals who have been assumed under-represented in science careers. As Linn et al., (2000) indicate, although there is often a belief that male students may make greater usage of information technology, some studies suggest that females and members of various cultural groups who have fewer opportunities for learning science and mathematics in the elementary and middle school in traditional practice can benefit from new technological innovations where students work in groups, and therefore males and females have equal contributions to the discussion. Likewise, these reports support the second finding of this study. The interaction between gender and treatment variable indicated that the project had a stronger effect on females and led to the higher mean score for the PBES group in comparison to the PB group. In contrast, the PBES treatment had no significant effect on the male students (Figure 1). Basically, this interesting finding supports a few researchers’ views and requires additional investigation. Figure 1 shows the average score gained for males and females by treatment group.

PBL and Critical Thinking

Overall, the findings of this study supported the use of the project-based experimental simulation on learning outcome. The result of this study indicated that the students' comprehension of the watershed concept had improved as a result of the innovative approach, although it did not improve students' content knowledge (multiple-choice questions) significantly. One possible explanation for this finding is the lack of sufficient time to cover multiple concepts in the PBES groups. While the PBES groups spent their time on various activities to comprehend a few concepts deeply, the PB group focused on direct instruction for receiving and memorizing multiple concepts. This in turn will explain why PBES students performed better in comprehension. Examination of these students' responses indicated that they achieved better and deeper understanding of the watershed concept. They were able to interpret the graphs of runoff, absorbed water in the ground and inflow correctly and in more detail. Similar results were obtained in the study by Şendağ and Ferhan (2009). This study investigated how the online PBL approach employed in an online learning environment influenced undergraduate students' critical thinking skills and content knowledge acquisition. The results indicated that learning in the online PBL group did not have a significant effect on the content knowledge acquisition scores but it had a significant effect on increasing the critical thinking skills.

Furthermore, this finding is consistent with the report provided by NSF (Linn et al., 2000) on The Middle School Math through Applications Project (MMAP), a series of project-based units that offers the most technology-intensive middle school curriculum. MMAP uses the HabiTech application, a simplified version of STELLA. According to NSF, the evaluation of this unit of MMAP was very positive, indicating that the unit met the visions of the 1989 NCTM standards, engaged students and enhanced mathematical communication in classrooms. Roschelle et al., (2000) also agreed that innovative computer-based simulation demonstrated significantly higher gains compared to those receiving only traditional instruction.

Moreover, it has been concluded that the gain in students' comprehension of the watershed concept was independent of the students' grade-level. This finding further emphasized that the multi-age computer-mediated group as a whole performed better in comprehension than the multi-age traditional group due to the treatment effect.

The promising students' attitudes toward the project, especially toward STELLA simulation, might be due to their positive perception of the whole project afterwards. The NSF's evaluation report also indicated that MMAP has positive effects on students' attitudes as well as on their performance (Linn et al., 2000).

Research Role

Research is moving ahead to introduce the computer into the classroom to ensure this technology is used effectively to enhance learning. Research should also help teachers to teach with technology rather than to use computers for personal productivity. Teachers, especially, need pedagogical content knowledge which refers to knowledge about how students learn from materials infused with technology. Successful technology use and effective learning for science teaching is dependent on the teachers' knowledge of the technology itself, and how a particular tool is best utilized for particular purposes, classroom or laboratory settings, and students themselves (Hennessy, 2006). Simulation encourages the student to interact with the variables, understand their sensitivities and appreciate how a change in one variable results in changes in other variables. However, we have shown here that the success of simulations as effective learning tools is dependent on how simulations are used. Finding ideal uses of technology in science instruction remains an active research area, and the technology itself is a "moving target," as new projects emerge on a regular basis. As Chiocchio and Lafrenière (2009) recommend, teamwork and technology are becoming important components of PBL in academic settings but fostering computer-assisted teamwork is complex and time consuming. Knowing how and when to intervene would prove useful. Finally, research demonstrates that technological tools can enhance learning in science and mathematics, in a PBL setting, since they allow more personalized and project-oriented commitments (Linn et al., 2000). According to Hakkarainen (2009), PBL offers a good model to support students' knowledge and skills, and students will benefit from learning with and about technology such as computer-based simulations in science and mathematics instruction. Nevertheless, effective incorporation of these technologies into the curriculum has been controversial, difficult and demanding.

Conclusion

In conclusion, the method by which features of project-based learning should be implemented to have its full impact on learning requires further investigation. Also as it was noted, many studies suggest that project-based simulations for visualization and modelling have transformed some fields of science and seem promising for elementary and middle school instruction. However, research indicates that incorporating computer-simulation modelling into project-based learning requires careful planning and implementation. Recent research demonstrates simulations of complex relationships, such as graphs of change over time, connections between components of a geometric construction, and location of a local minimum in a three dimensional surface can help students learn these difficult topics. At the same time, considerable research suggests that all simulations are not successful. Often simulations fail because they are too complex or too difficult to understand. In fact, in replicated studies, researchers have noted significant improvements in students' understanding of scientific concepts and motivation when using simulation software, but few studies elaborate on reasons of its failure or underling success.

In this study, the researcher administered the same simulation model to both groups PBS and PBES and conduct an experimental model only to one group (PBES) in order to demonstrate that although the structure and the design of the STELLA models were the same and presented similarly to both groups, the PBES group significantly outperformed PBS. Nevertheless, the number of interactions of students with the simulation models and themselves also increased as a result of deeper insight gained with experimental model in PBES.

This research suggests that once the learners learn the bases of using simulation, in this case through experimental model, it enables them to interact efficiently and effectively with the more complicated models. In other words, experimental model in this study helps the students to understand how primary data and flows operate in the first simulation model (STELLA1) because the model was built upon those data, therefore allowing them to build up deeper intuitions into the simulation model and its role, which in turn pursue them to appreciate more advanced subjects that is impossible to perform in natural settings, through a more complicated simulation model (STELLA2). This is an analogy for learning a new language by learning bases of vocabularies and grammatical rules which results in writing more sophisticated sentences, in comparison with learning through environment. This study also suggests that students learn best by actively constructing knowledge from a combination of experience, interpretation and structured interactions with peers when using simulation in a PBL setting. This is only possible by a careful step by step instructional design in which simulation model should solely be considered as one component of whole pedagogical structure of the PBL. The challenge is to ensure that simulation is used effectively to enhance how and what children learn. Careful planning is required for constructing a project to be purposeful and uses technology specially computer simulations such as STELLA in a constructive, real-world manner. Simulations do not work on their own, there needs to be some structuring of the students' interactions with the simulation to increase effectiveness, as it was initiated through experimental model causing to raise the quality and quantity of interactions with the STELLA models.

Hence, the first important finding of this research indicates that simulation models such as STELLA are used to expand students' experience of experimental science and should not be used as a substitution for basic understandings and intuitions; one can promote students to higher level STELLA model only when they learned the basic role of STELLA. . The second important finding advocates that in spite of common belief that the use of technology is male dominant activity**,** a few studies such as this paper suggest females and those who are at risk for learning less in traditional practice can benefit from new technological innovations. This study strongly suggests further research in this area.

In addition, it should be noted that the experimental model may only be one way to promote students' understanding of and interactions with STELLA model in this PBL. As an enhancement to this project, development and evaluation of further PBL in which students are required to create their own STELLA models (integrating developing simulations not interactive simulations) instead of using experimental models is highly recommended. Computer model developments are mirrors of one's own mental development. Model building is an interactive process; moving from identification of causal loops to computer simulation provides deep involvement in the topic and consequently deep understanding of the subject as well as STELLA mission.

Furthermore, it is recommended that further research be conducted with randomly selected participants to be truly representative for the study. Also, as this study suggests, it would be desirable to have a clear measure of effectiveness before committing to continual investment in technology.

Nevertheless, it is important to note that the results of this study are limited to the integration of simulation model, in particular STELLA, in a PBL settings and cannot be extended to all kind of educational technologies nor to other learning strategies.

Acknowledgment

We would like to acknowledge and extend our gratitude to Dr. V. Simonite and Dr. R. Long for their comments and corrections.

References

Alessi, S.M., & Trollip, S.R. (2001). *Multimedia for Learning*, Boston, MA: Allyn and Bacon.

Balasooriya, C. D., Hughes C., & Toohey S. (2009). Impact of a new integrated medicine program on students' approaches to learning. *Higher Education Research & Development, 28*(3), 289-302.

Blake, C., & Scanlon, E. (2007). Reconsidering simulations in science education at a distance: features of effective use. *Journal of Computer Assisted Learning, 23*(6), 491-502.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(l), 32-42.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist, 26* (3&4), 369-398.

Chiocchio, F., & Lafrenière, A. (2009). A project management perspective on student's declarative commitments to goals established within asynchronous communication. *Journal of Computer Assisted Learning, 25*(3), 294-305.

Chu, H., Chen, T., Lin, C., Liao, M., & Chen, Y. (2009). Development of an adaptive learning case recommendation approach for problem-based e-learning on mathematics teaching for students with mild disabilities. *Expert Systems with Applications, 36*(3), 5456-5468.

Clinchy, E. (1989). Education in and about the real world. *Equity and Choice, 3*, 19-29.

Downing, K., Kwong, T., Chan, S., Lam, T., & Downing, W. (2009). Problem-based learning and the development of metacognition. *Higher Education, 57*(5), 609-621.

Dyrli, O.E. & Kinnaman, D.E. (1994). Integrating technology into your classroom curriculum. *Technology and Learning, 14*(5), 38-44.

Ehrmann, S.C. (1995). Asking the right questions. *Change, 27*(2), 20-27.

Green, K., & Gilbert, S. (1995). Great expectations: Content, communications, productivity and the role of information technology in higher education. *Change, 27*(2), 8-18.

Hakkarainen, P. (2009). Designing and implementing a PBL course on educational digital video production: lessons learned from a design-based research. *Educational Technology Research & Development, 57*(2), 211-228.

Hennessy, S. (2006). Integrating technology into teaching and learning of school science: a situated perspective on pedagogical issues in research. *Studies in Science Education, 42*, 1-48.

Jackson, D., Bourdeau, G., Sampson, A., & Hagen, T. J. (1997). Internet Resources for Middle School science: Golden Opportunity or “Silicon Snake Oil”? *Journal of Science Education and Technology, 6*(1), 49-57.

Krajcik, J. S., Blumenfeld, P. C., Marx, R. W., & Soloway, E. (1994). A collaborative model for helping teachers learn project-based instruction. *Elementary School journal, 94*, 483-497.

Lebow, D. G., & Wager, W. W. (1994). Authentic activity as a model for appropriate learning activity: Implication for design of computer-based simulations. *Paper presented at the 1994 National Convention of the Association for educational Communications Technology Sponsored by the research and Theory Division*, February 16-20, Nashville, TN, USA.

Linn, M. C., Kessel, C., Lee, K., Levenson, J., Spitulnik, M., & Slotta, J. D. (2000). *Teaching and learning k-8 mathematics and science through inquiry: Program reviews and recommendations.* Retrieved January 10, 2010, from http://www.metiri.com/ Solutions/k8ReportSubmitted.doc.

National Research Council (1999). *Being fluent with information technology*, Washington, DC: National Academy Press.

Omale N., Hung W., Luetkehans L., & Cooke-Plagwitz J. (2009). Learning in 3-D multiuser virtual environments: Exploring the use of unique 3-D attributes for online problem-based learning. *British Journal of Educational Technology*, *40*(3), 480-495.

Penrose, R. (1989). *The emperor’s new mind: Concerning computers, minds, and the laws of physics,* Oxford: Oxford University Press.

Peterson, S. (1985). Using STELLA in environmental Education. *Environmental education report and newsletter*, *14*(2), 13-18.

Polychronopoulou, A., & Divans, K. (2009). Dental Students' Perceived Sources of Stress: A Multi-Country Study. *Journal of Dental Education, 73*(5), 631-639.

Roschelle, J. M., Pea, R. D., Hoadley, C. M., Gordin, D. N., & Means, B. M. (2000). Changing How and What Children Learn in School with Computer-Based Technologies: The Future of Children. *Children and Computer Technology, 10*(2), 76-101.

Santos-Trigo, M., & Camacho-Machín, M. (2009). Towards the Construction of a Framework to Deal with Routine Problems to Foster Mathematical Inquiry. *Problems, Resources & Issues in Mathematics Undergraduate Studies, 19*(3), 260-279.

Şendağ, S., & Ferhan, O. H. (2009). Effects of an online problem based learning course on content knowledge acquisition and critical thinking skills. *Computers & Education, 53*(1), 132-141.

Steed, M. (1992). Stella, A simulation construction kit: Cognitive process and educational implications. *Journal of Computers in Mathematics and Science Teaching, 11*(1), 39-52.

Schutte, J. G. (1997). *Virtual Teaching in Higher Education: the new intellectual superhighway or just another traffic jam*? Retrieved May 1, 2009, from http://www.csun.edu/sociology/virexp.htm.

Teaching Professor (2009). *Why Group Work Improves Problem-Solving Abilities*. Retrieved May 1, 2009, from http://www.magnapubs.com/issues/magnapubs_tp/23_5/.

Yew, E., & Schmidt, H. (2009). Evidence for constructive, self-regulatory, and collaborative processes in problem-based learning. *Advances in Health Sciences Education, 14*(2), 251-273.

Chen, L.-J., Ho, R.-G., & Yen, Y.-C. (2010). Marking Strategies in Metacognition-Evaluated Computer-Based Testing. *Educational Technology & Society*, *13* (1), 246–259.

Marking Strategies in Metacognition-Evaluated Computer-Based Testing

Li-Ju Chen, Rong-Guey Ho* and Yung-Chin Yen
Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taipei Taiwan // ljchen@ice.ntnu.edu.tw // hrg@ntnu.edu.tw // scorpio@ice.ntnu.edu.tw
*Contact author

ABSTRACT

This study aimed to explore the effects of marking and metacognition-evaluated feedback (MEF) in computer-based testing (CBT) on student performance and review behavior. Marking is a strategy, in which students place a question mark next to a test item to indicate an uncertain answer. The MEF provided students with feedback on test results classified as correct answers with and without marking or incorrect answers with and without marking. The study analyzed 454 ninth graders randomly assigned to three groups: G_{mm} (marking + MEF), G_{mu} (marking), and G_{uu} (none). Each group was further categorized into three subgroups based on their English ability. Results showed that marking improved medium-ability examinees' test scores. This was a promising finding because the medium-ability students were the very target group that had the most potential for improvement. Additionally, MEF was found to be beneficial as well in that it encouraged students to use marking skills more frequently and to review answer-explanations of the test items. The follow-up interviews indicated that providing adaptive and detailed AEs for low-ability students were necessary. The present study reveals the potential of integrating marking and adaptive feedbacks into the design of learning functions that are worth implementing in CBT systems.

Keywords

Computer-based testing (CBT), Test-taking behavior, Marking behavior, Metacognition evaluation, Confidence rating technique

Introduction

Computer-based testing (CBT) has been widely used since information technology became popularity. Such tests are easily administrated by computer or an equivalent electronic device, and students can immediately access their test results. Many researchers claimed that CBT systems were valuable self-evaluation tools for self-managed learning (Croft, Danson, Dawson, & Ward, 2001; Peat & Franklin, 2002). However, studies indicated that, for effective and efficient use as self-managed learning tools, CBT systems must provide adaptive feedback for future learning (Souvignier & Mokhlesgerami, 2006; Thelwall, 2000; Wong, Wong, & Yeung, 2001). They must also provide information that enables students to control their own pace during the test (Parshall, Kalhn, & Davey, 2002, p. 41).

Adaptive feedback enabled students to learn according to provided instructional strategies (Collis & Messing, 2000; Collis & Nijhuis, 2000). According to Collis, De Boer and Slotman (2001), giving adaptive feedback after a test was one strategy for helping students learn effectively. It could help underachievers extend their learning. For example, giving answer-explanations (AEs) related to key knowledge concepts of test items after a CBT could help students to understand what they have learned and to identify their mistakes (Wang, Wang, Huang, & Chen, 2004); that is, AEs was a metacognitive strategy (Rasekh & Ranjbary, 2003). Answer-explanations offered via automatic evaluation tools could correct student mistakes, reinforce their memories, and support their learning as well as reduce teacher workload so that individual students could receive adaptive compensatory instruction in a forty-student class. Therefore, if CBT systems only displayed scores without feedback, the "teachable moment", or the moment of educational opportunity when students were disposed to learn, might not be used effectively (Collis *et al*., 2001; Ram, Cox, & Narayanan, 1995).

To help students control their own pace, CBT systems could provide the information needed to navigate a test, such as reminders of unanswered items. Gibson, Brewer, Dholakia, Vouk and Bitzer (1995) showed that such information could help students complete the CBT efficiently and reduce their frustration and anxiety. Another mechanism for controlling the testing process within the CBT environment was the marking function. Marking was a skill used to increase the efficiency and effectiveness of self-managed learning (Parshall *et al*., 2002, p34). In the present study, marking referred to a test-taking behavior, in which the student placed a question mark next to a test item to indicate an uncertain answer, and it also served as a reminder to review, check or revise the answer. According to Higgins and Hoffmann (2005), students rarely marked test items when they were sure of their answers. Therefore, marking could be considered one alternative to the confidence rating technique conventionally used to measure the metacognition monitoring ability of students. Students applying confidence rating technique were required to check the confidence degree of their answers. Their metacognition monitoring ability was then evaluated by the matching the confidence degree with the test results (Baker & Brown, 1984; Vu, Hanley, Desoete, & Roeyers, 2006; Strybel, & Proctor, 2000). For example, choosing a correct answer and marking it high on confidence level suggested good metacognition monitoring ability whereas choosing a wrong answer and marking it high on confidence level indicated poor metacognition monitoring ability.

This study proposed metacognition-evaluated feedback (MEF), a new feedback mode for CBT systems displaying AEs integrating student answer responses and marking records. This study had two purposes. First, it explored whether marking could improve the test scores of examinees. Second, it investigated how MEF affected the review behavior of students after completing a CBT. To achieve these two purposes, an experiment was designed to address the following questions:

1. Does marking improve student scores?
2. Does MEF increase use of marking skills and review behavior?

Related research

Test-taking behavior and marking

Test-taking behavior varies among students. Researchers generally classified test-taking behaviors into nine types: (1) browsing items, (2) clarifying meanings of item body and options, (3) knowing the answer, (4) not knowing the answer and guessing, (5) omitting, (6) abandoning, (7) not reaching, (8) having partial knowledge that might be right or wrong, and (9) changing answers (Brown, 1980; Lazarte, 1999; Lord, 1975; McMorris & Leonard, 1976). Examinees usually used marking skills under type (8) and (9) conditions (Burton, 2002; Parshall *et al*., 2002, p.34) because marking was a helpful test-taking technique for checking answers. However, most CBT systems described in the literature did not incorporate the marking function (Gibson *et al*., 1995; Parshall *et al*., 2002, p.34).

Marking was a direct test-taking strategy used by students. It helped examinees remember the test items they skipped or wanted to recheck. The marked test items could then be changed according to partial knowledge or the test-taking skills of the examinees (Burton, 2002). Rogers and Bateson (1991) concluded that good test-taking skills and knowledge of a certain subject could help examinees improve their scores by identifying clues embedded in the test items. Therefore, marking was likely to enhance student performance because it could make them focus on specific items. However, current CBT systems such as Mklesson, Tutorial Gateway, Eval and Open Learning Agency of Australia (Gibson *et al*., 1995), LON-CAPA (http://www.lon-capa.org), and TopClass (http://www.websystems.com) (Bonham, Beichnen, Titus, & Martin, 2000; Wang *et al*., 2004) did not analyze marking behavior. Briefly, a noticeable problem of the current CBT systems was that they did not incorporate the marking function. In CBT systems without the marking function, examinees might not focus on the items they needed to reconsider. Therefore, this study attempted to overcome this problem by designing a CBT system with marking function.

Confidence rating technique

Marking indicated student confidence as well as a remainder to recheck test items (Parshall *et al*., 2002, p.34; Higgins *et al*., 2005). For example, students might put a check mark beside a test item to indicate that they were not sure of the answer. Restated, marking was an alternative approach for judging the confidence level of examinees, which was traditionally measured by using confidence rating technique to estimate metacognition monitoring ability (Baker & Brown, 1984). Other measurement methods, such as interview, observation, thinking aloud, self-reporting and questionnaire survey, have also been used in past studies (Desoete & Roeyers, 2006; Elshout-Mohr, Meijer, van Daalen-Kapteijns, & Meeus, 2003; Garner, 1988, p61). However, each had drawbacks. The analytic results of interview, observation and thinking-aloud were accurate but time-consuming. Moreover, coherent results were difficult to obtain because these measurement methods often involved subjective evaluations (Veenman, 2003). Also, the results of self-reporting and questionnaire survey might induce ‘response set’ problems such as careless answering or acquiescence and social expectations (Garner, 1988, p.61; Linn & Gronlund, 2000, p.182). Therefore, this study employed marking as a confidence rating technique for the benefits of its stability, efficiency and practicality.

Confidence rating technique was performed as follows. Examinees estimated their confidence in their answers by ticking one of the three levels: ‘sure correct’, ‘not sure’, or ‘sure incorrect’. Their metacognition monitoring ability was then measured by matching their confidence degree (‘sure correct’ or ‘sure incorrect’) with their test results (‘correct’ or ‘incorrect’). For example, students who chose a correct answer but marked ‘sure incorrect’ on the confidence level suggested that they had poor metacognition monitoring ability. Conversely, students choosing a wrong answer and marking ‘sure incorrect’ on the confidence level showed that they had good metacognition monitoring ability. However, noted that students who marked ‘not sure’ were excluded from the analysis of metacognition monitoring ability regardless of whether their answers were correct. This approach provided simple and quick measures, which were expected in computer-based adaptive learning environments (Kalyuga, 2006). However, the problem with this confidence rating technique was that low-ability examinees were most likely to choose ‘sure incorrect’ in tests, and most indeed ended up having incorrect answers. Therefore, they were mistakenly interpreted as students with high metacognition monitoring abilities. The method applied in this study avoided this problem since MEF can clearly identify this particular group.

In short, using marking as an alternative confidence rating technique was not only a good way to measure the metacognition monitoring abilities of examinees; it was also rather easy to incorporate into CBT systems (Parshall *et al*., 2002, p.34). Therefore, if the CBT was designed to employ marking, the confidence rating technique could be applied, and data for metacognition monitoring abilities could be attained.

Design of metacognition-evaluation feedback

This study proposed metacognition-evaluated feedback (MEF), a new feedback mode integrating student marking records and responses. Before the CBT starts, students were instructed to place a mark on the test item where they were unsure of the answer. As soon as students completed the CBT, they obtained the MEF. The marking and correctness of their answers were the criteria used to classify their test results into four categories: Category I, II, III, and IV (Chen, Ho, & Yen, 2006). Category I represented correct answers with marking while Category II denoted correct answers without marking. Category III included incorrect answers with marking whereas Category IV was incorrect answers without marking. Since the presence of marking indicated whether or not students were sure of their answers, Categories I and III could therefore be defined as unsure-correct and unsure-incorrect. Further, Category II and IV were defined as expected-correct and unexpected-incorrect according to failure-driven learning theory (Pacifici & Garrison, 2004; Schank, 1995). This learning theory claimed that mistakes, including unsuccessful results and unmet expectations, were failures that could promote advanced learning. For instance, students made predictions about their test results and then observed what happened to check their predictions. If their predictions failed, they tried to determine how these mistakes occurred and then solved their problems. In MEF, further classification of incorrect responses as either unsure or unexpected might motivate students to practice further.

In this study, MEF adopted marking as an indicator of student confidence level. Compared with traditional confidence rating technique, marking was more straightforward, and it reduced interference because it did not require students to check confidence level on each test item during the test (Jacobs & Chase, 1992; Wise & Plake, 1989). Also, by excluding Category III (incorrect answers with marking) from the score for metacognition monitoring ability, MEF avoided a common problem in traditional confidence rating technique: misinterpreting low-ability students as having high metacognition monitoring abilities.

As Wang *et al*. (2004) indicated, CBT systems that collected and analyzed student responses and answering processes could identify student learning outcomes and subject matter misconceptions. Therefore, the AEs in the MEF were designed to incorporate the above information to provide useful adaptive feedback so that students could understand their performance, clarify their mistakes, and increase their learning motivation.

Methodology

Participants

A total of 454 ninth-graders participated in this experiment. All participants had over five years of formal computer literacy instruction (more than 180 hours), which confirmed that they had basic computer skills required to take a

CBT. One reason they volunteered to take part in the experiment was because they wanted to prepare for the English Basic Competence Test (EBCT) given by the Committee of the Basic Competence Test for Junior High School Students three months later. The participants were randomly assigned to G_{uu} (four classes, 145 students), G_{mu} (four classes, 139 students), and G_{mm} (five classes, 170 students).

Three versions of CBT system

Three versions of CBT system, labeled G_{mm}, G_{mu}, and G_{uu}, were designed based on two factors, marking and MEF. The G_{mm} (See Figure 1) adopted marking; MEF; G_{mu} (See Figure 2) only adopted marking, and G_{uu} adopted neither of them (See Figure 3). Figure 1 shows an example of MEF test results in G_{mm}: examinee X had twenty correct responses and ten incorrect responses, *i.e.*, 67% of responses by X were correct. The test results were then categorized as follows: the 2nd, 5th, 12th, 15th, 21st, and 25th test items were marked and correct (Category I); the 3rd, 6th, 8th, 9th, 11th, 14th, 17th, 18th, 22nd, 24th, 26th, 28th, 29th, and 30th test items were unmarked and correct (Category II); the 7th and 27th test items were marked and incorrect (Category III), and the 1st, 4th, 10th, 13th, 16th, 19th, 20th, and 23rd were unmarked and incorrect (Category IV). However, Figure 2 shows an example of feedback for the test results of examinee X in G_{mu}. The displayed information was identical to that in G_{mm} but not sorted into four categories. Figure 3 shows an example of feedback for the test results of examinee X in G_{uu}. The displayed information was similar to that in G_{mu}, except that the summary did not include marking records. Three versions of CBT system recorded examinee scores, answer responses, time consumed and review records in each action. The examinee test results and responses were examined for effects of marking on student scores, test-taking time, and MEF on marking skills and review behavior.

Briefly, the three versions of CBT system were as follows:

1. G_{mm}: Examinees could place or remove a question mark on any items, indicating they were ‘unsure’. The examinee responses, results, marks, and scores for each item were shown on the screen and sorted into four categories after the test was completed.
2. G_{mu}: The marking method was the same as that in G_{mm}, and so was the displayed information. However, the information was not sorted into four categories in this version.
3. G_{uu}: Examinees could not mark any items. Except for marking, the displayed information was the same as that in G_{mu}.

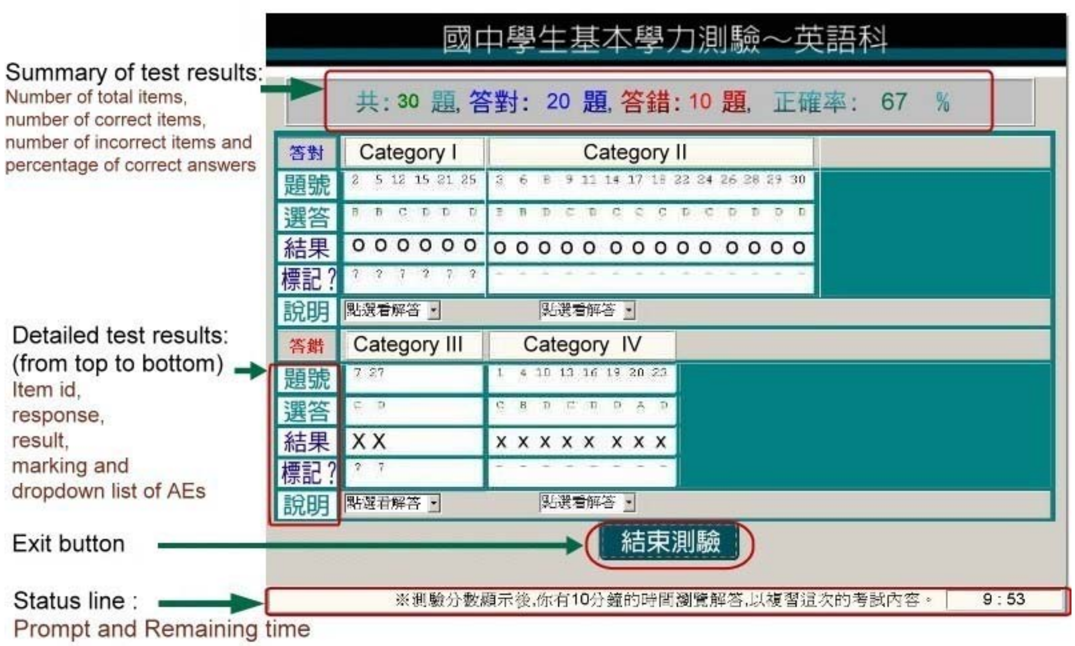

Figure 1. Example of MEF screen in G_{mm}

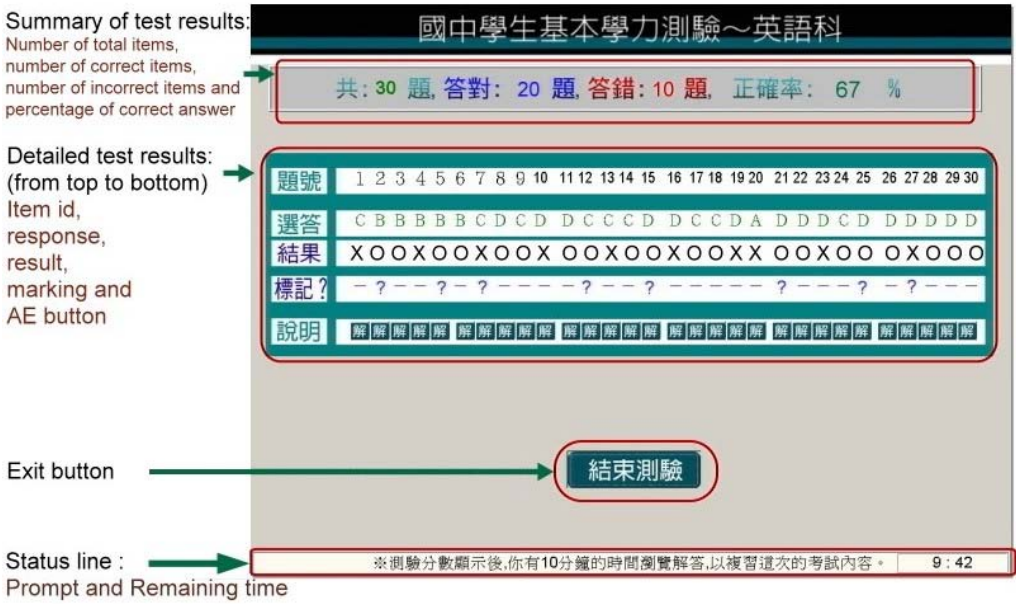

Figure 2. Example of feedback screen in G_{mu}

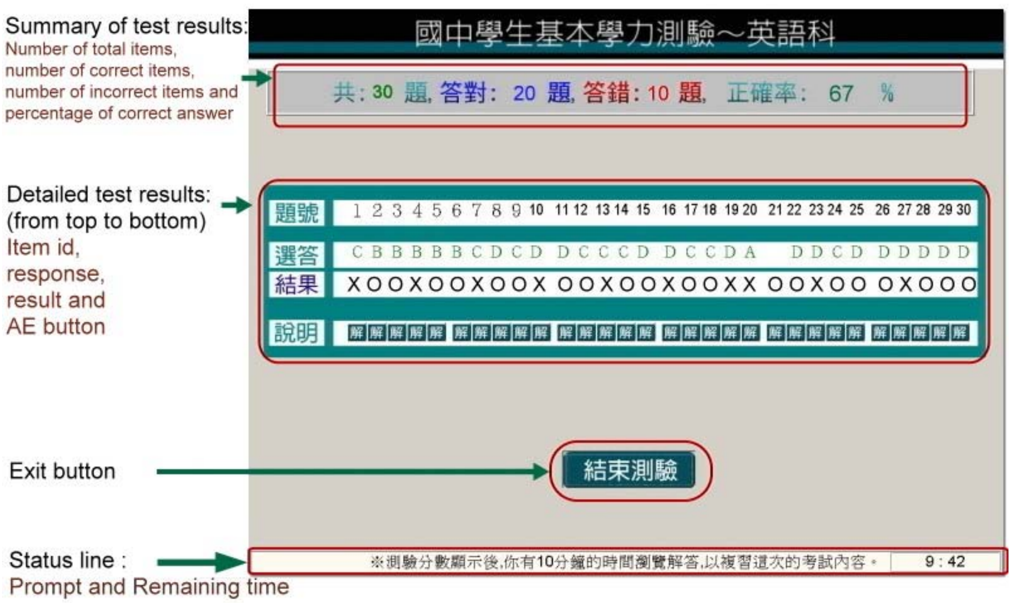

Figure 3. Example of feedback screen in G_{uu}

Test items

Three versions of CBT system were adopted in a test comprised of thirty multiple-choice items selected from the vocabulary and reading comprehension sections of the EBCT in Taiwan. Because less than 500 participants were sampled, analyzing the parameters of test items based on three-parameter model in item response theory was unsuitable (Hambleton & Swaminathan, 1985, p.227, p.308; Mason, Patry, & Bernstein, 2001). Therefore, classical test theory was used; the item difficulty index and item discrimination index of the test items were calculated. As Table 1 shows, both indices had means above .5, and item discrimination indices were above .4. The reliability of internal consistency (KR-20) was .926. These three figures indicated that the quality of the test items was acceptable (Ahmanan & Glock, 1981, p163; Ebel & Frisbie, 1991, p.231-32; Noll, Scannell, & Craig, 1979, p.109). The following is an example item in the reading comprehension section:

In 1999, there were about 2,482 traffic accidents in Taiwan. Most of the accidents happened because _________1_________. For example, some drivers drove too fast. Some drivers drank too much wine or beer before they got into the car. And some drivers never tried to stop when the traffic lights went from yellow to red.

Most of the accidents happened because _________1_________.
(1) motorcyclists were riding too fast
(2) the MRT system was not built yet
(3) drivers didn't follow the traffic rules
(4) there were too many traffic lights on the road

Table 1. Statistical properties of the test items (number of items=30, number of examinees =454)

Parameter	Mean	Std Dev	Minimum	Maximum
Item difficulty index	.56	.098	.37	.71
Item discrimination index	.70	.095	.44	.89

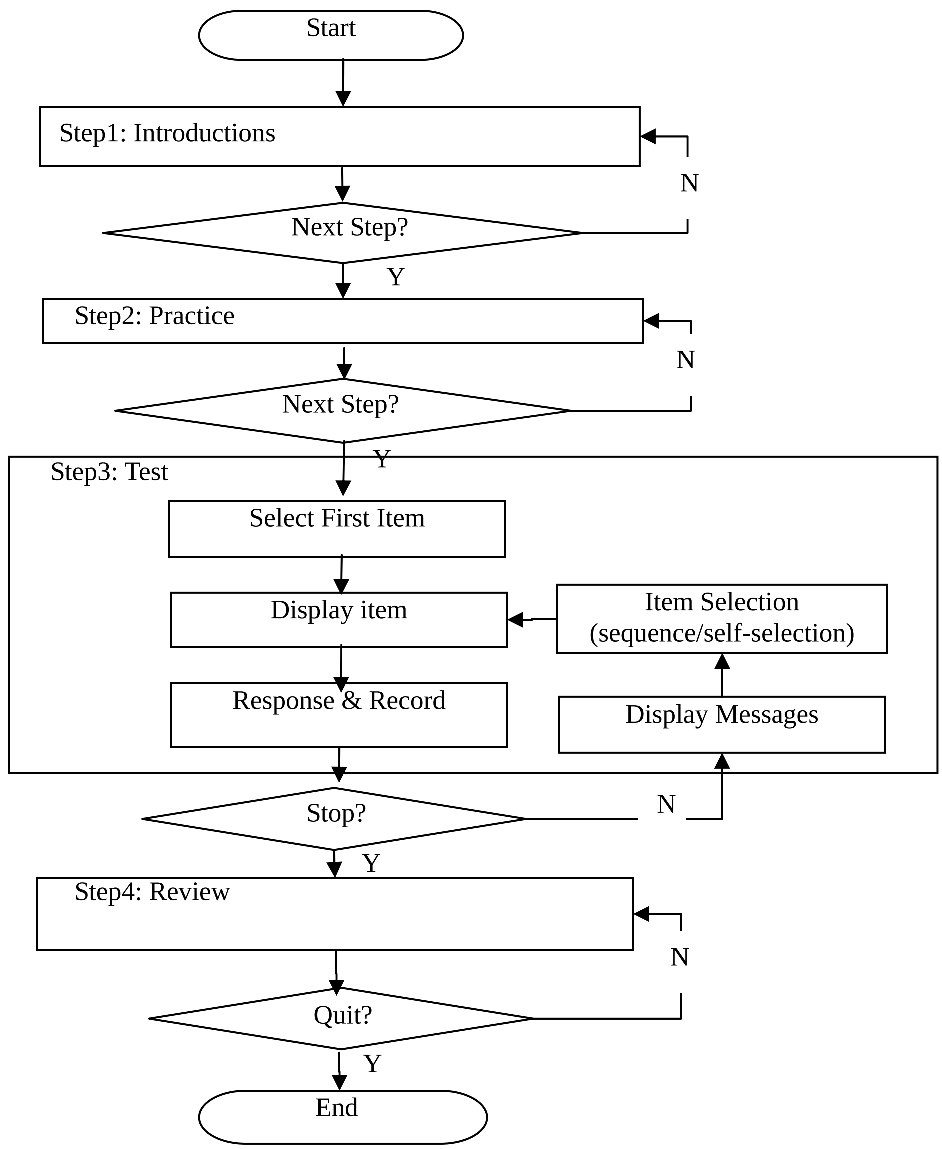

Figure 4. Testing procedure of CBT system

Procedure

Design three versions of CBT system

As Figure 4 shows, three versions of CBT system first displayed instructions and demo clips page by page in the Instruction session. During the Practice session, examinees could answer the four sample items repeatedly to familiarize themselves with the CBT system interface. In the Test session, examinees could change their answer(s) by selecting items from a pop-up menu. The participants in G_{mm} and G_{mu} were also instructed to put a question mark next to items in which they were unsure of the answer. They could also remove question marks if they later felt certain that their answers were correct. The stopping conditions of the test were activated when the test time was finished or when all the items were completed and the ‘Finish’ button was clicked. The scores were calculated as soon as the test ended.

In the Review session, the test results were shown and examinees could review the AEs including supplementary materials related to major knowledge concepts written by several junior-high-school English teachers. Also, according to suggestions from these teachers, the important concepts, words, and keys in the CBT system were highlighted with bright colors. Figure 5 illustrates an example of the AEs for test items in the Review session.

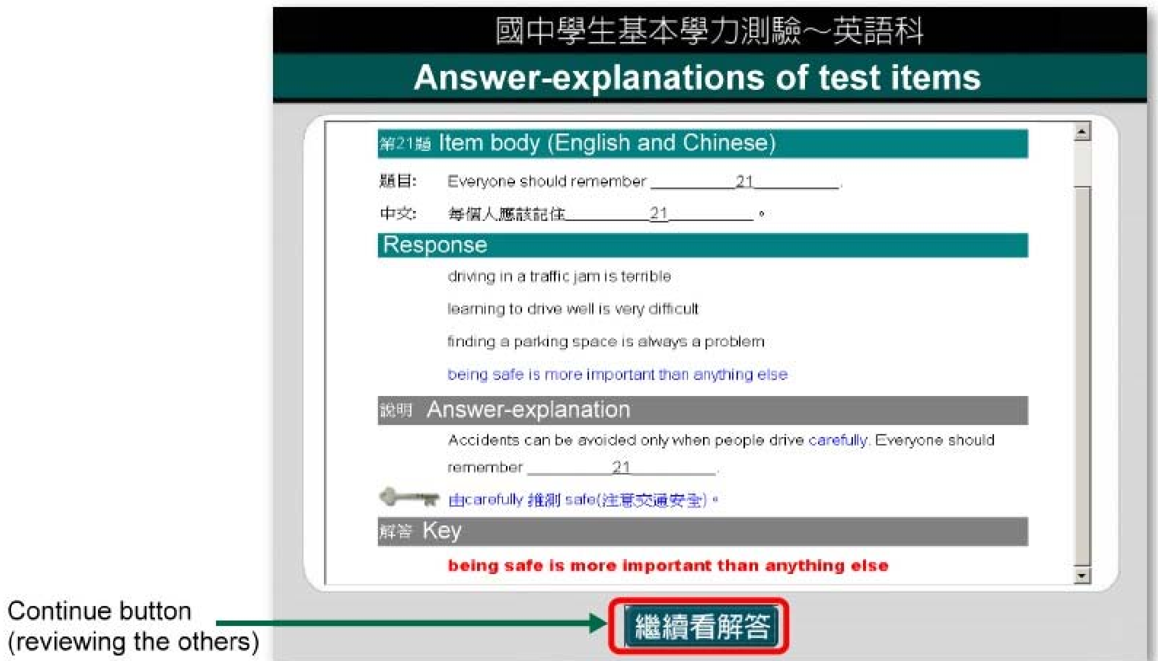

Figure 5. Screenshot of an AE for a test item

Tryout

Six ninth-graders and three eighth-graders from another school in the same district were recruited for system testing and tryout. The student with the lowest EA in the tryout group was able to complete a thirty-item test in 30 minutes. Therefore, the test-taking time in the main study was set to 30 minutes to ensure that all participants could finish the test within the time limit. Problems such as unclear instructions, blurred pictures, and misspelled items were corrected after the tryout.

Main study

All participants were volunteers and were randomly assigned to the G_{mm}, G_{mu}, or G_{uu} group. They were coached briefly on the testing procedure, answering method, and how to get AE feedback on each CBT systems before taking the EBCT. Students in G_{mm} and G_{mu} were instructed to place a question mark next to any item when they were not sure of the answer. The participants in the G_{mm} were told that the Review session would display their test results and sorted AEs according to their marking records and answer responses while those in G_{mu} and G_{uu} were told that their test results and AEs given in the Review session would not be sorted. The participants took the ECBT in the computer classroom in their school to control for the anxiety associated with testing in an unfamiliar environment. The three versions of CBT system recorded the responses and the time administered. After the experiment, the participants received their test results and reviewed their record reports.

Results and Discussion

To investigate the effects of marking and MEF on examinees with different levels of English ability, the examinees in G_{mm}, G_{mu}, and G_{uu} were further classified by their test scores. The top 25% scorers were labeled as high English ability (H-EA) group; the bottom 25% scorers were labeled as low English ability (L-EA) group, and the scorers from 37.5% to 62.5% were labeled as middle English ability (M-EA) group. The sampling procedure skipped the scorers from 25% to 37.5% and from 62.5% to 75% (totaling 25% of the whole sample, *i.e.*, 112 participants). This was to reduce possible influence between two successive EA groups on test scores and review behavior. Restated, the number of sampling participants was 342: 114 examinees in H-EA, M-EA, and L-EA, respectively. For H-EA examinees, thirty-two, forty-four, and thirty-eight were in G_{mm}, G_{mu}, and G_{uu}, respectively. Forty-four, thirty-two, and thirty eight students in the M-EA were classified as G_{mm}, G_{mu}, and G_{uu}, respectively. For L-EA examinees, forty-five, thirty-nine, and thirty were in G_{mm}, G_{mu}, and G_{uu}, respectively. Compared with the total number of participants, the number of sample in each subgroup was relatively small, and the statistical reliability of the analysis would decrease. However, the sample size in each subgroup was still more than 30 which satisfied the criteria of minimum numbers (15 subjects in each subgroup) in an experimental study recommended by Gall, Borg, and Gall (1996, p.229). Therefore, the level of reliability to explore the effects of marking and MEF on test scores and review behavior for examinees with different EA levels was acceptable. Table 2 shows the descriptive statistics for test scores in each subgroup. The average test scores of H-EA examinees in three groups were, from highest to lowest, G_{mm}, G_{uu}, and G_{mu}. However, the average test scores of M-EA and L-EA examinees in three groups were, from highest to lowest, G_{mu}, G_{mm}, and G_{uu}. Additionally, the average test scores of examinees in three groups were, from highest to lowest, G_{mu}, G_{uu}, and G_{mm}.

Table 2. Descriptive statistics of test scores in each subgroup (N=342)

Groups	H-EA			M-EA			L-EA			Total		
	n	Mean	Std Dev	n	Mean	Std Dev	n	Mean	Std Dev	n	Mean	Std Dev
G_{mm}	32	27.97	1.58	44	16.52	3.35	45	7.11	1.42	121	16.05	8.56
G_{mu}	44	27.16	4.14	32	17.25	5.13	39	7.62	3.99	115	17.77	9.40
G_{uu}	38	27.53	1.59	38	14.95	3.42	30	6.27	1.39	106	17.00	8.95
Total	114	27.97	1.61	114	15.79	3.31	114	6.81	1.62	342	16.92	8.97

The following section presents the results and related discussions to answer the two research issues: the effect of marking on three levels of EA examinee performance in G_{mm}, G_{mu}, and G_{uu} and the effects of MEF on their marking frequency and review behavior.

Effects of marking on examinee test scores

To investigate the effects of marking on examinee test scores, the *t*-tests were conducted for three successive EA levels. As Table 3 shows, marking significantly affected the test scores of M-EA examinees ($t_{.05\,(112)} = 2.4$, $p < .05$), but had no significant effect on those of H-EA ($t_{.05\,(112)} = -.05$, $p > .05$) or L-EA ($t_{.05\,(112)} = 1.95$, $p > .05$) examinees. The test results indicated that middle EA examinees benefited significantly from the marking function, which suggested that CBT systems incorporating marking function could improve average student performance. This finding was rather promising because classroom intervention was typically aimed at average level students, since this target group had the most potential for improvement compared to their high ability and low ability counterparts. In the present case, the high ability students were confident of their own answers or they had already understood the important concepts prior to the test; therefore, marking skill was not an immediate need for them. Similarly, marking did not improve the performance of low ability students. They lacked enough knowledge to answer correctly even though they had good marking skills. However, marking might have encouraged average ability students to seek the clues among test items and assist them in answering correctly, which would thus have improved their performance. Therefore, marking should not be neglected in CBT systems design.

Table 3. Descriptive statistics of examinee test scores and results of three *t*-tests (N=342)

Examinees' English ability	Marking						*df*	*t* value
	With			Without				
	n	Mean	Std Dev	n	Mean	Std Dev		
H-EA	76	27.50	3.32	38	27.53	1.59	112	-0.05
M-EA	76	16.83	4.18	38	14.95	3.42	112	2.4*
L-EA	84	7.35	2.90	30	6.27	1.39	112	1.95

*p<.05

Effects of MEF on examinee marking skill and review behavior

To evaluate how MEF affected the marking skills and review behavior of examinees, the ANOVAs used MEF (with and without) and examinee EA levels (H-EA, M-EA and L-EA) as independent variables. Dependent variables were the marking frequency, number of reviewed AEs, and time (in seconds) spent reviewing AEs. The sampling participants in G_{mm} used CBT with MEF function, in which test results were sorted into four categories (See Tab. 2). Thirty-two H-EA, forty-four M-EA, and forty-five L-EA examinees were included. However, the sampling participants in G_{mu} received the same information, but it was not sorted into the same four categories as that in G_{mm} (See Tab. 2). Forty-four H-EA, thirty-two M-EA, and thirty-night L-EA examinees were included in this group.

Marking frequency

Tables 4 and 5 present the descriptive statistics for the marking frequency of examinees and the ANOVA summary, respectively. The statistical results showed that the interaction of MEF and examinee EA was significantly associated with marking frequency ($F_{(2, 230)}$ =4.06, p<.05). *Post hoc* analysis further showed that the marking frequency in M-EA ($t_{.05\ (73)}$ =2.13, p<.05) and L-EA ($t_{.05\ (74)}$ =3.55, p<.001) examinees was significantly higher than that in H-EA examinees when MEF was adopted ($F_{(2,\ 230)}$ =6.31, p<.05) (See Tab. 6). However, marking frequency did not significantly differ among the three levels of examinee EA without MEF ($F_{(2, 230)}$ =.55, p>.05). More importantly, as Figure 6 shows, MEF significantly increased the marking frequency in M-EA ($F_{(1,\ 230)}$ =4.10, p<.05) and L-EA examinees ($F_{(1,\ 230)}$ =20.14, p<.05). That is, MEF motivated low and middle examinees to use marking skills when they knew in advance that the test results would be sorted into four categories. However, MEF did not significantly affect H-EA examinees, possibly due to their mastery of the subject matter, as stated above.

Table 4. Descriptive statistics of marking frequency (N=236)

Examinees' EA	MEF					
	With			Without		
	N	Mean	Std Dev	n	Mean	Std Dev
H-EA	32	4.00	4.72	44	3.57	5.17
M-EA	44	6.95	6.84	32	3.72	4.11
L-EA	45	9.24	12.04	39	2.33	5.27

Table 5. Summary of ANOVA for marking frequency (N=236)

Source	*SS*	*df*	*MS*	*F*
MEF	718.21	1	718.21	13.99*
Examinees' EA	169.97	2	84.98	1.66
MEF × Examinees' EA	417.23	2	208.62	4.06*
w. cell (error)	11810.15	230	51.35	

* p<.05

Table 6. Summary of ANOVA of simple main effects for marking frequency (N=236)

Source	*SS*	*df*	*MS*	*F*	*Post hoc*
Examinees' EA					
With MEF	648.46	2	324.22	6.31*	H-EA,L-EA>H-EA
Without MEF	56.54	2	28.27	0.55	
MEF					
H-EA	12.57	1	12.57	0.24	
M-EA	210.37	1	210.37	4.10*	MEF> without MEF
L-EA	1033.96	1	1033.96	20.14*	MEF> without MEF
w. cell (error)	11810.15	230	51.35		

* $p<.05$

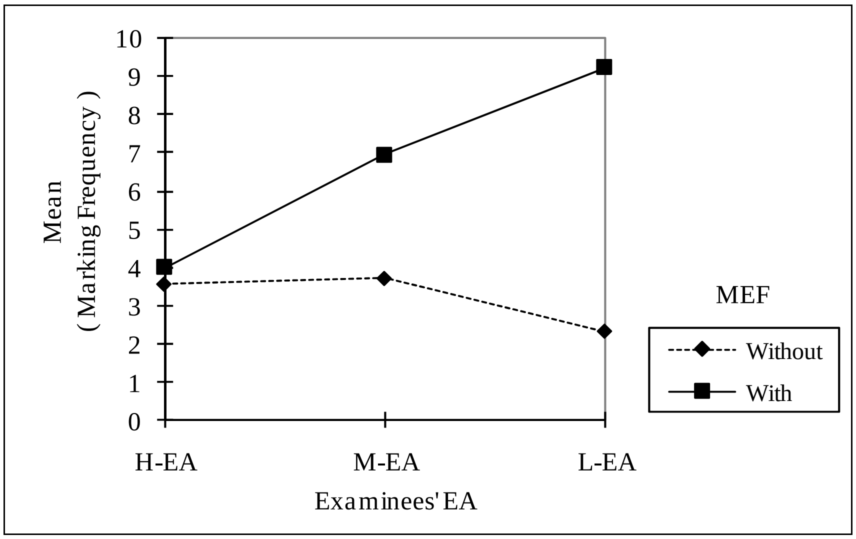

Figure 6. Marking frequency (MEF x Examinees' EA)

Number of reviewed AEs and review time

Table 7 presents the descriptive statistics of the number of reviewed AEs and review time. As Tables 8 and 9 show, no significant interactions were detected between MEF and examinee EA for number of reviewed AEs ($F_{(2,\ 230)}$ =.88, p>.05) and review time ($F_{(2,\ 230)}$ =.57, p>.05). In addition, no significant main effect of examinee EA was found in these two dependent variables (number of reviewed AEs: $F_{(2,\ 230)}$ =1.63, p>.05; review time: $F_{(2,\ 230)}$ =1.75, p>.05). This suggested that low ability students might have reviewed fewer AEs while making far more mistakes might because the design of the AEs did not meet their needs. Further, the significance of the main effects of MEF on both dependent variables were examined (number of reviewed AEs: $F_{(1,\ 230)}$ =4.87, p<.05; review time: $F_{(1,\ 230)}$ =8.33, p<.05) (See Figs. 7 and 8). In short, the CBT incorporating MEF increased the frequency and length of time spent reviewing AEs.

Table 7. Descriptive statistics of number of reviewed AEs and review time (N=236)

	MEF	n	Number of reviewed AEs		Review time	
			Mean	Std Dev	Mean	Std Dev
H-EA						
	With	32	2.44	2.91	63.12	89.62
	Without	44	2.07	2.11	35.34	34.69
M-EA						
	With	44	4.32	5.88	86.82	111.00
	Without	32	2.84	3.88	51.38	60.39
L-EA						
	With	45	3.89	7.51	61.09	83.82
	Without	39	1.51	2.68	33.23	75.69

Table 8. Summary of ANOVA for number of reviewed AEs (N=236)

Source	*SS*	*Df*	*MS*	*F*
MEF	109.26	1	109.26	4.87^{*}
Examinees' EA	73.13	2	36.56	1.63
MEF x Examinees' EA	39.58	2	19.79	0.88
w. cell (error)	5162.28	230	22.44	

* $p<.05$

Table 9. Summary of ANOVA for review time (N=236)

Source	*SS*	*Df*	*MS*	*F*
MEF	53245.19	1	53245.19	8.33^{**}
Examinees' EA	22321.18	2	11160.56	1.75
MEF x Examinees' EA	731.82	2	365.91	0.57
w. cell (error)	1470344.00	230	6392.40	

$^{**}p<.01$

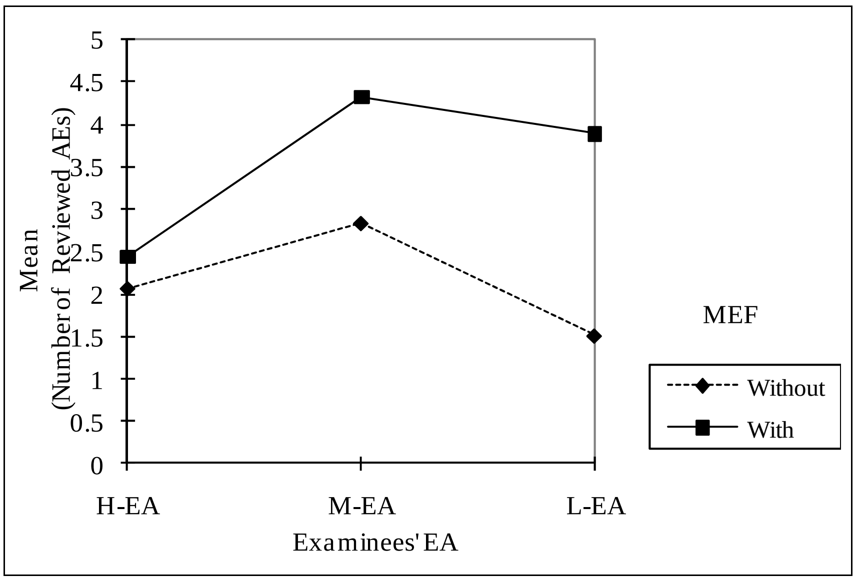

Figure 7. Number of reviewed AEs (MEF x Examinees' EA)

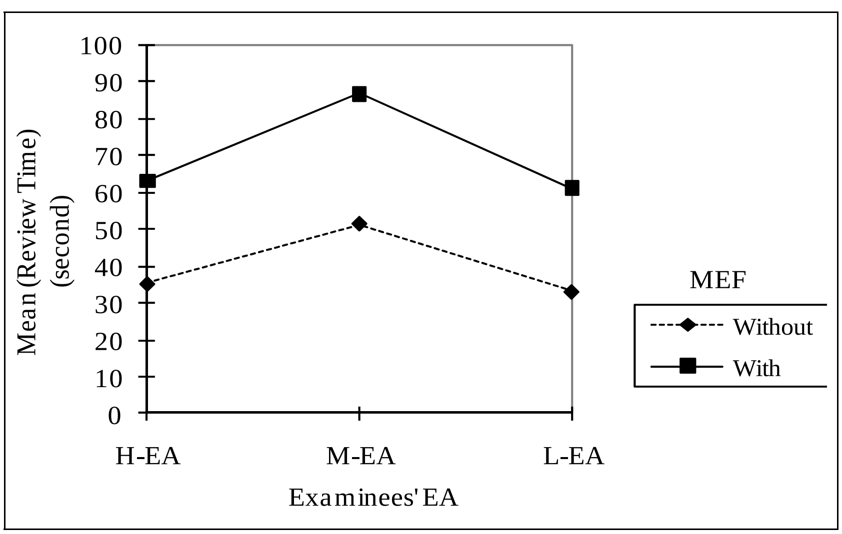

Figure 8. Review time (MEF x Examinees' EA)

In summary, the statistical results showed that marking and MEF had different effects on the scores and review behavior for different levels of EA examinees (See Tab.10). First, for H-EA examinees, marking did not significantly increase test scores. Additionally, MEF did not encourage them to apply marking skills more frequently, but it improved their review behavior, including the number of reviewed AEs and review time. This suggested that the high ability students might have understood important concepts of the subject matter prior to the experiment; thus, they did not need marking skills to help them to recheck their answers. However, the new MEF feedback mode effectively prompted students to review supplementary knowledge both quantitatively and qualitatively. Second, for M-EA examinees, marking significantly increased test scores, possibly because the students tended to spend more time deliberating on the marked items during the test. For this group of students, marking therefore was a reminder to further ponder unsure test items. Besides, MEF motivated them to use marking skills more frequently and promoted their review behavior. This suggested that both marking and MEF were effective functions for the M-EA students. Third, for L-EA students, marking did not significantly increase their test scores. This suggested that low ability students had insufficient knowledge to correct their answers despite the help of marking. Further, the effects of MEF on their marking frequency and review behavior were the same as those of M-EA students. In short, student ability was a critical threshold for marking behavior to improve test performance.

Table 10. Summary results for three levels of examinees' EA

Factor (s)	Independent variable	Results
Marking	Test scores	+ (only M-EA examinees)
MEF and Examinees' EA	Marking frequency	MEF + (under M-EA and L-EA conditions) M-EA, L-EA > H-EA (under MEF condition)
	Number of reviewed AEs	MEF (Main effect) +
	Review time	MEF (Main effect) +

+ : significant increase

Interview with examinees

The qualitative results of interviews in this study also revealed important implications. Ten examinees who ranked in the bottom 25% of scores with low AE review rates were selected and interviewed to explore why they reviewed fewer AEs and made far more mistakes. The students stated that the AEs were too brief to understand and what they needed were more detailed explanations. However, interviews with another ten examinees with average scores indicated that the AEs for test items were sufficiently clear. This suggested that the AEs should be revised for underachievers to encourage review behavior. For example, the content of AEs could be made more detailed, or the illustrations for supplementary materials could be increased. Besides, interviews with six examinees who correctly answered all test items with 100% review rate in G_{uu} indicated that they had to review all AEs in order to find the explanations for the questions they had guessed. This suggested that marking function in CBT systems could be helpful to increase the efficiency of compensatory instruction.

Conclusion

In this study, a metacognition-evaluated CBT system integrating marking and MEF was employed to evaluate their effects on self-managed learning for junior high school students. The MEF was designed using confidence rating technique, failure-driven learning theory, and suggestions from the participating English teachers, which served to support teachers' role of giving individual students adaptive compensatory instruction. In metacognition-evaluated CBT system, marking might also help students navigate lengthy tests. The present study, therefore, investigated the following two issues: (1) whether marking affects students' test scores and (2) whether MEF affects student marking skills and encourage student in review behavior.

Findings were summarized as follows. First, marking enhanced test scores in middle level EA students, but not in high and low level EA students. In other words, of the three groups, the middle level EA students benefited the most from using marking skills. Some low EA students, for example, demonstrated good marking skills but did not significantly improve their test scores. Generally, low level EA students spent less time on testing than high or middle level EA students, which revealed that low level EA students tried to complete items that were far above their ability. Thus, they had little patience to complete difficult test items. Secondly, MEF not only effectively encouraged three levels of EA students to apply marking skills more frequently, but it also motivated their review behavior. Also, follow-up interview indicated that providing adaptive and detailed AEs in CBT systems for low level EA students were necessary.

In sum, the findings in the present study indicate that, marking should be incorporated in CBT systems because it may encourage average students to deliberate upon marked test items, which can then increase their scores. Also, MEF is effective for improving students' review behavior. Additionally, providing AEs of test items after completing a CBT helps students learn and review knowledge concepts related to test items and can also lighten the workload of teachers by minimizing the need for compensatory instruction. To design AEs, future studies should engage students to review their mistakes.

Some limitations are presented in this study as well. The conclusions are based only on the analysis of junior high school students taking EBCT. Experimental results may differ from students in other age groups or with different knowledge domains. These issues should be considered in further studies to obtain more comprehensive results. Future research should investigate how MEF influences student performance and learning motivations as well as how AEs affect student performance.

Acknowledgements

This work was supported in part by the Science Education Division of Taiwan National Science Council under the Grant No. NSC 93-2520-S-003-010-. The project leader, Rong-Guey Ho, is also the corresponding author of this article. The authors would like to thank the NSC of Taiwan for financial support and the anonymous reviewers for constructive suggestions.

References

Ahmanan, J. S., & Glock, M. D. (1981). *Evaluating Student Process: Principles of Tests and Measurement* (6th Ed.), Boston: Allyn & Bacon.

Baker, L., & Brown, A. L. (1984). Metacognitive skills in reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of Reading Research* (pp.353-394), New York: Longman.

Bleuer, J. C., & Walz, G. R. (2002). *New perspectives on Counseling Underachievers*, ERIC Document Reproduction Service No. ED 470602.

Bonham, S. W., Beichner, R. J., Titus, A., & Martin, L. (2002). Education research using web-based assessment systems. *Journal of Research on Computing in Education, 33*(1), 28-43.

Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension* (pp. 454-481), Hillsdale, NJ: Lawrence Erlbaum.

Brown, A. L. (1987). Metacognition, executive control, selfregulation and other more mysterious mechanisms. In F. E. Weinert, & R. H. Kluwe (Eds.), *Metacognition, Motivation, and Understanding* (pp. 65-116), London: Lawrence Erlbaum.

Burton, R. F. (2002). Misinformation, partial knowledge and guessing in true/false tests. *Medical Education, 36*, 805-811.

Chen, L. J., Ho, R. G., & Yen, Y. C. (2006, Octber). Effects of metacognition evaluated strategy on computer-based test system. *Paper presented at the E-Learn 2006 Conference*, October 13-17, Honolulu, Hawaii, USA.

Collentine, J. (2000). Insights into the construction of grammatical knowledge provided by user-behaviour tracking technologies, *Language Learning and Technology, 3*(2), 44-57.

Collis, B., De Boer, W., & Slotman, K. (2001). Feedback for web-based assignments, *Journal of Computer Assisted Learning, 17*, 306-313.

Collis, B., & Nijhuis, G. G. (2000). The instructor as manager: Time and task. *The Internet in Higher Education, 3*, 75-97.

Collis, B., & Messing, J. (2000). Usage, attitudes and workload implications for a Web-based learning environment. *Journal of Advanced Learning Technologies, 9*(1), 17–25.

Croft, A. C., Danson, M., Dawson, B. R., & Ward, J. P. (2001). Experiences of Using Computer Assisted Assessment in Engineering Mathematics. *Computers & Education, 37*, 53-66.

Desoete, A., & Roeyers, H. (2006). Metacognitive macroevaluations in mathematical problem solving. *Learning and Instruction, 16*(1), 12-25.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th Ed.), Englewood, NJ: Prentice Hall.

Elshout-Mohr, M., Meijer, J., van Daalen-Kapteijns, M., & Meeus, W. (2003). A self-report inventory for metacognition related to academic tasks. *Paper presented at the 10th Conference of the European Association for Research on Learning and Instruction (EARLI)*, August 26-30, Padova, Italy.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational Research: An introduction* (6th Ed.), White Plains, NY: Longman.

Garner, R. (1988). *Metacognition and Reading Comprehension*, Norwood, NJ: Ablex.

Gibson, E. J., Brewer, P. W., Dholakia, A., Vouk, M. A., & Bitzer, D. L. (1995). A comparative analysis of web-based testing and evaluation systems. *Proceedings of the 4th WWW conference*. Retrieved September 3, 2009, from http://renoir.csc.ncsu.edu/MRA/Reports/WebBasedTesting.html.

Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*, Boston: Kluwer.

Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation on reading test performance. *The Journal of Technology, Learning, and Assessment, 3*(4), 1-35

Jacobs, L. C., & Chase, C. I. (1992). *Developing and Using Tests Effectively: A Guide for Faculty*, San Francisco, CA: Jossey-Bass.

Kalyuga, S. (2006). Rapid cognitive assessment of learners' knowledge structures. *Learning and Instruction, 16*(1), 1-11.

Lazarte, A. A. (1999). Modeling time to respond and probability of correct answer in a simulated computerized test-taking situation. *Paper presented at the Annual Meeting of the American Educational Research Assoc*, April 19-23, Montreal, Canada.

Linn, R. L., & Gronlund, N. E. (2000). *Measurement and Assessment in Teaching* (8th Ed.), Upper Saddle River, NJ: Prentice-Hall.

Lord, F. M. (1975). Formula-scoring and number-right scoring. *Journal of Educational Measurement, 12*(1), 7-12.

Mason B.J., Patry M., & Bernstein D.J. (2001) An examination of the equivalence between non-adaptive computer-based test and traditional testing. *Journal of Educational Computing Research, 24*, 29–39.

McMorris, R. F., & Leonard, G. (1976). Item response changes and cognitive style. *Paper presented at the 22nd Annual Meeting of the National Council on Measurement in Education*, San Francisco (ERIC Document Reproduction Service ED 12991839).

Miller, T. W., & Weiss, D. J. (1976). Effect of Time Limits on Test-taking Behavior, *Research Report 76-2*, Minneapolis: Minnesota University, Dept. of Psychology (ERIC Document Reproduction Service No. ED126123).

Noll, V. H., Scannell, D. P., & Craig, R. C. (1979). *Introduction to Educational Measurement* (4th Ed.), Boston: Houghton Mifflin.

Pacifici, L., & Garrison, J. (2004). Imagination, emotion and inquiry: The teachable moment. *Contemporary Pragmatism, 1*(1), 119-132.

Parshall, C. G., Kalhn, J. C., & Davey, T. (Eds.). (2002). *Practical Considerations in Computer Based Testing*, New York: Springer-Verlag.

Peat, M., & Franklin, S. (2002). Supporting student learning: The use of computer-based formative assessment modules. *British Journal of Educational Technology, 33*(5), 515-23.

Ram, A., Cox, M. T., & Narayanan, S. (1995). Goal-driven learning in multistrategy reasoning and learning system. In A. Ram & D. B. Leake (Eds.), *Goal-driven Learning* (pp421-437), Cambridge, MA: MIT Press/Bradford Books.

Rasekh, Z., & Ranjbary, R. (2003). Metacognitive strategy training for vocabulary learning, *TESL-EJ, 7*(2), 1-18.

Rogers, W. T., & Bateson, D. J. (1991). Verification of a model of test-taking behavior of high school seniors. *Journal of Experimental Education, 59*(4), 331-350.

Schank, R. C. (1995). End run to the goal line. *EDU- COM Review, 30*(1), 14-17.

Souvignier, H., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction, 16*(1), 57-71.

Thelwall, M. (2000). Computer-based assessment: A versatile educational tool. *Computers & Education, 34,* 37-49.

Veenman, M. V. J. (2003). The assessment of metacognitive skills: what can be learned from multi-method designs? *Paper presented at the 10th Conference of the European Association for Research on Learning and Instruction (EARLI)*, August 26-30, Padova, Italy.

Vu, K. L., Hanley, G. L., Strybel, T. Z., & Proctor, R. W. (2000). Metacognitive processes in human-computer interaction: Self-assessments of knowledge as predictors of computer expertise. *International Journal of Human-Computer Interaction, 12*(1). 43-71.

Wang, T. H., Wang, K. H., Wang, W. L., Huang, S. C., & Chen, S. Y. (2004). Web-based assessment and test analyses (WATA) system: Development and evaluation. *Journal of Computer Assisted Learning, 20*(1), 59-71.

Wise, S. L., & Plake, B. S. (1989). Research on the effects of administering test via computers. *Educational Measurement: Issues and Practice, 8*(3), 5-10.

Wong, C. K., Wong, W., & Yeung, C. H. (2001). Student behavior and performance in using a Web-based assessment system. *Innovations in Education and Teaching International, 38*(4), 339-346.

Liu, G.-Z. (2010). Book review: Handbook of Research on Educational Communications and Technology (Eds. J. M. Spector, M. D. Merrill, J. van Merrienboer, & M. P Driscoll). *Educational Technology & Society, 13* (1), 260-263.

Handbook of Research on Educational Communications and Technology

(Book Review)

Reviewer:

Gi-Zen Liu
Faculty of ET & CALL, Foreign Language & Literature Department
National Cheng Kung University, Tainan City, Taiwan
gizen@mail.ncku.edu.tw

Textbook Details:

Handbook of Research on Educational Communications and Technology (3rd Ed.)
Editors: J. Michael Spector, M. David Merrill, Jeroen van Merrienboer, and Marcy P Driscoll
2008, Lawrence Erlbaum, ISBN: 978-0-415-96338-1
Chapters: 56; Pages: 928

As a variety of information and communication technologies (ICT) have been emerging and evolving in different contexts and fields, it is estimated that ICT integrated education will become normal in entirely online learning environments and in blended courses over the next five to ten years (Mayadas *et al*, 2009). In terms of research and development in educational technology (ET), the pioneers and some practitioners have already been experiencing it for several decades ; however, in terms of the paradigm shift in ET research and development, we have travelled only a short distance down the path of a thorough educational and conceptual reconfiguration (Ely, 2008). Hence, research concerning how to choose, use, design, develop, implement, manage, and evaluate appropriate ICT with cutting-edge methodology and theory—in learning, instruction, training, and beyond—has become necessary and crucial in the diverse and broad field of ET (Liu, 2008).

This is the third edition of a *Handbook*, with the first appearing in 1996, and the second in 2004. The latest edition reflects the fact there that there have been a number of new technological developments and innovative educational utilizations of emerging ICT over the last few years. There are 56 chapters in the six major parts of the *Handbook*, with a total of almost one thousand pages. The wide-ranging contributions in this third edition show that it has met the needs of the nearly 200 members of the Association for Educational Communications and Technology (AECT), who provided (a) their feedback on the second edition, and (b) details of what they would like to see in this new edition, which were collected via an online survey. This strategy for updating the work effectively illustrates the thoughtful efforts of the four co-editors, all of whom are well-established ET scholars.

This *Handbook* may be viewed as one of the most thought-provoking works of the current ET research paradigm. It is not only a collaborative and insightful academic handbook comprised of numerous research inputs, but is also a professional association-led (AECT) collection of many research outputs. Obviously, the aim of this comprehensive *Handbook* is to provide state-of-the-art analyses, syntheses, and summaries of the theory and practice of ICT use in education and educational research in the U.S. and other countries. It systematically introduces and discusses the relevant (a) foundations, (b) strategies, (c) technologies, (d) models, (e) design and development, and (f) methodological issues of ET in education and educational research, with a special focus on each of these six major areas. Therefore, novice and experienced practitioners and researchers, as well as other interested faculty members, graduate students and readers, may find much of value in the way that this work highlights the meaningful interrelationships and relevant forms among users (including instructors and learners), ICT applications, and situated learning contexts that promote and enhance traditional and innovative technology-enhanced learning and instruction.

What makes this edition especially valuable is that the theoretical focus of the *Handbook* is provided in a series of chapters on the historical and theoretical foundations and evolution of this broad and engaging topic. As a trans-disciplinary field, ET has been criticized by some for a lack of solid theoretical foundations (Mishra & Koehler, 2006). Fortunately, the seven chapters in “Part I: Foundations” provide a series of convincing historical, theoretical, and philosophical concepts, along with their background and development, in order to offer multiple perspectives on ET. The main purpose of the these chapters is to help this professional field establish a fundamental knowledge base—as drawn from the past and present academic, educational, and inquiring experiences in ET and other relevant disciplines. Consequently, interested researchers and practitioners can utilize this knowledge to meaningfully construct their own views on ET research and development. For instance, Chapter 1 reviews the historical foundations of this field, taking a broad view of educational media; Chapter 2 analytically explores the theoretical foundations of this field in four relevant areas (including the psychology of learning, communications theory, human-computer interaction, and instructional design and development); Chapter 7 meaningfully constructs the philosophical foundations for the ET field with the seven philosophical perspectives (including Objectivism, Realism, Empiricism, Rationalism, Idealism, Relativism, and Pragmatism) and the five psychological perspectives (including Behaviorism, Cognitivism, Cognitive contructivism, Sociocultural/historicism, and Situativity theory). Readers will benefit from the clearly structured overview of the milestones and key players in the foundation, development and evolution of ET that is presented in this part of the work.

Most of us realize that the aim of ET is making effective and efficient use of appropriate ICT in various situated learning contexts for educational purposes. In "Part II: Strategies", readers may learn a variety of helpful instructional or learning strategies that can be used in specifically defined learning environments. The authors do a great job in these seven chapters of clearly demonstrating the use of various strategies, with vivid and inspiring descriptions, explanations, and examples. For example, Chapter 8 uses illustrations and figures to demonstrate how to apply research-based guidelines to support learning with various types of media. These guidelines are derived from four related learning theories: (a) information-processing theory, (b) dual-coding theory, (c) cognitive load theory, and (d) Baddeley's model of memory. In addition, Chapter 14 discusses comparisons of Merrill's first principles of instructions and other recently developed instructional design principles, in order to stimulate more rigorous research to evaluate the validity of such principles. These research-driven strategies and guidelines for real practices consciously and concisely offer critical thinking with regard to the design and development of ET, all embedded in the form of the tactics, models, figures, and tables that are used throughout this part of the book.

The sixteen chapters in "Part III: Technologies," which constitute the largest part of the *Handbook*, identify, introduce and discuss the so-called "upstream technologies" (referring to analysis, planning, and design) as well as "downstream technologies" (referring to development, deployment, and evaluation), by considering various ICT uses in education. Some prominent topics in this part include programmed technologies, computer-mediated technologies, knowledge-based technologies, blended learning, adaptive technologies, learning objects, and open source and open standards. These upstream and downstream technologies can be treated as what Heinich *et al.* (1999) term "soft technologies" (which refers to well-designed instructional processes, models and techniques that are developed with behavioral and social methods and theories in mind in order to bring about desired outcomes with the use of hard technologies, which are composed of hardware and software). In these chapters, readers will learn from the clear descriptions and commentary on the current use of various ICT tools with upstream and downstream technologies, as well as their impact on education.

In recent years, autonomous learning has received growing emphasis, so models that can guide, promote, and enhance effective and efficient learning are desired. Interested readers will find "Part IV: Models" valuable and important, because it introduces various new approaches to facilitate learning that are designed to be used in schools, universities, workplaces and beyond. These new models and approaches are well-designed and -developed in terms of soft technologies, with or without the use of ICT. Readers will explore the current prevailing research topics, including cooperative learning models, cognitive apprenticeship approaches, adaptive instructional systems, problem-based learning, performance improvement approaches, resource-based training, and domain specific approaches in this profession-oriented part of the work.

Practitioners and researchers may be curious as to why some stakeholders are doing better than others in similar contexts and situations, what the best practices are for these professional activities, and how they can develop themselves to become professional educational technologists or instructional designers of ET. The eleven chapters in "Part V: Design and Development" focus on the research towards professional practice and development in this field. Readers will realize the high-quality know-how from several significant dimensions of instructional design and technology, including instructional design competencies, task analysis, performance assessment, evaluation models and methodology, system design for changes, and others.

One of the apparent advantages of this *Handbook*, and of "Part VI: Methodological Issues," is that the research paradigm in technology-enhanced learning and teaching, and the relevant research on it, are consciously and explicitly introduced and discussed. Readers will benefit from the four chapters on theory development, research designs, data collection and analysis, and foundations for the future in this part of the work. Chapter 54 especially focuses on the recognized research paradigms and the paradigm shift in research with regard to instructional design and methodological approaches of ET in this ever-changing professional field. This chapter also provides help on how to identify, design, and investigate research questions—in order to choose the appropriate method with regard to quantitative, qualitative, or other inquiry research. Just as Shih *et al*. (2008) try to identify the current research trends and possible new research directions for e-learning studies, the eight co-authors of Chapter 54 discuss the research papers published in the *Educational Technology Research & Development* between 1994 and 2005 to provide new research directions and research topics, along with the relevant research methodologies and issues. Moreover, readers will learn much important know-how with regard to how to collect and analyze formative data in various phases of their inquiries in Chapter 55.

In terms of structure in this *Handbook*, the four co-editors arrange related topics in the same part in order for readers to systematically capture the overview and the main ideas of the related chapters. In terms of format within each chapter, the co-authors thoughtfully provide an abstract, keyword definitions, an introduction, the main texts, and references, so that readers can better understand the content and be able to continue reading further on their own. Theoretically speaking, this well-chunked food for thought will certainly inspire readers' reflections on the information it contains. Technically speaking, the add-on information in each chapter means that they are all clearly organized texts that provide the key terms and main concepts in each of the well-specified domains and areas that they cover.

One minor shortcoming of this *Handbook* is the lack of an introduction to cutting-edge technologies and their possible applications in education, as well as how we can learn from them. For example, context-aware ubiquitous computing technologies (Hwang *et al*., 2009; Liu & Hwang, 2009) foster immediate learning by using sensors and RFID readers and tags, or by using the Global Positioning System (GPS; Ogata *et al*., 2008). But educational applications of these two types of new ICT are lacking in the *Handbook*. Although such technologies are not yet in common use, possibly due to issues of cost or inaccessibility, I believe that most readers would enjoy the opportunity to learn more about possible directions in formal and informal learning with the most advanced technologies, and hope that the next edition of the *Handbook* will provide such a chapter.

I would also like to suggest that a more comprehensive survey should be conducted in order to receive a broader perspective concerning the research methodologies and topics that should be included in the next edition, as well as new ways to develop or identify such research tools and directions. Maybe the editor of the next edition of the *Handbook* or other interested authors should conduct an online survey of AECT members and non-members, and of researchers and practitioners in developed, developing, and under-developed countries, in order to uncover unexplored or unidentified research issues and directions in ET and ICT. This would then make the next edition even more useful in enabling readers to conduct novel or innovative studies based on emerging ICT in a broad range of contexts, learning situations and educational settings.

However, these are minor criticisms, and overall I feel that the *Handbook of Research on Educational Communications and Technology (third ed.)* could well be *Educational Technology & Society*'s Best Research Handbook for 2009, if such an award existed. I highly recommend this collection of so many excellent works at such an affordable price to all novice and experienced stakeholders, graduate programs, and university libraries in our field and beyond, as I am convinced that it will enable readers to conduct innovative and beneficial research in educational communications and technology.

Acknowledgements

This work was supported by the NCKU Project of Promoting Academic Excellence (R021) and the National Science Council (NSC 98-2631-S-024-001) in Taiwan.

References

Ely, D. (2008). Frameworks of educational technology. *British Journal of Educational Technology, 39*(2), 244-250.

Heinich, R, Molenda, M, Russell, J. D. & Smaldino, S. E. (1999). *Instructional media and technologies for learning (6th edn.)*. Prentice-Hall, Upper Saddle River, NJ.

Hwang, G. J., Yang, T. C., Tsai, C.C. & Yang, S. J. H. (2009). A context-aware ubiquitous learning environment for conducting complex science experiments. *Computers & Education, 53*(2), 402-413.

Liu, G. Z. (2008). Innovating research topics in learning technology: Where are the new blue oceans?. *British Journal of Educational Technology, 39*(4), 738-747.

Liu, G. Z. & Hwang, G. J. (2009). A key step to understanding paradigm shifts in e-learning: Towards context-aware ubiquitous learning. *British Journal of Educational Technology, 40(6),* doi: 10.1111/j.1467-8535.2009.00976.x

Mayadas, A. F., Bourne, J. & Bacsich, P. (2009). Online education today. *Science, 323*(5910), 85-89.

Mishra, P. & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017-1054.

Ogata, H., Saito, N. A., Paredes J. R. G., San Martin, G. A., & Yano, Y. (2008). Supporting Classroom Activities with the BSUL System. *Educational Technology & Society, 11* (1), 1–16.

Shih, M., Feng, J. & Tsai, C. C. (2008). Research and trends in the field of e-learning from 2001 to 2005: A content analysis of cognitive studies in selected journals. *Computers & Education, 51*(2), 955-976.

http://www.ifets.info

ISSN: 1436-4522 (online)
ISSN: 1176-3647 (print)