Supporting collaborative inquiry during a biology field trip with mobile peer-to-peer tools for learning: A case study with K-12 learners

Jari Laru*, Sanna Järvelä

Department of Educational Sciences and Teacher Education, University of Oulu, Oulu, Finland

Roy B. Clariana

School of Graduate Professional Studies, Penn State University, USA

*Mr. Jari Laru (M.Ed)
Learning and Educational Technology Research Unit (LET)
Department of Educational Sciences and Teacher Education
P.O. Box 2000
FIN-90014 University of Oulu
Email: jari.laru@oulu.fi
Fax: +358-8-553 3744

(Submitted January 23, 2009; Accepted March 8, 2010)

This study explores how collaborative inquiry learning can be supported with multiple scaffolding agents in a real-life field trip context. In practice, a mobile peer-to-peer messaging tool provided meta-cognitive and procedural support, and tutors and a nature guide provided more dynamic scaffolding in order to support argumentative discussions between groups of students during the co-creation of knowledge claims. The aim of the analysis was to identify and compare top- and low-performing dyads/triads in order to reveal the differences regarding their co-construction of arguments while creating knowledge claims. Although the results revealed several shortcomings in the types of argumentation, it could be established that differences between the top performers and low performers were statistically significant in terms of social modes of argumentation, the use of warrants in the mobile tool and in overall participation. In general, the use of the mobile tool likely promoted important interaction during inquiry learning, but led to superficial epistemological quality in the knowledge claim messages.

Keywords: collaborative learning; mobile learning; inquiry learning; argumentation; scaffolding

Introduction
Visits to natural settings provide enjoyable, rich, authentic, and salient real-life experiences that can enhance science learning (Anderson, Thomas, & Nashon, 2008) and are a fundamental part of many science courses (Soloway et al., 1999). While learning activities already occur outdoors, new mobile technologies can help students carry out
scientific inquiry in the context of discovering and exploring an environment (Price & Rogers, 2004; Looi et al., 2009). However, outdoor field trips and computer-based indoor learning activities are typically performed separately. For example, observations and data collections are done on a field trip, and then further analyses are carried out back in the classroom with the help of various software packages (Novak & Krajcik, 2005; Price & Rogers, 2004).

In general, the various educational affordances of wireless technologies suggested by researchers (Looi et al., 2009; Roschelle & Pea, 2002) have paved the way for the emergence of so-called mobile learning or ubiquitous learning initiatives, such as G1:1 learning (Chan et al., 2006). However, much of the research has been driven by the technical capabilities of new devices, while their true strengths and weaknesses and the application of theory in the use of these technologies for educational purposes must be further explored (Fischer & Konomi, 2007; Naismith, Lonsdale, Vavoula & Sharple, 2005).

The main task of this study was to explore how dyads and triads of learners can be supported by multiple scaffolding agents in an outdoor inquiry learning context. In practice, a mobile peer-to-peer messaging tool provided meta-cognitive and procedural support, while tutors and a nature guide provided more dynamic scaffolding in order to support the groups’ dialectical argumentation and co-creation of knowledge claims. The aim of the analysis was to identify and compare top- and low-performing dyads/triads in order to reveal the differences in their co-construction of arguments and knowledge claim messages.
Collaborative Argumentation During Inquiry Learning

Progress has been made over the last ten years in regards to understanding how argumentative interaction leads to cognitive changes during collaborative learning (e.g., Andriessen, Baker & Suthers, 2003). For example, one particular form is termed ‘broadening and deepening’ a space of debate, resulting in a richer representation of ideas (van Amelsvoort, Andriessen & Kanselaar, 2007). Furthermore, researchers have developed instructional design principles from a socio-constructivist approach for inquiry learning that includes principles of student-centered, collaborative, and problem-driven learning. For example, during inquiry, learners are offered ill-structured but authentic learning tasks that are designed to trigger higher-order thinking and reasoning skills, rather than just memorisation of facts (Clarebout & Ellen, 2001).

In general, inquiry learning emphasises active learning, encouraging students to ask questions, formulate hypotheses, and to test the hypotheses through experimentation. In a more detailed way, the process of inquiry can be characterised by repeatable phases as described; for instance, in a model of progressive inquiry (Hakkarainen, 2002). Collaborative argumentation is claimed to be a core activity practiced by learners during inquiry. In formulating arguments and debating with peers about which piece of evidence supports a particular premise, learners can acquire both argumentation skills and domain-specific knowledge, but may also even construct new knowledge (Weinberger, Stegmann, Fischer & Mandel, 2006).

In this study, groups of learners were engaged in an argumentative discussion (i.e., dialectical argumentation) (Driver, Newton & Osborne, 2000) in order to collaboratively co-construct knowledge claims during a task (data, claim, warrant, as
cited in Toulmin, 1958) based on a condensed argumentation model that was suggested by Weinberger, Stegmann, Fischer, and Mandl (2006).

**Scaffolding Collaborative Argumentation with Mobile Technologies**

In order to favour the emergence of productive interactions and to improve the quality of argumentation, adequate scaffolds must be provided (Sharma & Hannafin, 2007; Weinberger et al., 2006). Scaffolding makes learning more tractable for students by changing complex and difficult tasks in ways that make these tasks accessible, manageable, and within the students’ zone of proximal development (ZPD) (Vygotsky, 1978).

Although the original concept of scaffolding addressed learning in face-to-face situations (Wood, Bruner & Ross, 1976), technology research in the learning sciences has illuminated some ways in which technological learning tools may provide different scaffolding functions (Joolingen, Jong & Dimitrakopoulou, 2007; Sharma & Hannafin, 2007). Instead of a single knowledgeable person or software providing support, there are multiple ZPDs consisting of tools and resources, peers, or the learning environment itself (Puntambekar & Hubscher, 2005). In such cases, technology can provide procedural and metacognitive support for routine tasks and allow the teacher to provide dynamic support (Kuhn, 1991, as cited in Jonassen & Kim, in press; Sharma & Hannafin, 2007).

In this study, tutors and a nature guide scaffolded an argumentative discussion by asking questions that were based upon students’ argumentative skills. In addition, mobile technology provided both metacognitive and procedural scaffolding in the form of a knowledge claim message template and storyboard message.
This study considers how collaborative inquiry learning can be supported with multiple scaffolding agents within an outdoor field trip context. We anticipate that top performers will create better knowledge claim messages, and that the structural quality of their argumentative discussion will be better than that of low performers.

The specific research questions are:

1) What were the differences between top and low performers in regards to collaborative inquiry learning during the field trip? Specifically:
   a. To what extent were learners jointly engaged in collaborative inquiry learning during the different task phases?
   b. What was the nature of the argumentative discussion in top- and low-performing groups?
2) What was the difference between top and low performers in regards to the structural quality of knowledge claim messages?
3) How much did the top and low performers learn about biology during the field trip?

Methods
This study followed the principles of the case study method, which is defined as an empirical study that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between the phenomenon and the context are not clearly evident (Yin, 2003). The contemporary nature of the case study, combined with its practice and action-oriented approach, fits well with the purpose of this study.

Participants and Research Setting
The participants were primary school students (n=22, all 12 years of age) who participated in a one-day learning project during a field trip to a nature park in a wilderness forest setting in northern Finland. The field trip activities in this investigation were designed and developed by the research team in collaboration with the nature park’s local expert, who is a biologist. The students were randomly assigned to eight groups (six
triads and two dyads), and each group was provided with a mobile phone. Before the experiment, the principles and procedures of collaborative inquiry learning and argumentation were presented, and practical training for the field trip was given in the classroom by the researchers and the biologist.

**Task**
The student groups explored traces of inanimate and animate nature within a transect of a wilderness environment that straddled three distinct areas: (a) a densely wooded forest with dead trees; (b) a moderately wooded forest and swampy ground cover; and (c) a cliff-top area with several visible features from the ice age.

The pedagogical design consisted of four recurring phases: (1) encountering a problem presented by the scientist on the storyboard messages and the creation of group level presumptions by the dyads/triads; (2) small-group argumentative discussions and explorations in order for the students to produce their own knowledge claims about the findings in the form of a message; (3) review and comparison of the knowledge claims between dyads/triads; and (4) a teacher-led conclusive discussion between the dyads/triads.

In practice, phases 1-2 were introduced to students as short snippets (Figure 2B) of the storyboard narrative. Other snippets (task feedback, information, and conclusion) were presented when the task was completed, as students were walking between task areas, and at the end of the nature trail, respectively.

The storyboard narrative was based on the premise that a scientist in a distant country was preparing a book about Finnish nature and he needed help from local assistants (story snippet: grounding). In real time, the scientist asked for clarification...
about previous findings, by inquiring, “In this region, one should find some features of animated nature. Please help me find them and explain the reasons for them!” (translated from Finnish) (story snippet: task introduction). In order to answer this vaguely defined goal, the groups of learners were required to engage in argumentative discussion, which was aimed at the collaborative co-construction of a knowledge claim message of each designated area (Figure 1B).

**Tools**
The tools used were mobile phones and a prototype of a mobile Peer-to-Peer messaging application, called *Flyer*. Flyer is an example of a social proximity application, which belongs to an emerging class of mobile networks: mobile encounter networks (MENs). The technology used created a digital “sphere”, “field”, or “aura” surrounding each group by enabling their phones to broadcast information to, and fetch information from, nearby groups or storyboard phones (See Figures 1-2) directly without connection to a network or server. (cf. Volovikov, Juonoja, Weber, Kotilainen, Vapa & Vuori, 2008). Information in this digital realm was used to support and augment existing collaborative inquiry learning practises in real space, instead of using it as a collaboration tool per se.

The application of Flyer was adapted to suit collaborative and argumentative inquiry learning by employing a design that embedded procedural and meta-cognitive scaffolding into the interface design of the system in the form of knowledge claim message templates and fixed storyboard messages. The following section describes the
features of the Flyer prototype.

Figure 1 - Left: students working with a phone, A-B: Editing a knowledge claim message, C-D: Publishing and receiving a message.

*Creating Flyers.* In practise, students were asked to edit Flyer templates from saved Flyer folders (See Figure 1A). The Flyer editor allowed users to add a title, text, image, and choose the background colour. In order to constrain the argumentative discussion, student groups were cued to the main components of the knowledge claim message (metacognitive scaffolding) by the templates, which specified the components in reasoning from data to claim in the form of embedded sentence openers. Furthermore, sentence openers were provided as suggestions; students were able to ignore the openers, change them, or create new ones. Suggested sentence openers were always present and available to the students through the learning phases in each template, and provided five pre-defined structural components (see Figure 1B): (field 1) a research question for expressing group level presumptions; (fields 2 through 4) sentence openers for knowledge claim creation (e.g., claims, grounds, warrants); and (field 5) a photographic image (visual representation that supported the group’s claim).

Figure 2 - Left: Pushing storyboard messages, Right (A-B): Receiving a storyboard message

**Background receiving.** This feature scanned the environment for other Flyer users and Bluetooth devices, and presented found storyboard and knowledge claim messages in a list (Figure 2A, Figure 1D). The list displayed the subject of the message and the date and time the message was received. In practice, the *storyboard messages* functioned as activity placeholders (procedural scaffolding) for each of the four learning phases, while knowledge claim messages were artefacts created by the students. The former were automatically pushed to student groups’ phones at appropriate phases or places along the nature trail before and after students’ activities and the latter were spread to peer phones after they were published manually by the student dyads/triads.

**Data collection and analysis**
Audio recordings of group interactions, the knowledge claim messages, and paper-and-pencil based mindmaps were analysed. The audio recordings and messages were the primary measures of the influence of this pedagogical approach on student engagement (discussed in greater detail below). The pre- and post-mindmaps provided a measure of biology content knowledge and a means of forming top- and low-performing dyads/triads.
Analysis of mindmap tasks to answer the question: What did students learn during the field trip?

A pre-test/post-test quasi-experimental design was used for the mindmap task. First, the paper-and-pencil mindmap task was used as a measure of the students’ learning during the field trip. In the pre- and post-mindmapping tasks, the individual student was asked to write the two super-ordinate category terms ‘animate’ (living or once living) and ‘inanimate’ (inorganic) on paper, and then link concrete examples from nature to the appropriate category. This open-ended mindmap approach allowed students to show their understanding (for example, by assigning a bird’s track in the sand to the animate category), but also their misunderstandings (such as listing a bird in a tree as being inanimate). An increase in correctly organised knowledge, as well as a decrease in misunderstandings, signals student learning (Novak, 2003).

The mindmap scoring rubric considered both correct and erroneous associations. In each mindmap, in order to reduce the statistical influence of the participants who included many responses in their mindmaps, the sum of the errors was subtracted from the sum of the correct relationships, and this value was then converted to a scale from -3 to +3 using the following point system: -3 awarded for values less than or equal to -10; -2 awarded for values between -5 and -10; -1 awarded for values between 0 and -5; 0 awarded for 0 values, +1 awarded for values between 0 and 5; +2 awarded for values between 5 and 10; and +3 awarded for values greater than or equal to 10. Two raters scored all mindmaps individually using the system, and then met to reach a consensus when the initial scores were not in agreement with an inter-coder reliability of Cohen’s Kappa = 0.72, p < 0.0001. In terms of reliability, the pre-test animate and inanimate
subtest scores \(r = .86\) and the post-test animate and inanimate subtest scores \(r = .95\) were strongly related. Cronbach alpha reliability of the four subtest scores was 0.67.

Next, average mindmap scores for each dyad/triad were used to identify top-performing and low-performing groups for further qualitative analysis. Note that contrasting the activity and artifacts of top performers to those of low performers is intuitively appealing (Jonassen, Tessmer & Hannum, 1999), and has been shown to reveal important characteristics and aspects that are not uncovered using other approaches (Wyman & Randel, 1998). Specifically, this analysis focused on group differences in collaborative argumentation activities (i.e., verbal interactions and also the structural quality of knowledge claim messages). Given the relatively small sample size and skewed nature of some distributions, inter-group differences were analysed by non-parametric Mann-Whitney U-tests (Gliner & Morgan, 2000).

Analysis of group interactions in order to answer the question: What are the differences between top and low performers regarding collaborative inquiry learning during the field trip?

The audio recorded data was drawn from transcribed digital recordings of authentic interactions during the field trip and captured by means of personal digital audio recorders and lapel microphones that students (one student per dyad/triad) wore on their pockets during the field trip. The total data included eight 160-minute group recordings, of which approximately eight 120-minute recordings dealt with task-related interactions (‘off-task’ interactions, recorded while students were walking between locations, were excluded). Turn-taking was used as a criterion for the transcript corpus segmentation. The entire corpus comprised 3,593 utterances, and 30 percent of the data were coded by two independent coders with an inter-coder reliability of Cohen’s Kappa = 0.81, \(p < 0.0001\).
First, audio transcripts were analysed in order to answer the subquestion: To what extent were learners jointly engaged in collaborative inquiry learning during the different task phases? This analysis was adapted from the method that focuses on the duration of on-task and off-task episodes. (For details of this method, see Järvelä, Veermans & Leinonen, 2008.) In this analysis, the focus was placed on the number of social-mode utterances, which were used as a measure of on-task activities, while off-task activities were coded as their own off-task category. Second, on-task activities were further segmented in order to reveal the division of the argumentation into the respective task phases.

Second, in order to answer the subquestion: What was the nature of the argumentative discussions during the field trip?—the transcripts were analysed using categories of social modes of co-construction, which were developed originally as a framework for argumentative knowledge construction (Weinberger & Fischer, 2006). This categorisation was grounded on the idea that social modes differ to the degree in which learners refer to previous contributions of their learning partners, including: (1) **externalisation**, which refers to situations where she/he is articulating thoughts to the group; (2) **elicitation**, which includes utterances where she/he is questioning learning partner(s) or provoking a reaction from her/his learning partner(s); (3) **quick consensus building**, which refers to cases where she/he is accepting or affirming the contributions of her/his learning partner(s); (4) **integration-oriented consensus building**, which contains situations where a learner is taking over, integrating, and applying multiple perspectives of her/his learning partners; and (5) **conflict-oriented consensus building**, where a learner
Analysis of the knowledge claim messages in order to answer the question: What is the difference between top and low performers in terms of the structural quality of knowledge claims?

First, the structural components of the knowledge claims were analysed (Stein & Albro, 2001). The scoring rubric for analysis consisted of five structural components: (1) the research question; (2) the claim; (3) the ground; (4) the warrant; and (5) the photographic image. Each was worth of one point. Thus, the scoring rubric considered both correct and incorrect answers in each message, and the sum of the incorrect components was subtracted from the sum of the correct components. Second, in order to reveal the epistemological nature of the claims, warrants and grounds were evaluated against the condensed argument model (Weinberger et al., 2006).

The stored data consisted of all knowledge claim messages made by the students. The total data included 20 messages, as 4 messages were not published due to technical difficulties within their groups during the tasks. The entire data set consisted of 82 structural components, which were analysed by two independent coders, with an inter-coder reliability of Cohen’s Kappa  = 0.80, p <0.0001.

Results

How much did the top and low performers learn about biology during the field trip?

The mindmap data were analysed by a 2 (Group: eight dyads/triads) x 2 (Gain: pre-test and post-test) x 2 (Subtest: animate and inanimate) mixed ANOVA. The first is a between-subjects factor, and the second and third are within-subjects factors. Only the gain from pre-test to post-test was significant, with F(1,14) = 26.509, MSe = 1.545, p <

.001, indicating a substantial gain in mindmap scores from pre-test to post-test. This finding provides support for the use of software-supported argumentation scaffolds in field trip settings, and warrants close analysis of the learning process qualitative data that may explain these gains.

With the help of the mindmap gain scores [note that gain scores provide a measure of engagement in the task and are not really a measure of achievement, and so it makes good sense to use gain scores to determine top and low performer groups], top-performing and low-performing dyads/triads were identified for further analysis. The results show that there were four dyads/triads that were categorised as top performers (triads G2 and G6; dyads G4 and G5; gain score ≥ 1.5; X = 2.08; SD = 0.85), and four groups were categorised as low performers (triads G1, G3, G7, and G8; gain score < 1.49; X = 0.75; SD = 0.37).

**What were the differences between the top and low performers regarding collaborative inquiry learning during the field trip?**

First, in order to investigate to what extent learners were jointly engaged in collaborative inquiry learning during the different task phases, the amount and division of on-task utterances were analysed. The total sum of utterances (n = 2.069; M[SD] = 517.3[56.4]) in top-performing dyads/triads and their high-percentage share of the total number of utterances (70 percent of 2,970) reveal that top-performing groups engaged in twice the number of small-group interactions as their low-performing peers did (n = 901; M[SD] = 332.3 [67.4]). A Mann-Whitney U-test shows that this difference is also quite statistically significant (Z = -2.3; p < 0.01***). Furthermore, the Mann-Whitney U-test suggests that there were also significant differences in terms of on-task (Z = -2.309; p < 0.05) and off-
task ($Z = -2.309; p < 0.05$) interactions. However, the proportion of off-task interactions was quite low in both groups (less than 22 percent of the total share within each group, see Table 1), which suggests a high share of on-task dialogue during the activity.

Unexpectedly, almost all on-task interaction (89 percent) concentrated on the small group interactions of dyads/triads during task phase 2, while whole class activities led to only minor interactions and an almost total lack of counter-argumentation.

Table 1 - Social modes of co-construction of knowledge

<table>
<thead>
<tr>
<th>Social Mode</th>
<th>High Total</th>
<th>M</th>
<th>SD</th>
<th>Low Total</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>180</td>
<td>46.5</td>
<td>12.9</td>
<td>108</td>
<td>35</td>
<td>13.9</td>
</tr>
<tr>
<td>EL</td>
<td>505</td>
<td>126</td>
<td>39.4</td>
<td>279</td>
<td>97</td>
<td>33.5</td>
</tr>
<tr>
<td>QCBO</td>
<td>31%</td>
<td>58%</td>
<td>70%</td>
<td>37%</td>
<td>76%</td>
<td>76%</td>
</tr>
<tr>
<td>IOCB</td>
<td>27%</td>
<td>36%</td>
<td>76%</td>
<td>25%</td>
<td>19%</td>
<td>17%</td>
</tr>
<tr>
<td>COCB</td>
<td>10%</td>
<td>7%</td>
<td>71%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

| Abbreviations: E: Externalization; EL: Elaboration; QCBO: Quick Consensus Building; IOCB: Integration-Oriented Consensus Building; COCB: Conflict-Oriented Consensus Building |

Second, in order to investigate the nature of the argumentative discussions within the top- and low-performing groups, the audio transcript data were analysed using categories of social modes of co-construction. In spite of significant differences between high and low performers in total interactions, the overall patterns of social modes reveal strong similarities between these groups (see Table 1). Specifically: (1) in all groups, questioning and provoking reactions from learning partners or tutors was a very common social activity (elicitation, \( \geq 31\% \)); (2) students also frequently accepted or affirmed peer learners’ contributions in order to move on with each task (quick consensus building,
≥25%); (3) results further show that disagreeing or replacing the perspectives of learning partners remained low in both high- and low-performing groups (conflict-oriented consensus building, ≥17%); (4) in addition, in all groups, students made only a few contributions to the discourse without reference to other contributions (externalization, ≥11%); and yet, (5) the amount of integration of each other’s perspectives—and thus making sense of the task—remained very small (integration-oriented consensus building, ≥7%).

However, when the social modes of co-construction were further analysed using the Mann-Whitney U-test, differences were found between high and low performers. Although the number of integration-oriented consensus building and conflict-oriented consensus building utterances were low, differences were shown to be significant in these categories (IOCB: Z = -2.309; p < 0.05 and COCB: Z = -1.73, p < 0.1), and also in the quick consensus building category (Z = -2.309; p < 0.05).

What is the difference between top and low performers in regards to the structural quality of knowledge claim messages?
In terms of the completion rate of message templates there were differences between high and low performers (e.g., 80 percent compared to 56 percent, respectively). However, these differences were not statistically significant according to the Mann-Whitney U-test (see Table 2). Considering specific components in the knowledge claims both high- and low-performing groups included claims and pictures in their messages equally. However, the rate of warrants and of grounds by the low-performing groups was considerably lower than that by the high performers. This difference in the usage of warrants was statistically significant according to the Mann-Whitney U-test (Z = -1.95, p = < 0.05).
Epistemological analysis of the content of the knowledge claim messages (see example: Figure 3) revealed that knowledge claims included students’ guesses and suggestions, which were little more than observable surface features. Students were not necessarily producing high-level theoretical explanations of the inanimate or animate environment behind the observed areas during the field trip. Most of the messages (n = 14, or 58 percent) contained argumentation about possible reasons for dysmorphism, or the dead trees in the forest, while 21 percent (n = 5) included reasoning about possible causes of cracked stones or bedrock. The other 21 percent (n = 5) of messages were either observations regarding human footsteps, logging areas, or fungi (n = 3, or 13 percent), or were missing (n = 2, or 8 percent).

Table 2. Structural components of knowledge claim messages

<table>
<thead>
<tr>
<th>Structural components</th>
<th>Title</th>
<th>Claim</th>
<th>Ground</th>
<th>Warrant</th>
<th>Pic</th>
<th>TOTAL ³</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Subtotal¹</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td>M</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2.5</td>
<td>2.5</td>
<td>12</td>
</tr>
<tr>
<td>SD</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9</td>
<td>0.9</td>
<td>4.8</td>
</tr>
<tr>
<td>% of components¹</td>
<td>83%</td>
<td>83%</td>
<td>83%</td>
<td>75%</td>
<td>75%</td>
<td>80%</td>
</tr>
<tr>
<td>Low Subtotal¹</td>
<td>7</td>
<td>9</td>
<td>6</td>
<td>2</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>M</td>
<td>2</td>
<td>2.5</td>
<td>1.5</td>
<td>0</td>
<td>3</td>
<td>8.5</td>
</tr>
<tr>
<td>SD</td>
<td>1.5</td>
<td>0.9</td>
<td>1.7</td>
<td>1</td>
<td>1</td>
<td>5.3</td>
</tr>
<tr>
<td>% of components¹</td>
<td>58%</td>
<td>75%</td>
<td>50%</td>
<td>17%</td>
<td>83%</td>
<td>56%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>17</td>
<td>19</td>
<td>16</td>
<td>11</td>
<td>19</td>
<td>82</td>
</tr>
<tr>
<td>M</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1.5</td>
<td>3</td>
<td>10.3</td>
</tr>
<tr>
<td>SD</td>
<td>1.2</td>
<td>0.9</td>
<td>1.4</td>
<td>1.3</td>
<td>0.9</td>
<td>5</td>
</tr>
<tr>
<td>% of components</td>
<td>71%</td>
<td>79%</td>
<td>67%</td>
<td>46%</td>
<td>79%</td>
<td>68%</td>
</tr>
<tr>
<td>Z</td>
<td>-0.83</td>
<td>-0.5</td>
<td>-1</td>
<td>-1.95</td>
<td>-0.33</td>
<td>-1.315</td>
</tr>
</tbody>
</table>

Significance levels: *P < 0.1, **P < 0.05, ***P < 0.01
¹Component usage rate: 4 teams x 3 tasks x 1 component
²Completeness rate: 4 teams x 3 tasks x 5 components

<table>
<thead>
<tr>
<th>Group</th>
<th>Area</th>
<th>Question</th>
<th>Claim</th>
<th>Ground</th>
<th>Warrant</th>
<th>Pic</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>G3</td>
<td>Densely wooded forest with dead trees</td>
<td>Footprint (no explicit question) [-]</td>
<td>Someone has made a trace [+], Trace [-]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>G5</td>
<td>Moderately wooded forest and swampy cover</td>
<td>Why are there traces on the tree? [+], A woodpecker has made the traces [+], There are holes on the tree [+], Woodpeckers knock trees [+],</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. An example of epistemological analysis of knowledge claim messages

**Discussion**

Overall, the use of the mobile messaging tool likely promoted important types of argumentative interactions during inquiry learning. However, results revealed several shortcomings in the structure and division of the argumentative discussions throughout the task phases, and in regards to the superficial epistemological quality of the knowledge claim messages for some of the teams.

First, content knowledge gain from pre-test to post-test was statistically significant, indicating a substantial gain in mindmap scores from pre-test to post-test. This finding provides support for the use of software-supported argumentation scaffolds in field trip settings, and was further used to contrast top and low performers in the following analyses.

Second, statistically significant results from the Mann-Whitney U-test revealed that top-performing dyads/triads engaged in twice the number of argumentative
discussions during co-construction of knowledge claim messages as their low-performing peers did. However, almost all argumentative activities in each dyad/triad were concentrated in the second phase of the pedagogical design, while whole class activities led to only minor interactions and had an almost total lack of counter-argumentation. One reason for this finding is the problematic application of Toulmin’s model to situations that involve two or more arguments (Jonassen & Kim, in press). In the other words, that model was applicable during the small group phase, where the main target was to co-construct one knowledge claim message, but not during the latter activities where messages were evaluated and compared.

Third, analysis of the social modes of argument co-construction revealed statistically significant differences between top and low performers in the integration- and conflict-oriented consensus building categories. This finding suggests that the superior results of the high performers were the consequence of their slightly greater willingness to integrate, but also of their willingness to criticise each other’s contributions, which has been considered an important component in the socio-cognitive perspective of collaborative learning (Teasley, 1997). These two categories of argumentation may be subtle but strong indicators of task and social engagement, and the lack of these elements indicates task avoidance.

Although these differences were statistically significant, all student groups shared the same overall pattern of social modes. Elicitation and quick consensus building constituted over half of the total argumentation. Such a high proportion of elicitation and quick consensus building (i.e., uncritical affirmation) may have been detrimental for learning (Joolingen et al., 2007; Weinberger & Fischer, 2006), and may be one

explanation as to why the epistemological quality of the knowledge claim messages was poor. The high numbers for elicitation and quick consensus building suggests that the likely aim of learners was “completing the learning environment” rather than participating in the inquiry process itself (Joolingen et al., 2007, p. 113).

Alternately, these results may suggest that the learners’ main challenge or goal had been to meet the perceived requirements posed by the design of the experiment by using tutors and peers as a shortcut learning resource—an approach referred to as ‘soft’ scaffolds (Sharma & Hannafin, 2007, p. 39), in contrast to the ‘hard’ argumentation scaffolds provided by the messaging tool. Contrary to the explicit failures with the ever-present technological scaffolds for co-construction of knowledge claims, co-construction dominated the learning utterances and did drive some of the argumentation. It can be argued that, without it, there would have been almost no argumentation and no learning gain by any of the students.

Argumentation (and especially counter-argumentation) has been proven to be difficult for students at this age, and thus requires more exposure than one brief field trip before it becomes common practice for them (Jonassen & Kim, in press). The role of inexperience as a major factor in undermining productive inquiry learning is further reinforced by the essential role of prior domain knowledge, inquiry skills, and appropriate epistemological beliefs, while lack of such knowledge or motivation will likely lead to superficial results (Edelson, Gordin & Pea, 1999; Jonassen & Kim, in press; Joolingen et al., 2007).

The mixed results of this study further reinforce the key findings and questions presented by Puntambekar and Hubscher (2005, p.10): “…effective scaffolding needs to
be distributed, integrated, and multiplied so that students have more chances to notice and take advantage of the environment’s and activity’s affordances… How can we design scaffolds that are based on multiple ZPDs found in a classroom?… Are there strategies (or aspects of the domain) that are best scaffolded by a teacher rather than by a tool?…

Each of these issues is complex.”

Furthermore, the findings from this study contribute to the emerging body of studies surrounding the demanding aspects of scaffold inquiry learning (Joolingen et al., 2007), but they also make a timely contribution to emerging discussions about mobile technology-supported collaboration and how productive and profound mobile technology-supported collaboration can be.

This study was limited by the small number of participants and the lack of a control group that would have completed the same tasks without mobile support. However, it has been argued that research designs in authentic contexts inevitably provide principles that can be localised for others to apply to new settings and to produce explanations of innovative practice (Fishman, Marx, Blumenfeld, Krajcik & Soloway, 2004). Therefore, research investigations conducted in authentic contexts are still needed as a first step in order to understand these new opportunities in terms of the learning interaction and collaboration that mobile technologies can provide.

ACKNOWLEDGEMENTS: We wish to thank the Doctoral Programme for Multidisciplinary Research on Learning Environments for supporting this work.

Jari Laru: M.Ed Jari Laru is a doctoral student in the field of learning and educational technology in the Doctoral Programme for Multidisciplinary Research on Learning Environments. He works as a researcher in the Learning and Educational Technology Research Unit (LET). His main research interest deals with developing and analyzing innovative ways of applying the ideas of distributed scaffolding and cognitive tools for learning activities in authentic real world contexts where collaboration is supported with mobile computers.

Sanna Järvelä: PhD Sanna Järvelä is a professor in the field of learning and educational technology and the head of the Learning and Educational Technology Research Unit (LET) (http://www.let.oulu.fi/) in the Department of Educational Sciences, University of Oulu. Her main research interests deal with learning processes in technology-based and virtual learning environments, social and motivational processes in learning, and self-regulated and computer-supported collaborative learning. Ms. Järvelä has published more than 80 scientific papers in international refereed journals, several book chapters, and two edited books.

Roy Clariana: Mr. Roy Clariana is an associate professor in the field of instructional systems and the head of the Education Division in the School of Graduate Professional Studies, Penn State Great Valley University, USA. His main research interests deal with classroom integration of computer-based instruction, the efficient design of instruction, and educational research and statistical analysis. Mr. Clariana earned his bachelor's degree in biology and master's degree in biology education from the University of Central Arkansas in 1975 and 1979, respectively, and a doctorate degree in education from the University of Memphis in 1990. He has done extensive research on e-learning, and published numerous articles on the subject.

References


