Evaluating ‘ThinknLearn’: A Mobile Science Enquiry Based Learning Application in Practice

Sohaib Ahmed  
Institute of Information & Mathematical Sciences (IIMS),  
Massey University, Auckland, New Zealand  
s.ahmed@massey.ac.nz

David Parsons  
Institute of Information & Mathematical Sciences (IIMS),  
Massey University, Auckland, New Zealand  
d.p.parsons@massey.ac.nz

ABSTRACT
There is growing interest from science educators and researchers to develop technology-assisted enquiry based learning environments in the domain of school science education. Traditionally, school science education has been dominated by deductive and inductive styles of enquiry investigations, while the abductive style of enquiry investigation has previously been sparsely explored in the literature related to technology enhanced learning. We have therefore designed and evaluated a mobile learning application ‘ThinknLearn’ for the abductive style of enquiry investigation. This study uses the M3 evaluation framework for evaluating this designed application with high school science students. The results indicated in this paper showed improvements in the students’ understanding of the learning domain as well as developing their positive attitudes towards mobile learning.

Author Keywords  
Abductive, enquiry based learning, mobile learning, science education, technology-assisted learning

INTRODUCTION
Enquiry based learning is a pedagogical approach in which learners get knowledge through exploration and investigation with authentic situations, and develop their higher level thinking skills (Lim, 2004; Shih et al., 2010). It is suggested that these learning activities foster learners’ motivation and interest in science (van Joolingen & Zacharia, 2009). Rapid advances in digital technology have increasingly attracted the interest of science educators and researchers for developing systems to support learning experiences about the sciences. In recent years, the use of mobile technologies has increasingly supported access to web-based contents on-the-go “anywhere, anytime” because of their portability (Svetlana & Yonglk-Yoon, 2009). Further, these technologies not only support the learning experience inside the school (e.g. lab, classroom, library) but allow learners to perform enquiry based learning activities in natural environments (e.g. park, woodland, museum) (Rogers et al., 2005). This is due to the affordances of these technologies that offer many different levels of engagement (Looi et al., 2010). This makes it possible to build learning environments that can enable enquiry based learning activities in multiple contexts.

In the literature related to school sciences, a number of mobile science enquiry based learning applications have been discussed which reflect the diversity of enquiry investigations and their use in both indoor and outdoor settings. Among these, the Ambient Wood Project (Rogers et al., 2005) and Savannah (Facer et al., 2004) are well-known mobile science projects in which learners are engaged in science learning activities by exploring virtual or natural environments in outdoor settings. There are some other recent mobile science projects in which learners are involved in both indoor and outdoor learning activities such as LET’s Go! (Vogel et al., 2010) and nQuire (Sharples et al., 2011). For indoor school settings, BioKIDS Sequence (Parr, Jones, & Songer, 2004) and WHIRL (Yarnall, Shechtman, & Pennel, 2006) use mobile technologies in science classrooms in order to support more frequent assessment practices.

There are many studies (Chen et al., 2008; Huang, Lin, & Cheng, 2010; Shih et al., 2010) showing that the participants enhanced knowledge significantly when they were equipped with mobile learning applications as compared to the traditional ways of science learning. Specifically to hypothesis formation activities, there are some studies found in the literature that highlight the importance of technology-assisted environments which can help students to construct scientific hypotheses and their explanations during science enquiry investigations (Mulder et al., 2010; Oh, 2011; Peker & Wallace, 2011).

Most of these mobile science learning applications follow a hypothetico-deductive or inductive means of enquiry investigation in which learners are required to process ideas (or hypotheses) (Grandy, & Duschl, 2007). In contrast, abductive scientific enquiry emphasizes the development of hypotheses observed from the natural environment (Oh, 2011). In the technology enhanced learning literature, this kind of enquiry has not been previously exploited (Grandy & Duschl, 2007; Oh, 2011). Since no previous studies have demonstrated the benefits of mobile learning in hypothesis formation activities in the context of abductive science enquiry investigations, this provides us with an opportunity to explore some new approaches to technology-assisted learning in the sciences.
**ABDUCTIVE SCIENCE ENQUIRY**

In enquiry based learning, one of the important learning activities is to provide scientific explanations of natural phenomena. An abductive science enquiry is also implemented in a similar way that leads learners towards new explanations on the basis of background theories and observations (Raholm, 2010). In this trait of enquiry, learners are not sure about the conclusions but they get some possible explanations of a given problem, and those potentially possible explanations guide learners to construct some meaningful learning (Eriksson & Lindstrom, 1997). Substantially, this is the essence of abduction that it starts with the incomprehensive nature of explanation and concludes with the construction of satisfactory new knowledge by relating observed phenomena and the underlying concepts of a given domain (Raholm, 2010).

The concept of abduction was coined by C.S. Peirce (1839-1914) who classified abduction as a form of inference. He further explained that the logic of scientific enquiry is divisible into three fundamental modes of inference (Raholm, 2010): (1) deduction or explicative inference (2) induction or evaluative inference and (3) abduction or innovative inference. The following example, taken from our domain of study, will show the relationships more clearly. Here, the ideas relate to black surfaced tins (cans) containing hot water losing heat more quickly than white or shiny surfaced tins. In these examples, the Case (Hypothesis), Result (Observation) and Rule (Condition or Suggestion) are defined to show the differences in order.

**Deduction:**
- **Rule** – The water particles in a black surfaced tin vibrate faster than the other tins.
- **Case** – A black surfaced tin absorbs more heat energy than the other tins.
- **Result** – A black surfaced tin cools more quickly.

**Induction:**
- **Case** – A black surfaced tin absorbs more heat energy than the other tins.
- **Result** – A black surfaced tin cools more quickly.
- **Rule** – The water particles in a black surfaced tin vibrate faster than the other tins.

**Abduction:**
- **Rule** – The water particles in a black surfaced tin vibrate faster than in the other tins.
- **Result** – A black surfaced tin cools more quickly.
- **Case** – A black surfaced tin absorbs more heat energy than the other tins.

From these examples, it can be observed that in both deduction and induction, a Case (Hypothesis) is processed with either a Rule or a Result to generate the other component, while in abduction, the Rule and Result are used together to find a Case. This trait of abduction is well-suited to enquiry problems in which learners are challenged to formulate scientific hypotheses and explain natural phenomena (Oh, 2011). Therefore, science educators and researchers have recently begun to study the process of hypothesis generation in the context of abductive enquiry investigations.

**THINKNLEARN: A MOBILE WEB APPLICATION**

In consultation with the science teachers from a local high school, we agreed on one of the science enquiry topics from the national standard science curriculum as the experimental context to test a mobile learning application that supports abductive enquiry. In this experiment, three tins with different surface colours are filled with boiling water in order to compare the way they radiate heat energy. Tin A is painted white, tin B black and tin C is shiny (unpainted). Learners have to formulate a hypothesis from collecting data about these tins and then explain it further as depicted in Figure 1a and 1b (further details about the application can be found in Ahmed, Parsons & Mentis (2012)).

This mobile web application ‘ThinknLearn’ follows the AIM (Abductive Inquiry Model) (Oh, 2011) which includes four phases: exploration, examination, selection and explanation. In the exploration phase, the application asks about the temperature of the various tins which were recorded by the students at a particular time interval after pouring boiling water in these tins as shown in Figure 1a. After submitting all values for the given tins, the application poses a series of Multiple Choice Questions (MCQs) regarding the collected values of these measures one-by-one in the examination phase. This feature makes students use their observational abilities to answer the given questions. Further, it gives suggestions based on the answers chosen by the students. This question-suggestion module of the application guides students towards a point where they are able to formulate hypotheses about the given measures and understand the knowledge presented in this application. These context-sensitive suggestions are generated from an ontology which may lead the students to think about the various aspects of heat energy related to different coloured surfaces. The ontology is used for the representation of the domain of interest (Uschold & Gruninger, 2004).

In the selection phase, students are asked to select one of the appropriate hypotheses about the observed phenomena as depicted in Figure 1b. There are two hypotheses defined in this application; one is related to the vibration of the water particles and the loss of heat energy from the different coloured tins while the second is about the heat absorption and the loss of heat energy from the different coloured tins. The application uses a random function to ask about one of these hypotheses. In addition, the application extracts all the possible hypotheses including one correct and other three
distracters by using the domain ontology and its inter-related concepts. At the end, students express their complete explanations of the observed phenomena in the explanation phase.

**EVALUATION**

It is suggested in (Sharples, 2009) that mobile learning applications evaluation can inform systems by examining how the learning activity and the underlying technology can be developed to enhance learning and offer new learning opportunities. For the purposes of evaluation, part of the M3 evaluation framework (Vavoula et al., 2009) is used in this study. This framework consists of three levels of granularity (Vavoula & Sharples, 2009); Micro, Meso and Macro. However, in this study, only two levels (i.e. Micro and Meso) are applied as defined in Table 1. The reason for not considering the Macro level at this stage is because this level is used to examine the longer term impact of the new technology on established learning practices (Vavoula et al., 2009). At this stage of the research, this level of evaluation is not yet possible.

<table>
<thead>
<tr>
<th>M3 Evaluation Framework Level</th>
<th>Evaluation Aspects</th>
<th>Form of Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Micro Level</td>
<td>Technology usability</td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td>Individual and group learning activities (Hypotheses Formation)</td>
<td>Semi-structured group discussion</td>
</tr>
<tr>
<td>Meso Level</td>
<td>Mobile learning experience as a whole</td>
<td>Pre-Post tests (Experimental and Control groups)</td>
</tr>
<tr>
<td></td>
<td>Learners’ cognitive skills and learning performance</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Evaluating 'ThinknLearn' using the M3 evaluation framework

**Experimental Design**

The rationale for this experimental design is to evaluate ‘ThinknLearn’. Evaluation at the Micro level includes the technological usability and utility of the application. The utility covers the guidance towards the construction of hypotheses about the underlying domain while the usability focuses on the capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions (Bevan, 2001). Three ISO metrics including learnability, operability and understandability are used (ISO, 2003) for that purpose. In addition, for exploring the quality of learners’ learning experiences in a mobile learning context, three ‘softer’ metrics of quality are applied; metaphor, interactivity and learning content (Parsons & Ryu, 2006).

Evaluation at the Meso level explores learners’ educational aspects such as mobile learning experiences, cognitive thinking skills and learning performance in abductive enquiries. This experimental design uses control and experimental groups to compare learning outcomes. A control group performed the heat energy experiment in the science laboratory using a “pre-test -> heat energy experiment -> post-test” method, where the participants carried out the learning activities without using ‘ThinknLearn’. The experimental group used the application ‘ThinknLearn’ while performing the same experiment in the science laboratory using a “pre-test -> heat energy experiment + using ThinknLearn -> post-test” method. The learning activities involved pre and post tests around hypothesis formation activities in the context of
abductive science enquiry. These tests consist of MCQs and open-ended question which assessed learners’ knowledge about the topic covered in their science class.

This experiment was a between-subject design. The two ways of generating hypotheses and improving learning are considered as the independent variables; using the application ‘ThinknLearn’ versus the traditional approach, while the dependent variable is learning performance in this experiment. The measurement of the learning performance assesses how well each participant has learnt the given science content (i.e. heat transfer energy) while performing abductive science enquiry investigations.

Participants
A total of 161 students from six science classes voluntarily participated in this experiment. They were all NCEA level 1 science students, from Albany Senior High school, Auckland, aged 15-16 years. One of the groups was treated as an experimental group which comprised 86 students from three science classes. The other 75 students were a control group. In the experimental group, 86 students filled in the questionnaire and participated in group discussions while 81 students participated in pre-post activities. In the control group, 75 students were involved in pre-post activities without using ‘ThinknLearn’.

For the distribution of the groups (i.e. experimental and control), science teachers were insistent on keeping the class structure intact. Therefore, students could not be randomly assigned to any of the groups. However, three classes apiece were selected as experimental and control groups respectively. In each class, there were previously 8-9 sub-groups for performing their science classroom activities. So, we continued with this distribution and conducted this experiment in the second week of February, 2012.

Apparatus
Both groups were provided with three different coloured surface tins including Black, Shiny (Silver), and White. In addition, the experimental group was equipped with WiFi enabled mobile devices. The control group were required to perform the experiment in the traditional way (i.e. without any mobile devices). Both groups had used the same concepts related to the given topic (i.e. heat energy transfer) which were already covered in their earlier classrooms. For the experimental group, the ‘ThinknLearn’ application was used to assist the participants to understand those concepts using their mobile devices. For pre and post activities, a MCQ quiz was provided with the instructions according to each group. Further, a questionnaire was also given to each participants in the experimental group, to investigate their individual learning experiences when using the application.

![Figure 2 Experimental group participants performing the experiment](image)

Procedure
Initially, science teachers introduced the information about the experiment, the purpose of the study, and the data collection process for each group. Both groups of participants were first asked to answer the pre-test that consisted of four MCQs. Following this, they were required to perform the heat energy transfer experiment as depicted in Figure 2. In this experiment, they found some data values related to each tin. These data values helped them to understand some key concepts discussed in the given topic. At the end of the experiment, they were asked again to answer the same MCQs with the addition of one open-ended question related to the hypothesis, with its explanation. This open-ended question was used to understand how well the participants engaged in the learning and thinking process during the enquiry investigation. For evaluating the usability and the utility of the application, experimental group participants were also required to individually rate a 9-question questionnaire on a five-point Likert scale. They were also involved in semi-structured group discussions. In these discussions, participants were posed three questions which were related to the usability and the softer aspects of the application.
RESULTS AND DISCUSSIONS

Micro Level Evaluation
In this micro level evaluation, the responses from a questionnaire and semi-structured group discussions were gathered from the experimental group participants about their learning experiences while using ‘ThinknLearn’. The control group participants were not involved at this level of the evaluation.

Questionnaire Responses
The questionnaire was filled in by the participants after they had finished their mobile science enquiry learning activities. The 9 questions in the questionnaire attempted to address different aspects of usability (learnability, understandability, operability) and softer aspects (metaphor, interactivity, learning content) as shown in Table 2. The questionnaire used a five-point Likert scale where 1 was ‘strongly disagree’ and 5 was ‘strongly agree’. When this application was tested with science students, their overall responses were encouraging.

<table>
<thead>
<tr>
<th>No.</th>
<th>Statements</th>
<th>Evaluation Aspects</th>
<th>Mean Response ± S.E (Standard Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>This mobile learning experience was enjoyable.</td>
<td>Learning Content</td>
<td>3.66 ± 0.11</td>
</tr>
<tr>
<td>S2</td>
<td>This mobile application was easy to use.</td>
<td>Learnability</td>
<td>3.66 ± 0.11</td>
</tr>
<tr>
<td>S3</td>
<td>Navigation through this application was easy.</td>
<td>Operability</td>
<td>3.87 ± 0.11</td>
</tr>
<tr>
<td>S4</td>
<td>This application guides me to formulate a hypothesis.</td>
<td>Understandability</td>
<td>3.45 ± 0.11</td>
</tr>
<tr>
<td>S5</td>
<td>The given suggestions in the application were relevant.</td>
<td>Metaphor</td>
<td>3.74 ± 0.10</td>
</tr>
<tr>
<td>S6</td>
<td>This application helps me understand the relationships between different variables.</td>
<td>Interactivity</td>
<td>3.50 ± 0.11</td>
</tr>
<tr>
<td>S7</td>
<td>The given suggestions help me to understand the topic.</td>
<td>Metaphor</td>
<td>3.54 ± 0.10</td>
</tr>
<tr>
<td>S8</td>
<td>This application helps me to improve my reasoning skills.</td>
<td>Interactivity</td>
<td>3.28 ± 0.12</td>
</tr>
<tr>
<td>S9</td>
<td>It is an effective learning application.</td>
<td>Learning Content</td>
<td>3.55 ± 0.11</td>
</tr>
</tbody>
</table>

Table 2 Questionnaire statements asked to the experimental group participants

The three questionnaire statements (S2, S3, and S4) were intended to investigate the mobile learning application from a usability perspective. The responses to statements ‘S2’ and ‘S3’ revealed that the participants found this application was not difficult to use and navigation was straightforward. The ratings on the statement ‘S4’ also revealed that our respondents perceived that the guidance towards hypothesis generation and the whole learning process was very easy to understand. A one sample T-test against the neutral value 3.00 confirmed these interpretations (t_{S4} = 6.14, p< .01 for S2; t_{S3} = 7.88, p< .01 for S3; t_{S4} = 4.20, p< .01 for S4).

The responses for the statements about softer aspects of the quality of the application showed positive attitudes from the participants. According to the results, the participants experienced ‘ThinknLearn’ as an interactive learning application (t_{S3} = 4.62, p< .01 for S6; t_{S3} = 2.38, p< .01 for S8). The statements S1 and S9 revealed that our participants considered that this application was an enjoyable learning experience and by and large an effective learning application (t_{S3} = 5.74, p< .01 for S1; t_{S3} = 5.11, p< .01 for S9). Similarly, the responses to the statements (S5, S7) showed that the participants experienced an overall vision of the learning processes (t_{S5} = 7.10, p< .01 for S5; t_{S5} = 5.19, p< .01 for S7). The processes include the relevancy of the given suggestions for constructing scientific hypotheses and the assistance provided by these suggestions to comprehend the given topic.

Semi-Structured Group Discussions
The group discussion questions were posed to the participants about their learning experiences as mentioned in Table 3. In these group discussions, 25 groups from three science classes participated. According to these discussions, the responses from the participants were straightforward and interesting.

With respect to question 1 responses, the most of the participants considered that ‘ThinknLearn’ was easy to use and they did not find any difficulty while using it. However, there were a few who found this application difficult in terms of its guidance towards hypothesis generation. One of the groups highlighted that “... questions were difficult and the given suggestions were not easy to understand”. It could be argued that they did not relate suggestions to understand the given topic. In another instance, one participant of another group described how “it was not difficult but confusing on some occasions”. Those participants who considered the application a bit confusing and difficult application did not understand the deliberate purpose of this application to exploit their high level skills of critical thinking in such science enquiries.
As far as the second question is concerned, almost all the participants were positive about their learning experiences and they enjoyed using ‘ThinknLearn’. One of the group participants stated that “we really enjoyed using it. This application was pretty good and engaging, it helped you to learn about your course (science)”. The other group participants gave an interesting comment about it as “this type of application keeps you on focus and requires better attention”. On the contrary, one group of the participants disliked this application. According to them, “it was boring and confusing and therefore, we did not like it”. Overall, participants valued the interactivity, enjoyed the innovative way of learning, and found the application engaging.

In the responses for question 3, the respondents believed that the given suggestions were relevant and made them think. One of the group participants indicated that “these suggestions are relevant to the answers but they make us to think”. On the other hand, there were a few groups who remarked that “…more detail should be provided” and “…relevant but they (suggestions) did not explain much”. These comments showed that this application presents some challenges to the participants and made learners to think about the given topic. It may be argued that a certain level of challenge was maintained in this application to make it more engaging and interesting. However, some ways may be needed to convince those participants about the value of this approach. Overall, the group discussion responses suggest that the application was engaging and the given suggestions make learners think about the knowledge space under investigation, and may exploit their cognitive thinking skills.

<table>
<thead>
<tr>
<th>No.</th>
<th>Group Discussion Questions</th>
<th>Software Quality Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What type of difficulty do you find in using this application?</td>
<td>Usability aspects</td>
</tr>
<tr>
<td>2</td>
<td>How do you feel after using this application?</td>
<td>Softer aspects</td>
</tr>
<tr>
<td>3</td>
<td>What do you think about the suggestions given in the application?</td>
<td>Softer aspects</td>
</tr>
</tbody>
</table>

### Table 3 Group discussion questions

#### Meso Level Evaluation

This level was used to examine the learning performance between experimental and control groups. It involved pre and post activities including answering MCQs and writing hypotheses with their explanations while performing science experiments.

#### Learning Performance

In pre-post tests, participants were asked to answer MCQs related to the learning domain. In comparing these two groups, an independent sample t-Test was used to find out the learning performance differences. The results showed a significant difference (p = .025) with and without using ‘ThinknLearn’. As a matter of fact, the control group participants got marginally better scores in their pre-tests as compared to the experimental group participants. However, in the post-tests, both groups improved but the experimental group gained more in learning performance than the control group as depicted in Figure 3.

![Learning Performance](image)

**Figure 3 Pre-Post tests comparison between experimental and control groups**

In the post-tests, both group participants were asked to write a hypothesis about the colour of any of these three tins and its explanation in the open-ended dialog box. As far as the marking of the open-ended question was concerned, it was mutually decided with the science teachers to mark thus: ‘0’ for wrong hypothesis or leave empty; 0.5” for correct hypothesis but wrong explanation; ‘1’ for correct hypothesis with its explanation. As an example, one of the answers from the participants who got ‘1’ mark for a correct hypothesis with its explanation was “Black tin absorbs more heat energy than the other tins and loses more heat energy than the others therefore it keeps the water cool from the inside”. Given that such answers are open to interpretation, there is the potential for bias which should be taken into account when analysing our results.
According to the applied independent sample t-Test, the results showed a significant difference ($p = .017$) between the experimental and control groups. The experimental group participants got improvements in their thinking and learning while formulating hypotheses about the learning domain. However, the control group participants did not appear to understand the given topic so well and therefore, were not able to formulate hypotheses and their explanations at the same level as the experiments group. The participants’ scores in percentages confirming these interpretations as illustrated in Figure 4.

**Figure 4** Comparison between experimental and control groups in hypothesis formation activity

**CONCLUSION AND FUTURE WORK**

The empirical data presented here make a case for the use of ‘ThinknLearn’ and provide some insights as to why it might be a more effective way of generating scientific hypotheses than the traditional pedagogy. This innovative application presented a case for the practical implementation of mobile abductive science enquiry applications. It can be suggested that this kind of application may be useful to enhance both learning performance and cognitive thinking skills where learners are engaged in exploring and experimenting in real environments. This can promote deeper understanding of a particular science domain and can guide learners in interpreting data to create meaningful hypotheses.

Applications of this kind can be extended further to target other professional fields of interest where researchers or educators want to explore abduction as a form of reasoning such as medical diagnostics, jury deliberation, scientific theory formulation, accidental investigations etc. Moreover, this study may be further developed to support the practical implementation of abduction theory in school sciences, which has not previously been explored.

**ACKNOWLEDGMENTS**

We acknowledge Diana Hartley and her colleagues for assisting us in conducting experiments with their students at Albany Senior High School, Auckland, New Zealand.

**REFERENCES**


