Computer-Based Collaborative Knowledge Mapping
to Measure Team Processes and Team Outcomes

CSE Technical Report 502

Harold F. O'Neil, Jr.
CRESST / University of Southern California

Gregory K. W. K. Chung and Howard E. Herl
CRESST / University of California, Los Angeles

March 1999

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 1.3 Technology in Action - Eva L. Baker, Project Director, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education, and under a subcontract to ISX (letter subcontract 95-D-8644-01) funded by Defense Advanced Research Projects Agency and administered by the Navy Research and Development Center.

The findings and opinions expressed in this report do not reflect the position or policies of ISX, Defense Advanced Research Projects Agency, the Department of the Navy, the Department of Defense, the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

COMPUTER-BASED COLLABORATIVE KNOWLEDGE MAPPING TO MEASURE TEAM PROCESSES AND TEAM OUTCOMES[1]

Harold F. O'Neil, Jr.

CRESST/University of Southern California

Gregory K. W. K. Chung and Howard E. Herl

CRESST/University of California, Los Angeles

Abstract

In this study we examined the feasibility and validity of using a computer-based, networked collaborative knowledge mapping system to measure teamwork skills. Student groups were assessed with our system twice in the same academic year, once in the fall and once in the following spring. Our study focused on the nature of the interaction between team members as they jointly constructed a knowledge map. Each student was randomly assigned to a team and communicated (anonymously) with other members by sending pre-defined messages. Teamwork processes were measured by examining message usage. Each message was categorized as belonging to one of six team processes: (a) adaptability, (b) communication, (c) coordination, (d) decision-making, (e) interpersonal, and (f) leadership. Team performance was measured by scoring each team's knowledge map using four expert maps as the criterion. No significant correlation was found between the team processes and team outcomes. This unexpected finding may be due in part to a split-attention effect resulting from the design of the user interface. However, student teams were able to successfully construct knowledge maps, suggesting that our general approach to using networked computers to measure group processes remain viable given existing alternatives.

Introduction

In this study we examined the feasibility and validity of using a computer-based, networked collaborative knowledge mapping system to measure teamwork skills. A particularly novel feature of our work is that we are refining an approach that employs networked computers to capture and measure - in real-time - team processes for individual students and teams. The current work extends our past efforts (e.g., O’Neil, Allred, & Dennis, 1997a; O’Neil, Chung, & Brown, 1997b) in two ways. First, we used a knowledge intensive task that is closer to instructional settings than our past work (see O’Neil, Chung, & Brown, 1997b, for a description of the union-management simulation). Second, we adminstered the same assessment in the fall and subsequent spring of the school year, which provided information on our teamwork assessment system. While some of the results raise questions about the validity of the assessment, our general approach to using networked computers to measure group processes remain viable given existing alternatives.

Teamwork Processes

Our work focuses on assessing team processes used by a group of individuals responsible for jointly constructing a knowledge map. A knowledge map is a node-link-node representation of content, where nodes represent concepts and links represent relationships between connected concepts (Dansereau, 1995; Jonassen, Beissner, & Yacci, 1993). We are interested in the nature of the interaction between team members, and how that interaction influences team performance on a computer-based, collaborative knowledge mapping task. We adopt Baker and Salas’ (1992) definition of a teams as “... two or more people who interact dynamically, interdependently, and adaptively and who share at least some common goals or purpose,” and where the team members have built-in dependencies and where each member has well-defined roles and positions.

We have drawn on the work of Morgan, Salas, and Glickman (1993); Salas, Dickinson, Converse, and Tannenbaum (1992); Burke, Volpe, Cannon-Bowers, and Salas et al. (1992); and others (O’Neil, Baker, & Kazlauskas, 1992; Webb, 1993; Webb & Palincsar, 1996) to aid our development of teamwork process measures. Morgan et al. provide insight into the nature of teams. In their model of team development, Morgan et al. (1993) postulate two tracks of team processes, a taskwork track and a teamwork track. The taskwork track accounts for specific activities unique to the task. Taskwork team skills influence how well a team performs on a particular task (e.g., how well a team of students are able to jointly construct a knowledge map). Taskwork skills are domain-dependent, task-related activities. The teamwork track or teamwork skills influence how effective an individual member will be as part of a team and are domain-independent.

Teamwork skills encompass skills such as adaptability, coordination, cooperation, and communication. Effective teams develop competence along both tracks. Members of effective teams possess basic skills required for the task and know how to coordinate their activities, communicate with each other, and respond effectively to changing conditions.

To measure teamwork processes, we used the taxonomy of teamwork process from prior work (O'Neil et al., 1997b). The taxonomy has six teamwork processes: (a) adaptability, recognizing problems and responding appropriately; (b) communication, the exchange of clear and accurate information; (c) coordination, organizing team activities to complete a task on time; (d) decision-making, using available information to make decisions; (e) interpersonal, interacting cooperatively with other team members; and (f) leadership, providing structure and direction for the team. Our prior research showed statistically significant positive relationships between decision making and team performance on a negotiation task (final contract offer and negotiation style), and a significant negative correlation between interpersonal processes and whether an agreement was reached (O'Neil et al., 1997b).

Team Processes in a Networked Knowledge Mapping Task

In this section we discuss each teamwork process in the context of a group knowledge mapping activity, where the activity takes place over networked computers. To the best of our knowledge, this has never been done before; we know of no theoretical framework that has been tested using a group knowledge mapping task. Thus, our expectations about what processes team members invoke are tentative at best.

Our group and task configuration is a weak implementation of (Baker & Salas, 1992) definition of a team. In particular, member interdependency is low (all members could complete the task independently), and the function of each member is non-specialized (all members eventually perform the same functions). Another difference is that our task is computer-based. Group members communicate anonymously by sending pre-defined messages to each other. Given these differences, we next examine each teamwork process with respect to our networked knowledge mapping task.

Our expectation of low member interdependency was based on our observation of how easy participants are able to grasp the idea of knowledge mapping and how easy they find using our computer-based system. We have found in previous work that participants as young as elementary school children can successfully use our system to construct knowledge maps. This situation also suggests that anyone can carry out the task, which means that there are few specialized skills necessary. Given our typical team scenario of one leader and two advisory members (see O'Neil et al., 1997b for a detailed description), we expected each person on a team to be able to perform either role equally well. This situation is different from more traditional team environments where each member has a well defined role (Morgan, 1993).

Finally, we expect mapping performance to be essentially knowledge driven. To the extent that teamwork processes emerge, we expect these processes in general to be mediated by the content expertise of team members. Teams with members low in content knowledge are expected to focus primarily on surface features of the task and far less on the substantive part (Chi, Glaser, & Farr, 1988). In contrast, in most team environments members have the requisite knowledge to carry out their function (Morgan, 1993). In our knowledge mapping task, we investigated how this relaxation of requisite knowledge would impact teamwork processes and performance.

Thus, given these characteristics of our group knowledge mapping - easy to carry out task, easy to use computer system, but content novices - we expect low performance and we do not in general expect to see strong relationships between many of the team processes and knowledge map performance. The one exception is decision making. Given that knowledge maps are knowledge dependent, we expect teams that engage in discussion about the content to make reasoned, informed decisions about concepts and their relations, and thus perform better on the task. Our tentative prediction of the role of team processes on performance follows.

Adaptability

In general, a team's adaptive capability affects the nature of the team's response to a given situation. Are team members able to detect problems as they arise and, once detected, respond effectively to resolve the problems? In the context of a group knowledge mapping task, adaptable teams should detect problems with their knowledge map at a deep (semantic) level and at a surface level. An effective team would be able to detect gross inaccuracies as well as the strength and significance of a relationship. This of course assumes adequate content knowledge. At a surface level, adaptable teams should recognize that the given set of concepts and links should be included in their map, and if not, to include them.

In our group knowledge mapping activity with novices, we did not expect any overall effect of adaptability on performance. Adaptable teams are able to detect and respond to problems as they arise. In our task there is no obvious indicators of whether a problem exists - there is no feedback available to students about their map. Participants have no way of knowing (other than their own background knowledge) the accuracy of their map. Given our presumption of familiarity with the content but not expertise, we expected little in terms of detection of content-based inaccuracies. However, we did expect to find some relationship between adaptability and the surface features of the task (e.g., the number of concepts and links). More adaptive teams should be more likely to detect whether their knowledge map has included all the terms and links.

Coordination

In general, a team's coordination capability affects how that team organizes its resources, activities, and responses. Teams with high coordination will carry out a task that is integrated, synchronized, and completed on time. Often well-coordinated teams will employ a divide-and-conquer strategy to complete a task (McIntyre & Salas, 1995). For a knowledge-based task like knowledge mapping, a reasonable strategy is to rely on members' domain expertise to determine the relationships between concepts. Another characteristic of well coordinated teams is that they are aware of time constraints of the task and respond appropriately.

Given that knowledge mapping is an open-ended task without an obvious terminal point, we expect well-coordinated teams to be sensitive to time constraints. We provide the team a clock on the computer screen that indicates the remaining time. Our task is timed and we have implemented leader-passing via a round-robin scheme. Every member gets control of the knowledge map twice during the task. Because our presumption is that participants are content novices, we do not expect them to develop any sort of division of responsibility along content expertise. Moreover, our task is a one-shot event and fairly short, so we do not expect to observe any team maturation effects (Morgan, 1993).

Decision Making

A team's decision-making capability affects its ability to capitalize on available information. Effective teams use information about their current situation to help them evaluate the utility of potential courses of action. In general, effective teams employ decision making that takes into account all available information. In terms of knowledge mapping, we expect decision-making processes to emerge when the nature of a relationship between two concepts is ambiguous, or when the relationship is unknown. Members need to assess the quality of the relationship based on their knowledge of the domain. As with adaptability, we expect content knowledge to play a significant role in how decision-making processes play out. Team members with relevant prior knowledge may be more likely to engage in substantive discussions about the relationships. Members with less background in the domain may not know enough to be able to engage in any substantive discussion about the concepts or relationships.

In our group knowledge mapping task, our assumption is that team members have sparse content knowledge. Thus, we expect the team discussion to be dominated less by substantive discussions and more by appeals for ideas on what to do procedurally. By substantive we mean considering alternative perspectives, assessing the quality of different relationships, assessing the relative merits of different relationships, and in general, engaging in reasoning about concepts and their relationships to each other. By procedural, we mean the mechanics of constructing a knowledge map.

Interpersonal

In general, a team's interpersonal capability affects how well team members work cooperatively. Do team members provide positive comments to each other? Do team members encourage participation? In the context of a group knowledge mapping task, we expect groups with a high interpersonal dimension to encourage participation and to engage in compliments and other cooperative discussion. We do not expect interpersonal processes to be dependent on content knowledge.

Leadership

In general, a team's leadership capability affects its ability to provide direction for the team. Does the team engage in planning and organization that reflect appropriate priorities? In the context of a group knowledge mapping task, teams that employ good leadership processes should provide appropriate direction to the team. We believe this will be reflected in a team that is progressing steadily toward completing the knowledge map. An effective team would have members' behaviors aligned toward a common goal. We expect that leadership processes in a knowledge mapping task to manifest itself primarily as directives to carry out knowledge map operations (e.g., adding concepts or links).

While getting the team to do something is a necessary condition for good performance on a knowledge map, directives to do something is not a sufficient condition for effective team performance. Good leadership results in part from requisite knowledge about the domain. Members who know the content can be direct and accurate about the content, which results in an accurate map. Members with little content knowledge can be just as direct, but it is unlikely the quality of their map will be similar. Thus, for our group knowledge map task, we expect leadership to be necessary but its effect on team performance will be neutral for knowledge-based tasks.

Measuring Team Processes Within a Networked Environment

Our work focuses on assessing team processes that emerge during a group knowledge mapping task. We are interested in the nature of the interaction between team members and how that interaction impacts team performance. Yet a critical measurement issue remains unresolved: How do we assess teamwork processes such that the measurement technique is reliable, valid and timely?

Existing approaches to measuring teamwork processes rely almost exclusively on observational methods (Baker & Salas, 1992). For example, behavioral checklists (e.g., Oser et al., 1989), videotaped and audiotaped observation (e.g., Brannick et al., 1993), and analysis of think-aloud protocols are the most common techniques to measure teamwork processes. These methods are labor intensive and time-consuming. Observations must be transcribed, coded, and analyzed post hoc. Such techniques offer no opportunity for rapid analysis and reporting of team performance. From an assessment perspective, these methods are unappealing because of the lag between test administration and reporting of test results. Further, these methods are neither practical nor cost-effective in large-scale test settings.

An alternative to observational methods is to provide participants with a set of messages to choose from (i.e., predefined messages). In our current version of the system we are providing predefined messages (O'Neil et al., 1997b) for three reasons. First, providing predefined messages eliminates any off-task discussions, which makes the assessment more efficient. In our pilot study (see next section), which allowed participants to type messages, off-task messages accounted for 26 percent of the total message usage. With predefined messages, participants are constrained to selecting complete messages, or message stems where participants type in text to complete the message stem (e.g., for the decision making message stem "If we [...], then [...]," participants type in the information for each "[...]").

Second, the use of predefined messages allows us to tailor the form of a message such that its intended use by participants is less ambiguous. For example, the message "vegetation leads to food resources" is potentially ambiguous - does the participant mean, "Add vegetation leads to food resources," or is the message a response to a question? With predefined messages, message stems could be created to disambiguate this situation (e.g., the message stem "Let's add [C]-[L]-[C]," where "[C]" and "[L]" are user selected concepts and links).

Third, the use of predefined messages provides the capability to measure team processes in real time. Messages are developed a priori based on a pre-defined taxonomy of teamwork processes; each message maps on to one of these processes (adaptability, coordination, decision making, interpersonal, and leadership). This taxonomy of teamwork is domain independent and independent of scenarios. By tracking the messages selected and sent (and hence, by definition, the teamwork process category) we get an index of the kinds teamwork processes members are using. We assume that each message in a category is as important as any other message, and thus all messages are equally weighed. This technique provides us with a real-time teamwork assessment system with the potential to administer, score and interpret in real time. Our prior work suggests that this is a feasible and promising approach to assessing teamwork skills (O'Neil et al., 1997b). The work reported here is an extension of our approach to a different domain and a different task.

Pilot Study

Two pilot studies were conducted to assess the feasibility of our approach. The first pilot study represented our initial attempt at assessing the functionality of the computer system and what kinds of messages participants use to construct a knowledge map in a group. We were interested in feedback from users regarding usage of the system, the messages, and task performance. The second pilot study reflected several major revisions to the system based on the results of the first pilot study, and verified that our technique was feasible (i.e., students could jointly construct knowledge maps using our system). Only the first pilot study is described here.

Participants

Participants were 30 ninth-grade male and female students (10 teams) randomly drawn from five ninth-grade science classes. The classes were located on an American army bases in Germany, and was part of the Department of Defense educational system. All students spoke English as their first language.

Team Knowledge Mapping Task

A knowledge map is a node-link-node representation of content, where nodes represent concepts and links represent relationships between connected concepts (Dansereau, 1995; Jonassen et al., 1993). In our pilot study we created a task where participants were encouraged to collaborate in three-person teams to jointly construct a knowledge map on environmental science. One reason we used environmental science as the topic was that it was part of the curriculum; thus, students were presumed to have some familiarity with the subject. (The pilot study was conducted near the end of the school year.) We provided the team with 15 terms (*atmosphere*, *bacteria*, *climate*, *CO2*, *decomposition*, *evaporation*, *food resources*, *oxygen*, *precipitation*, *respiration*, *sun light*, *vegetation*, *photosynthesis*, *waste*, *water*) and nine links (*contributes to*, *causes*, *leads to*, *part of*, *result of*, *similar to*, *produced by*, *influences*, *type of*, *prior to*). Teams were allowed 36 minutes to complete the map. The knowledge map was scored by the computer and in real-time using a scoring algorithm that compared students' maps with an expert's map (Herl, Baker, & Niemi, 1996).

We configured participants into three-person teams where each person had their own computer task and the computers were networked together. The scenario we used was one where there was one leader and two members. Only the leader could manipulate the knowledge map (e.g., adding concepts and links). Non-leaders could only send messages. To allow all members to participate in manipulating the map, leadership passed among team members such that each member was a leader once. Leadership passed from one member to another every 12 minutes (i.e., one-third of the task time).

Networked Knowledge Mapper

The system was developed on the Macintosh® with HyperCard 2.3®. We used the built-in networking capabilities of the operating system (System 7.5®) and HyperCard® to implement a rudimentary peer-to-peer system. Every member's screen was updated as changes occurred (e.g., someone sending a message, or the leader making changes to the knowledge map); thus, all computers were synchronized with each other. Communicating between team members occurred via typed messages. To send a message, members typed what they wanted to say in a special message box and then sent the message. The message was dispatched to other members' computers and appeared in the order they were sent. To make changes to the knowledge map, the leader could add, delete, or move knowledge map components around the screen. Each time a map event occurred, the event was dispatched to other members' computers and their maps automatically updated. All typed messages and map events were logged by the computer.

Results

All groups participated in the knowledge mapping task with sustained interest and effort. There was unanimous agreement from all participants (and their teachers) that the collaborative mapping task was interesting, engaging, and fun. All groups, with only a few minutes of instructions, were able to immediately use the system and engage in the task. All groups were able to construct a map within the allotted time.

Team Processes

Seven hundred ninety-eight messages were sent by all groups. All messages were rated by one experimenter. Messages were rated first as on-task or off-task.

Five hundred fifty-seven messages were rated as on-task (74 percent) and 211 (26 percent) messages were rated as off-task. On-task messages was defined as messages that were directly related to constructing a knowledge map (e.g., “Add evaporation and make a link to water”). Off-task messages was defined as messages that were unrelated to the construction of the knowledge mapping (e.g., “What was the last good movie you guys saw?”). On-task messages were then categorized into one of the teamwork process categories. Total message usage was: adaptability, 151 messages (26 percent); coordination, 12 messages (2 percent); decision making, 80 messages (14 percent); interpersonal, 15 messages (3 percent); leadership, 223 messages (38 percent); and communication, 87 messages (15 percent). Two additional categories were used to capture messages that reflected errors with using the system (e.g., a message reflecting a typing error, three messages, 1 percent) and a second category to capture uncategorizable messages (16 messages, 3 percent).

Team Knowledge Mapping Performance

Team performance was measured with knowledge mapping scores. The knowledge maps were scored by comparing the group map with an expert map in terms of the content and structure of the map (Herl et al., 1996). Herl et al. reported that this knowledge mapping scoring approach yielded reliable scores with strong positive correlations with other measures of content knowledge (e.g., essay writing and prior knowledge short answer questions). In this study, the mean team content score was 6.4 (*SD*=2.72, range=2 to 11) out of a possible 38 (16 percent). The mean team structural score was .29 (*SD*=.08, range=.14 to .37) out of a possible 1.0 (37 percent). The scores indicate little content knowledge of these students compared to an expert.

Individual message counts were also scaled by the total number of messages sent by the group. Non-parametric (Spearman) correlations were then examined between the six teamwork processes and the two outcome measures. For the content score, a significant negative correlation was found for communication (r_s(10)=-.74, p=.007). Communication was defined as the total number of messages sent. The more teams used communication messages, the lower they scored on the content measure. For the structural score, a significant positive correlation was found for decision making (r_s (10)=.59, p=.02), and a significant negative correlation for adaptability (r_s (10)=-.65, p=.02). The more teams used decision making messages (e.g., “Leader - is atmosphere and oxygen really related to climate?”), the higher they scored on the structural measure. Conversely, groups that used more adaptability messages (e.g., “What should we link next?”) tended to score significantly lower on the structural measure.

These results, while limited in scope, are consistent with our previous findings (O’Neil et al., 1997b). In particular, decision making seems to be a key factor in how a group performs. Groups that engaged in higher levels of decision making tended to perform higher on the task. Also consistent with our prior study was that no other significant correlation was found between other specific group processes and team performance.

MAIN STUDY

Method

Participants and Design

A pre-/post-test design was used to assess the reliability of our teamwork assessment. We were interested in how our assessment approach performed on two different occasions using participants randomly drawn from the same population.

Participants were drawn from the same classes from the same schools. The schools were located on two American army bases in Germany, and were part of the Department of Defense educational system. All students spoke English as their first language. For the pre-test, 23 groups (69 participants) were drawn from six middle and high school classes. Technical problems resulted in some computers crashing; thus, usable data for only 15 groups (45 participants) were available. For the post-test, 14 groups (42 participants) were drawn from the same classes seven months later. Students were randomly assigned to groups in both the pre- and post-test sessions. No attempt was made to maintain intact groups across the pre- and post-test data collections.

Networked Collaborative Knowledge Mapping System

Table 1 lists the specifications for our networked knowledge mapping system. Our system was developed using HyperCard® running on the Macintosh®. The software was designed to be simple and easy to use. Participants sent messages by clicking on numbered buttons. Concepts were added to the knowledge map via menu selections, and links were created by connecting two

Table 1

Domain Specifications Embedded in the Software

General domain specification	This software
Scenario	Create a knowledge map on environmental science by exchanging messages in a collaborative context
Participants	Student team (three members)
Knowledge map terms	Predefined. Eighteen important ideas identified by content experts: atmosphere, bacteria, carbon dioxide, climate, consumer, decomposition, evaporation, food chain, greenhouse gases, nutrients, oceans, oxygen, photosynthesis, producer, respiration, sunlight, waste, and water cycle
Knowledge map links	Predefined. Seven important relationships identified by content experts: causes, influences, part of, produces, requires, used for, and uses
Type of learning	Content understanding
Outcome measures	Semantic content score, organizational structure score, number of terms used, number of links used
Teamwork processes	Adaptability, coordination, decision making, interpersonal, leadership, communication

concepts and then selecting the desired link from a pop-up menu. Our pilot studies and in-house usability testing showed that participants of various ages (fifth graders to graduate students) could be trained to use the system within a few minutes.

User Interface

Figure 1 shows the user interface to the system. The display was partitioned into three major sections. The top fifth was reserved for knowledge mapping. The lower left of the screen displayed the messages. All messages sent were listed in the order sent by members. Thirty-five numbered buttons were provided to members. Each participant was given a paper handout that listed all messages, and the messages were numbered to correspond to the buttons on the computer screen. To send a message, participants clicked on a button and the corresponding message was sent to everyone's computers. Other information shown on the lower part of the screen was the remaining time as a leader or non-leader, and the remaining time for the entire task.

The majority of messages could be sent without typing. Participants could send a message by clicking on a numbered button. Some messages required user input. These messages were handled by the use of dialogs that required simple point-and-clicking. For example, Figure 2 shows how message number 1 ("Let's

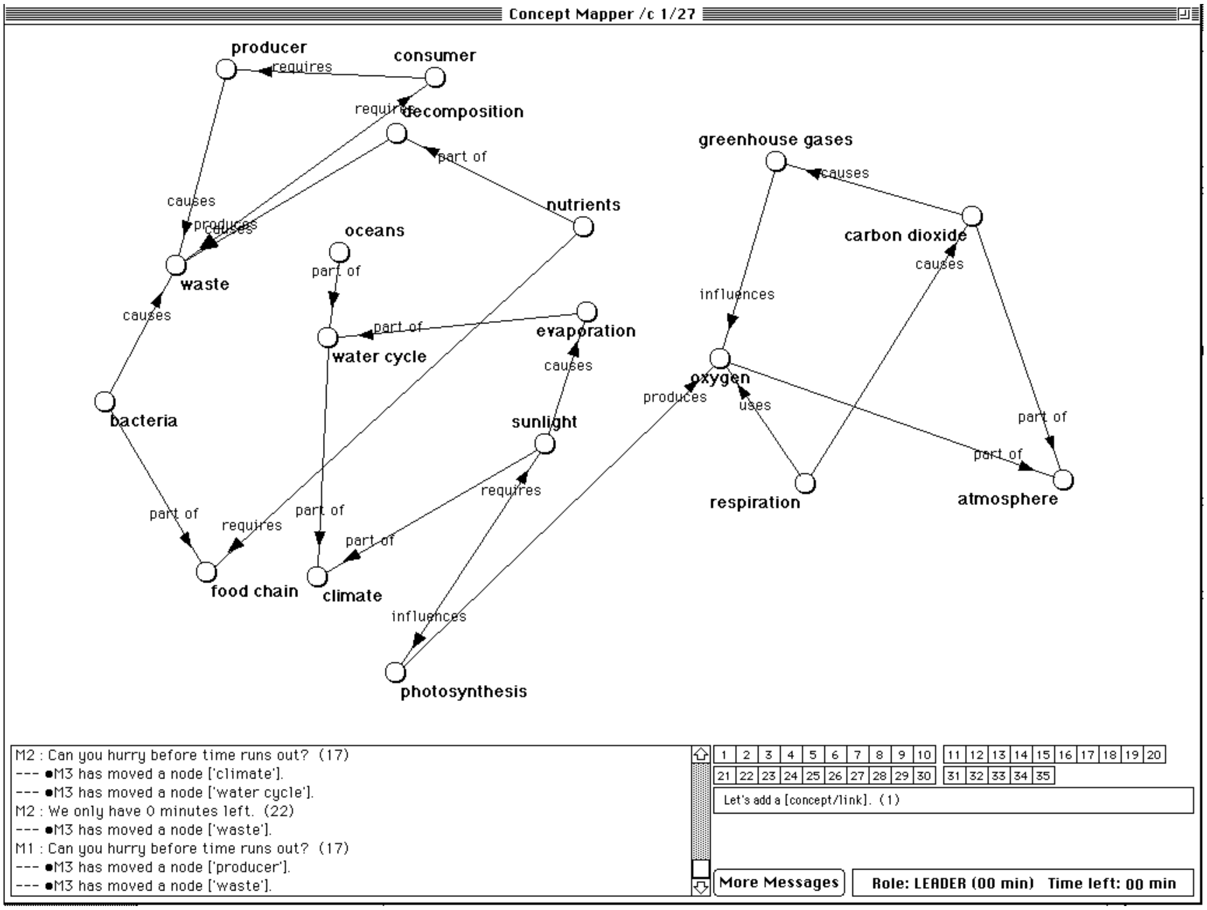

Figure 1. User interface for the system. The knowledge map is from an actual student group.

Figure 2. Message requiring user input: selecting values.

add a [concept/link].") was handled. In this case, the participant simply selects the desired option (concept or link) and the message was sent reflecting the choice. A second type of message required users to type text, as shown in Figure 3. In this case, the intent was for participants to complete the message stem.

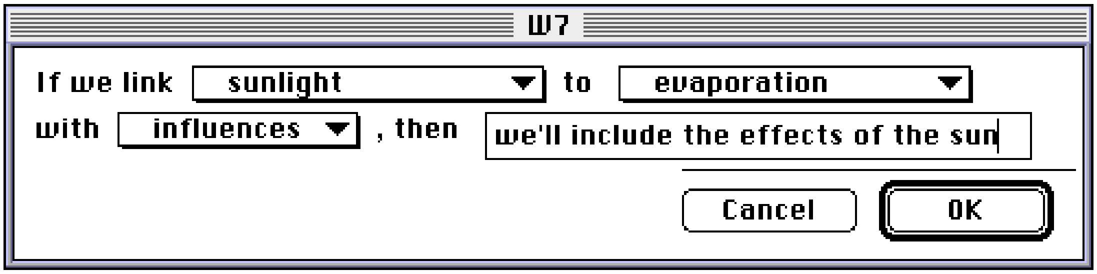

Figure 3. Message requiring user typed text.

Message Handout

Participants received a handout that listed the 37 messages, grouped by common functions (see Appendix B). The list of concepts and links were also provided.

Predefined Messages

All of the pre-defined messages used in the knowledge mapping task were selected from a larger pool of potential messages. We used four sources to develop our message set. Our first source was the message set from our previous study (O'Neil et al., 1997b). We examined each message to see if it could be used intact, or if the content or form could be converted for use in knowledge mapping. Our second source was based on the feedback from participants in our previous study (O'Neil et al., 1997b). In that study we asked participants what messages they would have liked to have. We reviewed participants' responses to this question to help develop additional messages. A third source of messages was the authors. We generated messages that we thought would represent a particular teamwork category in the context of a knowledge mapping task. The last source of messages was from the pilot study. In the pilot study participants were allowed to type messages to each other. We developed a set of pre-defined messages from participants' typed messages that we believed reflected the most commonly used messages.

After developing our preliminary set of messages, two independent raters sorted all into five teamwork process categories: adaptability, coordination, decision making, interpersonal, and leadership. Disagreements were resolved by discussion. We then conducted a series of in-house usability studies. We were interested in seeing if participants could jointly build a knowledge map using our set of pre-defined messages. We refined our message set based on feedback from participants and actual message usage. Any new messages generated were sorted into the teamwork process categories by two independent raters.

The final message set consisted of 37 messages (see Appendix A). Thirty-four messages belonged to the five teamwork process categories discussed earlier (adaptability, seven messages; coordination, seven messages; decision making, six messages; interpersonal, seven messages; and leadership, seven messages). Two messages (“Yes” and “No”) were part of the communication category. Another message was provided to signal that a previously sent message was mistakenly sent by the participant. This message was excluded from the message categories. Table 2 gives a brief definition and examples of one message from each category.

Table 2
Example of Messages Belong to Each Team Process Category

Team process	Definition and example
Adaptability	Recognizing problems and responding appropriately. What do you think, M1? (M1 refers to Member 1 of the team.)
Coordination	Organizing team activities to complete a task on time. We only have 5 minutes left.
Decision making	Using available information make decisions. What if we add “carbon dioxide”?
Interpersonal	Interacting cooperatively with other team members. I need to hear from all of you.
Leadership	Providing direction for the team. Let’s link “carbon dioxide” to “producer.”
Communication*	The clear and accurate exchange of information.

* Our definition of communication was modified from previous work (O’Neil et al., 1997b). Previously, we operationalized communication as the sum of all messages less the number of error messages. However, based on the feedback from our previous study and our pilot work for the current study, we added two messages to this category (“Yes” and “No”). Thus, communication was computed as the number of “Yes” and “No” messages sent, plus the total of all other messages minus the number of error message.

Measures

The measures used in this study focused on team level outcomes within a group knowledge mapping context. Performance was measured by how well the team’s knowledge map compared to experts’ maps. Process measures included team processes (as measured by message usage) and knowledge mapping activity (as measured by the concept mapping events that occurred). Each measure is described below.

Team Outcome Measures

Our team outcome measures were based computed by comparing a teams' knowledge map against a set of experts (Herl et al., 1996). A semantic content score, a structural score, and the number of concepts and links of the group map were used as outcome measures.

To derive semantic content scores, the following technique was used. First, links in both the group and expert maps were categorized into more abstract categories. The effect of the categorizing specific links into more abstract categories is to remove subtle differences between links (e.g., "causes" and "influenced" were classified as belonging to the "causal" category). Next, every proposition in the group's map (i.e., concept-link-concept) was compared against all propositions in an expert's map. A point was awarded for each match. This comparison was done against four expert maps and the group's knowledge mapping content score was averaged across the four experts.

The second measure was an organizational structure score, which reflected the similarity in network characteristics between a group's map and the experts' maps. Herl et al. (1996) gives a complete discussion of all scoring algorithms.

The two other outcome measures were simple counts of the number of concepts and the number of links used in the group knowledge map.

Individual- and Team-level Teamwork Process Measures

The individual and team-level teamwork process measures were designed to capture the degree to which individuals and their team engaged in each of the team processes (i.e., adaptability, cooperation, decision making, interpersonal, leadership, and overall communication). Individual team process measures were computed by counting the number of messages sent in each team process category. For example, if a team member sent 10 messages in the adaptability category, that individual's adaptability score was 10. This method was used for all categories except communication. Communication was measured by taking the total number of messages sent by an individual less the number of "I sent the wrong message" messages. Team-level teamwork process measures were computed by summing across individual categories. Thus, if all three team members each sent 10 messages in the adaptability category, then the team-level adaptability performance was 30.

Because the distribution of message usage highly skewed for most messages, an individual's frequency count on each message was scaled relative to the total usage across all individuals. This scaling is discussed in greater detail in the Results section.

Knowledge Mapping Activity Process Measures

Group knowledge mapping activity was measured by the number of concept and link events generated by the group. A concept related event occurred when a concept was added, deleted, and moved by the group. A link related event occurred when a link was created, deleted, or revised. All concept and link related events were logged by the computer.

Procedure

Each participant was randomly assigned to a team and a role (leader or member). There were no computer lab facilities at the school site so all data collection took place in classrooms that had ongoing instruction. Participants were arrayed in the back of the classroom to minimize the interaction with other students. The experimenters first introduced themselves and the study, then trained the students on the computer system and task, and then started the group knowledge mapping task. The training session took 13 minutes and the actual task 36 minutes. Leadership rotated every six minutes giving each participant the opportunity to control the knowledge map twice.

Training Task

Participants were trained on the system with a different set of concepts and a reduced set of links. The training system had three concepts (birds, eagle, and wings) and three links (part of, type of, and uses). Participants were shown how to carry out critical functions such as sending messages, adding concepts, and creating links between concepts. The experimenter also explained the different roles each member would play, and how the leadership would pass from each member such that each member would get to be a leader twice.

Team Scenario

Participants were informed that they were part of a three-person team comprised of one leader and two members. Only the leader could make changes to the map. The two members were responsible for advising the leader.

Data Collection Problems

During the pre-test data collection period 23 groups were run. The computers used were a combination of experimenter supplied computers and classroom computers. During the task many of the classroom computers crashed resulting in the loss of team process data. Thus, of the 23 groups only 15 groups produced data that was suitable for group analyses. Note that team performance data was recoverable and usable, resulting in 22 groups with group performance data. However, these performance data are not included in any of the analyses. The post-test data collection fared better, with 14 groups being collected without any problems.

Results

Overview of Results

Three issues were examined in this study: feasibility of a teamwork assessment system using networked computers; reliability of teamwork processes over two occasions; and reliability of teamwork performance measures over two occasions.

Measuring team processes and performance using a group knowledge mapping task over networked computers was demonstrated as feasible. The cost and effort put into the development of our prototype system was not prohibitive and technically possible. Clearly, all teams could use our system and could complete the task in the allotted time.

A second issue examined the reliability of the teamwork process and teamwork performance measures. Unfortunately, our sample sizes was too small to conduct a formal test-retest reliability analysis. Groups in the pre-test generally sent more messages on average, although the difference between pre-test and post-test groups were not significant. We did not find the expected relationships between teamwork process measures and teamwork performance for either the pre-test or post-test. Where significant relationships existed, the direction of the relationship was opposite of what we expected. In general, the relationship between teamwork processes and teamwork outcome measures was negative, suggesting that the more messages a team sent the poorer they performed on the knowledge mapping task. Follow-up analyses were conducted to explore possible explanations for these curious findings. Results of these analyses are consistent with a split-attention effect.

A third issue examined is the stability of the team performance measure. In this study we found that the teams increased their performance on all the knowledge mapping outcome measures from the pre-test to the post-test, although this increase was non-significant. All groups were able to successfully construct a knowledge map, although the range of performance on the content-related measure (i.e., the semantic content score) varied considerably from a low of 0.50 to a high of 10.30. In percentage terms the mean team performance was 19.3 percent (pre) and 26.6 percent (post) of the expert's mean performance.

Individual-level Measures

Analysis of the data showed that each of the 37 messages provided in the knowledge mapping task was used at some point. The observed frequency for each message is provided in Table 3. As Table 3 shows, although all messages were used, individual message usage varied. For example, in the pretest, message usage ranged from 1 (message #9: "Did we use all the concepts?") to 206 (message #34: "No"). The frequency distributions were highly skewed positively, with most items having a mode of zero. Thus, prior to combining the items to generate team process scales, the message counts were transformed from raw score frequency of usage counts for each participant to percent of overall item usage for that participant relative to all participants. That is, if Participant A had a frequency count of 6 on item X that had a total usage of 50, Participant A's score on item X was transformed from 6 to .12 (6/50).

Table 4 gives the frequency counts for each team process category. Interestingly, decision making was used far less than the other teamwork processes. The frequency counts of the other categories show fairly uniform usage. These results contrast with the data from the pilot study, which showed varied overall usage by category (adaptability, 26 percent; coordination, 2 percent; decision making, 14 percent; interpersonal, 3 percent; leadership, 38 percent).

For each participant, individual scores were calculated for the adaptability, coordination, decision making, interpersonal, leadership and teamwork processes. The scores were computed by counting the messages from each category. In addition, a total communication score was computed by summing each of the five teamwork process scores for each individual less the total

Table 3

Frequency Count of Messages by Occasion

	Occasion			Occasion			Occasion	
Message number	Pre	Post	Message number	Pre	Post	Message number	Pre	Post
1	60	32	12	47	30	25	16	13
2	41	50	13	13	8	26	27	25
3	47	47	14	9	3	27	23	28
4(adapt)[a]	31	20	15	4	3	28	50	53
4(dec)[a]	23	15	16	34	15	29	47	48
4(lead)[a]	14	17	17	65	47	30	54	23
5	30	30	18	24	15	31	108	70
6	91	90	19	47	31	32	112	80
7	18	13	20	31	13	33	167	109
8	46	27	21	36	51	34	206	200
9	1	10	22	82	37	35	46	31
10	2	1	23	58	30			
11	13	9	24	30	49	Off-task[b]	33	63

[a]Message 4 was presented to participants as a combination of three message types: adaptability, decision making, and leadership. The category of the message depended on the specific message selected.

[b]Off-task messages were possible with four messages. In this case, students were required to complete the message (e.g., "If we don't link [C] to [C] with [L], then [...]," where students replaced "[...]" with their own typed input.

Table 4

Frequency Count of Teamwork Processes, Individual-level

Team process	Occasion	Total (pct[a])	Mean	SD	Min.	Max.
Adaptability	Pre[b]	314 (17%)	6.98	4.59	1	18
(Msg #: 4(adapt), 5, 7, 9, 29, 30, 31)	Post[c]	214 (16%)	5.10	3.52	0	15
Coordination	Pre[b]	298 (17%)	6.62	5.45	0	22
(Msg #: 13, 17, 18, 19, 20, 21, 22)	Post[c]	202 (15%)	4.81	6.40	0	28
Decision making	Pre[b]	73 (4%)	1.62	1.79	0	6
(Msg #: 4 (dec), 10, 11, 14, 15, 16)	Post[c]	46 (4%)	1.10	1.86	0	8
Interpersonal	Pre[b]	316 (18%)	7.02	5.09	0	23
(Msg #: 23, 24, 25, 26, 27, 28, 32)	Post[c]	278 (21%)	6.62	6.02	0	31
Leadership	Pre[b]	346 (19%)	7.69	5.27	0	18
(Msg #: 1, 2, 3, 4(lead), 6, 8, 12)	Post[c]	293 (22%)	6.98	4.81	0	19
"Yes" or "No"	Pre[b]	373 (21%)	8.29	8.34	0	48
(Msg #: 33, 34)	Post[c]	309 (24%)	7.36	10.56	0	52
Error	Pre[b]	46 (3%)	1.02	1.48	0	7
(Msg #: 35)	Post[c]	31 (2%)	.74	.83	0	3
Off-task messages[d]	Pre[b]	33 (2%)	.73	2.33	0	15
	Post[c]	63 (5%)	1.50	2.62	0	13
Total	Pre[b]	1799				
	Post[c]	1312				

[a]Percentages do not total to 100% due to rounding errors; [b]n=45; [c]n=42; [d]Off-task messages were possible with four messages. In this case, students were required to complete the message (e.g., "If we don't link [C] to [C] with [L], then [...]," where students replaced "[...]" with their own typed input.

number of error messages for that individual. An individual's scale score was not adjusted for error messages. Means and standard deviations for each scale are presented in Table 5. As shown in Table 5, participants on average sent approximately 16 percent of the messages in each category. Note that our metric does not allow comparison between categories or across occasions.

Table 5

Descriptive Statistics for Pre- and Post Individual-level Teamwork Process Scales

Teamwork process	Occasion	Mean	SD	Min.	Max.
Adaptability	Pre[a]	0.16	0.12	0.01	0.51
	Post[b]	0.17	0.12	0.00	0.52
Coordination	Pre[a]	0.16	0.14	0.00	0.59
	Post[b]	0.17	0.21	0.00	0.98
Decision making[c]	Pre[a]	0.13	0.23	0.00	1.04
	Post[b]	0.14	0.24	0.00	1.00
Interpersonal	Pre[a]	0.16	0.13	0.00	0.55
	Post[b]	0.17	0.19	0.00	0.99
Leadership	Pre[a]	0.16	0.12	0.00	0.42
	Post[b]	0.17	0.12	0.00	0.48
Communication	Pre[a]	0.80	0.41	0.15	2.23
	Post[b]	0.86	0.53	0.03	2.39

[a]n=45; [b]n=42; [c]The decision making category contains six messages. Adaptability, coordination, interpersonal, and leadership contain seven messages. Communication contains two messages, plus the composite of all other categories.

Team-level Process Measures

Table 6 presents frequency team level counts for each team process category. As with individuals, decision making was used far less than the other teamwork processes. The frequency counts of the other categories show similar usage by process and across occasions.

Table 6

Frequency Count of Teamwork Processes, Team-level

Team process	Occasion	Mean	SD	Min.	Max.
Adaptability (Msg #: 4(adapt), 5, 7, 9, 29, 30, 31)	Pre[a] Post[b]	20.93 15.29	10.44 7.48	7 5	41 29
Coordination (Msg #: 13, 17, 18, 19, 20, 21, 22)	Pre[a] Post[b]	19.87 14.43	11.83 10.82	1 0	43 36
Decision making (Msg #: 4 (dec), 10, 11, 14, 15, 16)	Pre[a] Post[b]	4.87 3.29	3.18 2.81	0 0	13 9
Interpersonal (Msg #: 23, 24, 25, 26, 27, 28, 32)	Pre[a] Post[b]	21.07 19.86	10.13 12.15	3 6	41 46
Leadership (Msg #: 1, 2, 3, 4(lead), 6, 8, 12)	Pre[a] Post[b]	23.07 20.93	11.20 11.19	6 4	43 47
"Yes" or "No" (Msg #: 33, 34)	Pre[a] Post[b]	24.87 22.07	14.18 21.27	7 2	51 74
Error (Msg #: 35)	Pre[a] Post[b]	3.07 2.21	2.76 1.37	0 0	10 4
Off-task messages[c]	Pre[a] Post[b]	2.20 4.50	4.41 6.32	0 0	17 23

[a]n=15; [b]n=14; [c]Off-task messages were possible with four messages. In this case, students were required to complete the message (e.g., "If we don't link [C] to [C] with [L], then [...]," where students replaced "[...]" with their own typed input.

Each of the individual-level teamwork process measures was used to calculate comparable team-level scores. The individual scores for adaptability, coordination, decision making, interpersonal, leadership, and error messages were summed among the three members of each team to generate a team score for each process. Descriptive statistics are presented in Table 7, and intercorrelations for each of the team-level teamwork process measures are presented in Tables 8 and 9. Communication is a composite (simple sum) of the other five team process measures, which are restricted to a 0-1 range; thus the range for communication is from 1.19 to 4.49. Unlike the other team process measures, communication is a composite rather than a proportion.

Table 7

Pre- and Post-Test Descriptive Statistics, Team-level

Teamwork process	Occasion	Mean	SD	Min.	Max.
Adaptability	Pre[a]	.47	.25	.15	1.07
	Post[b]	.50	.26	.13	1.09
Coordination	Pre[a]	.47	.30	.02	1.06
	Post[b]	.50	.39	.00	1.35
Decision making	Pre[a]	.40	.30	.00	1.30
	Post[b]	.43	.37	.00	1.40
Interpersonal	Pre[a]	.47	.25	.04	1.02
	Post[b]	.50	.35	.16	1.30
Leadership	Pre[a]	.47	.22	.16	.83
	Post[b]	.50	.24	.10	.95
Communication	Pre[a]	2.33	0.60	1.19	4.49
	Post[b]	2.57	1.10	1.24	5.24

[a]n=15; [b]n=14

Table 8

Pre- and Post-Test Nonparametric (Spearman) Intercorrelations for Team-level Teamwork Process Measures. Pre-test correlations are in the lower half (pre-test n=15; post-test n=14)

Teamwork process	1	2	3	4	5	6
1. Adaptability		.49*	.48*	.40	-.05	.68**
2. Coordination	-.38		.25	.42	.30	.79**
3. Decision making	-.15	.55*		.40	-.36	.41
4. Interpersonal	-.20	.59**	.16		-.15	.78**
5. Leadership	.24	-.24	-.01	.15		.25
6. Communication	-.01	.63**	.70**	.69**	.09	

*p < .05; **p < .01 (two-tail)

Table 9

Nonparametric (Spearman) Intercorrelations for Team-level Teamwork Process Measures, Collapsed Across Occasions (n=29)

Teamwork process	1	2	3	4	5
1. Adaptability					
2. Coordination	.08				
3. Decision making	.29	.37*			
4. Interpersonal	.16	.49*	.29		
5. Leadership	.12	.08	-.12	.03	
6. Communication	.37	.71*	.55*	.75*	.23

*p < .05 (two-tail)

The correlations shown in Table 8 show a pattern that differed between the pre- and post-test. Significant moderate correlations were observed between coordination and (a) decision making and (b) interpersonal processes for the pre-test and overall (i.e., collapsed across occasions). For the post-test, significant moderate correlations were observed between adaptability and (a) coordination and (b) decision making. For communication, which is a composite of all other processes, the pre-test data show significant and high correlations with coordination, decision making, and interpersonal processes. For the post-test, significant and high correlations were observed between communication and adaptability, coordination, and interpersonal. Overall, communication correlated significantly with coordination, decision making, and interpersonal processes. Thus, the pattern of correlations among the teamwork processes differed by occasion with no obvious pattern of relations.

Team-level Outcome Measures

The knowledge mapping task generated a total of four team-level outcome measures: semantic content score, organizational structure score, number of terms used, and number of links used. Table 10 gives the descriptive statistics for each measure. Performance data for the experts used to score groups' maps are

Table 10

Descriptive Statistics for Team-level Outcome Measures

Outcome measure	Occasion	Mean	SD	Min.	Max.
Semantic content score	Pre[a]	4.35	2.62	0.75	9.50
	Post[b]	5.98	3.12	0.50	10.30
Organizational structure score	Pre[a]	0.19	0.07	0.10	0.30
	Post[b]	0.26	0.06	0.10	0.30
Number of terms used	Pre[a]	13.87	3.96	7	18
	Post[b]	15.36	2.62	10	18
Number of links used	Pre[a]	17.00	5.95	8	26
	Post[b]	22.21	7.68	10	39

[a]n=15; [b]n=14

given in Table 11. As Table 11 shows, group maps (Mpre=4.35, SDpre=2.62; Mpost=5.98, SDpost=3.12) were considerably lower than the expert criterion maps (M=22.5, SD=21.5). This result is consistent with our pilot data and our assumption that our participants were content novices. This results also suggests

Table 11

Knowledge Mapping Scores for Experts

Outcome measure	Expert 1	Expert 2	Expert 3	Expert 4	Expert average
Semantic content score*	25.3	23	20.7	21	22.50
Organizational structural score	.48	.51	.41	.45	.46
Number of concepts	18	18	18	18	18
Number of links	60	38	39	42	44.75

*Each expert was scored against the other three experts.

that the Herl et al.'s (1996) scoring technique can discriminate between expert and novice knowledge maps.

T-tests (two-tailed) on the outcome measures between the pre- and post-occasions showed a significant difference on the number of links used. The post-test group had more links in their knowledge maps than the pre-test group, $t(27)=2.05$, $p < .05$. There were no other significant differences on the other measures. Interestingly, the semantic content score increased across occasions, although this difference was non-significant ($t(27)=1.53$, $p=.14$).

Tests of Hypotheses

We hypothesized several relationships between the teamwork processes and outcome measures. Our unit of analysis was the team (n=15, pretest; n=14, post-test), not the individual. Table 12 summarizes the set of relationships we expected for our group knowledge mapping study. Table 13 shows the correlations between the team process and the outcome measures. These correlations were used as the primary test of the hypotheses listed in Table 12.

Table 12

Expected Relationships for the Group Knowledge Mapping Task

Teamwork process	Expected relationships
Adaptability	No overall effect on knowledge mapping performance. However, we expected more effective teams to detect problems with the surface features of the knowledge mapping task.
Coordination	We expected more effective teams to be sensitive to time constraints of the task.
Decision making	More effective teams should use more decision making messages than less effective teams.
Interpersonal	No effect on performance.
Leadership	No effect on performance.

Table 13

Nonparametric (Spearman) Correlations Between Team Processes and Outcome Measures by Occasion

Team process	Occasion	Semantic content score	Organizational structure score	No. of terms used	No. of links used
Adaptability	Pre[a]	-.11	-.29	-.41	-.46*
	Post[b]	-.65**	-.49*	-.62**	-.66**
Coordination	Pre[a]	-.27	-.32	-.14	-.10
	Post[b]	-.31	-.31	-.31	-.25
Decision making	Pre[a]	-.64**	-.67**	-.54*	-.44
	Post[b]	-.09	-.01	.02	-.37
Interpersonal	Pre[a]	-.24	-.30	-.21	-.10
	Post[b]	-.33	-.25	-.09	-.27
Leadership	Pre[a]	-.16	-.08	-.30	-.22
	Post[b]	.23	.37	.28	.03
Communication	Pre[a]	-.63**	-.69**	-.53*	-.37
	Post[b]	-.36	-.24	-.18	-.37

[a]n=15; [b]n=14; [c]n=29; *p < .05; **p < .01 (two-tail)

Our first hypothesis was that adaptability processes would have little effect on team performance. Given our assumption of content novices, we did not expect teams to be able have the requisite knowledge to detect inaccuracies in their map; thus, our expectation was that adaptability would have some relation with the surface features of knowledge mapping - the number of nodes and links. The results for this prediction was mixed. The correlations between a team’s use of adaptability messages and their performance, as measured by the content score of the knowledge map, was not significant for the pre-test group and overall (i.e., pre- and post-test groups combined). For the post-test group, however, a significant negative correlation was found (r_s=-.65, p <.01) between adaptability and content score. This suggests that the less a team used adaptability messages, the higher their score was on the knowledge map.

We also expected adaptability to be related to surface features of the map. One gross measure of surface features is the number of nodes and links in a knowledge map. For the number of nodes, significant negative correlations between adaptability and the number of nodes were found only for the post-test group (r_s=-.66, p <.01) and overall (r_s=-.50, p <.05). For the number of links, use of adaptability messages correlated significantly with the number of links for all three groups (pre-test, r_s=-.46, p <.05; post-test, r_s=-.66, p <.01; overall, r_s=-.54, p <.05).

Our second hypothesis was that the use of coordination messages would be associated with performance. We did not find evidence for this in our data. There was no significant correlations between use of coordination messages and the content score of the knowledge map, or any other outcome measure.

Our third hypothesis was that more effective teams would use of decision making messages would be positively associated with performance. The results for this prediction was puzzling. The correlations between a team's use of decision making messages and their knowledge map performance was significant but negatively correlated for the pre-test group (r_s=-.64, p <.01), but not significant for the post-test group and overall. In addition, the direction of the relationship was negative, indicating that the *less* a team used decision making messages, the higher their score was on the knowledge map.

For the remaining team processes, interpersonal and leadership, we did not expect to observe any relationship to team performance given the nature of the task. As expected, the correlation between the use of interpersonal messages and knowledge mapping score (for any group) was not significant. A similar result was found for use of leadership messages, with no significant correlation observed between use of leadership messages and team performance for any group.

Our findings are puzzling. Inspection of Table 13 show *negative* correlations between all teamwork process measures and all outcomes measures, across occasions. While only few correlations are statistically significant, the overall pattern of a negative relationship suggests the less a team used messages, the higher they performed on the knowledge mapping task. One possible interpretation of this finding is the split-attention effect (Sweller, 1994). Given the nature of our task - participants must divide their attention among three elements of the computer system, this interpretation is reasonable. Thus, the finding that more use of messages resulted in lower team performance is explored in the next section.

Exploratory Analyses

We conducted two additional analyses to investigate the nature of the relationships we observed in the previous analyses. As mentioned earlier, one possible explanation of the puzzling results is the split-attention effect. In this case, participants divided their attention among different elements of the user-interface: (a) the text box that showed the messages sent among team members, (b) the knowledge map, or (c) the message handout that contained the list of messages. Given this scenario, it may have been the case that high performing teams spent more time concentrating on their knowledge map compared to low performers. To examine these effects teams were classified as high or low performing, based on the team's semantic content score. High performing teams were those in the upper third of the distribution (n=11), and lower performing teams were those in the lower third of the distribution (n=12).

In order to test for the split-attention effect, two measures were derived from the message usage data and a set of measures from the online computer trace data. To get an estimate of the extent to which a team focused on the text box (see Figure 1), messages that suggested they were reading the text were summed. The messages comprising this measure is listed in Table 14. To test the idea that less effective teams focused more on reading the messages in the text box (and thus spent less time focusing on the knowledge map), a t-test was run on the reading-text box measure. A significant difference was found between low and high performing groups on this measure, $t(21)=2.36$, $p<.03$. Low performing groups used significantly more messages that suggested they were focusing on the text box (M=.75, SD=.24) than high performing groups (M=.48, SD=.30). This finding supports the idea that low performing groups were focused more on the text box than high performing groups.

Table 14

Messages that Suggested Participants were Focusing on the Text Box

Message number	Message	Category
7	Link [C] to what? (7)	Adaptability
23	I need to hear from all of you. (23)	Interpersonal
24	I think [M1 / M2 / M3] should contribute more. (24)	Interpersonal
25	It's important that we all contribute. (25)	Interpersonal
26	We all need to participate. (26)	Interpersonal
29	I [agree / disagree] with [M1 / M2 / M3]. (29)	Adaptability
32	Good idea. (32)	Interpersonal
33	Yes (33)	Communication
34	No (34)	Communication
35	I sent the wrong message (clicked on the wrong number). (35)	Communication

Another measure that examined participants' activity was derived from the computer trace data. Each knowledge mapping event was logged by the computer. An event was defined as any computer activity related to adding, deleting, or moving concepts, or any computer activity related to creating, revising, or deleting a link. Tables 15 and 16 show descriptive statistics for the event data and the relationship between map event data and outcome measures. When the knowledge mapping activity of the high and low performing groups were compared, high performing groups generated more knowledge mapping events (*M*=67.1, *SD*=5.23) than low performing groups (*M*=53.83, *SD*=5.11),

Table 15

Pre- and Post-Test Descriptive Statistics for Knowledge Mapping Activity, Team-level

Knowledge mapping activity	Occasion	Mean	SD	Min.	Max.
Number of add-concepts events	Pre[a]	16.67	3.90	11	26
	Post[b]	8.93	4.14	10	28
Number of delete-concepts events	Pre[a]	1.87	2.67	0	8
	Post[b]	3.00	3.01	0	8
Number of move-concepts events	Pre[a]	11.80	12.98	0	42
	Post[b]	17.14	14.04	0	38
Number of create-link events	Pre[a]	21.93	7.86	10	40
	Post[b]	26.36	7.90	12	41
Number of delete-link events	Pre[a]	2.60	5.79	0	23
	Post[b]	1.21	1.19	0	4
Number of revise-link events	Pre[a]	0.20	0.56	0	2
	Post[b]	0.43	0.65	0	2

[a]n=15; [b]n=14

Table 16

Nonparametric (Spearman) Correlations Between Outcome Measures and Knowledge Mapping Activity

		Concept events				Link events				
Outcome measure	Occasion	Add	Delete	Move	Total	Create	Delete	Revise	Total	Outcome total
Content score	Pre[a]	.40	-.45*	.50*	.61**	.48*	-.58*	-.24	.41	.71**
	Post[b]	.10	-.33	.09	.18	.43	-.19	.01	.38	.33
Structural score	Pre[a]	.45*	-.54*	.32	.42	.60*	-.38	-.08	.52*	.68**
	Post[b]	.11	-.30	-.14	-.03	.41	-.16	.13	.39	.16
No. of concepts	Pre[a]	.53*	-.60**	.32	.41	.72**	-.23	-.08	.66**	.71**
	Post[b]	.43	-.23	.12	.31	.50*	-.25	-.17	.44	.42
No. of links	Pre[a]	.39	-.66*	.18	.26	.81**	-.25*	.04	.76**	.58*
	Post[b]	.20	-.13	.0	.1	.72**	-.36	-.11	.64**	.38

[a]n=15; [b]n=14; [c]n=29; *p < .05; **p < .01 (two-tail)

although this difference was not significant ($t(21)=1.80$, $p=.09$). These results strongly suggest that high performing groups were spending more of their time mapping while low performing groups were spending more of their time focusing on the text box.

Table 16 gives the correlations between the outcome measures and knowledge mapping events. Note that some of the outcome measures covary with the knowledge map activity data. For example, the number of concepts and the add concept event naturally covary as does the number of links and the create link event. Thus, these correlations are not interesting. Interestingly, we found significant positive correlations between all outcome measures and total map activity for the combined pre- and post-test sample.

These findings are unsatisfying, yielding no information about the nature of a team's interaction. Given that low performing teams were spending more of their time reading messages, which presumably meant that they were engaging in discussion (shown to be helpful in general [Webb & Palincsar, 1996]), why were these teams performing so low? To get some insight into this question, participants' use of knowledge mapping messages was examined. The purpose for this analyses was to see whether teams profited from discussions about the knowledge map. Thus, messages that were related to knowledge mapping was summed to form a knowledge-mapping discussion measure. The set of messages selected for this measure is shown in Table 17.

No significant difference between high and low performing groups was found on use of knowledge-mapping focused messages, although the direction favored low performing groups (mean rank (low)=14.17 vs. mean rank (high)=9.64). Apparently, the amount of knowledge mapping messages a team sent had little bearing on their performance, which suggests low performing teams were not able to profit from the discussions. This interpretation is supported by the pattern of correlations within each group (see Table 18). For low performing groups, message usage was negatively and low to moderately correlated with the performance measures, although non-significantly. For high performing groups, however, the correlations are positive and in general moderate to high, although non-significant with the exception of the structural score ($r_s=.69$, $p<.05$).

Table 17

Knowledge Mapping Focused Messages

Msg. No	Message	Category
1	Let's add a [concept / link]. (1)	Leadership
2	Let's try to include [C]. (2)	Leadership
3	Let's [add / erase] [C]. (3)	Leadership
4*	What about [C]-[L]-[C]? (4)	Adaptability
4*	What if we add [C]-[L]-[C]? (4)	Decision making
4*	Let's add [C]-[L]-[C]? (4)	Leadership
5	How should we link [C] and [C]? (5)	Adaptability
6	Let's link [C] to [C]. (6)	Leadership
7	Link [C] to what? (7)	Adaptability
8	Make a link to something. (8)	Leadership
9	Did we use all the concepts? (9)	Adaptability
10	If we don't link [C] to [C] with [L], then [...]. (10)	Decision making
11	If we link [C] to [C] with [L], then [...]. (11)	Decision making
12	Move some concepts around to make the map clearer. (12)	Leadership
13	We should focus on one [concept / link] to complete this task. (13)	Coordination
14	We should think how [C], [C], and [C] relate to each other. (14)	Decision making
15	What if we change [L] to [L]? (15)	Decision making
16	What if we [add / erase / change / move] [C / L]? (16)	Decision making
28	You're doing great - keep going. (28)	Interpersonal
31	Any ideas? (31)	Adaptability

*Message 4 was a combination of three message types, adaptability, decision making, and leadership. The category of the message depended on the message option selected. The format of the actual message was "[What about / What if we add / Let's add] [C]-[L]-[C]? (4)."

Table 18

Nonparametric (Spearman) Correlations Between Message Usage and Performance by Low and High Performing Teams

	Group	Semantic content score	Organizational structure score	No. of terms used	No. of links used
Knowledge mapping focused messages	Low[a]	-.27	-.07	-.37	-.28
	High[b]	.30	.69*	-.24	.58

[a]n=12; [b]n=11; *p < .05 (two-tail)

Thus, these data support the idea that low performing groups were unable to profit from discussions about the knowledge map, while high performing groups did to a higher extent. What may have been occurring, consistent with a split-attention effect, was a cycle mediated by the number of messages sent. Participants in the low performing groups were in general sending more messages. More messages meant messages scrolling through the text box faster, thus compelling participants to pay close attention to the text box. Sending more knowledge mapping messages meant little if the message was missed or read but not comprehended. On the other hand, participants in the high performing groups sent less messages overall. Less messages sent meant not only less messages scrolling through the text box (giving participants time to read and comprehend the messages), but also that participants were probably able to spend more of their time focusing on their knowledge map.

Discussion

Our findings support the idea that students using our networked knowledge mapping system were able to jointly construct a knowledge map. The number of concepts used by groups (across occasions) ranged from 7 to the maximum of 18, and the number of links ranged from 8 to 39. Thus, while some groups had sparse maps, others had fairly complex ones. There was nothing in the data, observed by the experimenters, or reported during debriefing sessions that participants had difficulty grasping the notion of knowledge mapping or that they had difficulty (procedurally) making knowledge maps with our computer system. Participants did, however, consistently comment that they did not like the use of predefined messages. Participants reported that they felt the messages were too constrained and they wanted very much to type their own messages.

In terms of team performance, our results are consistent with our assumptions about participants. We assumed that participants had little content knowledge, and participants' semantic content scores were much lower than our experts. This suggests that Herl et al.'s (1996) knowledge map scoring algorithm can discriminate between experts and novices.

Our findings of no significant correlations between most team processes and outcome measures were unexpected. In particular, we expected the use of decision making messages to play a critical role in our group knowledge mapping task. Our decision making messages were designed to give participants the opportunity to consider alternatives among different links and concepts (e.g., “What if add ‘producer’ - ‘requires’ - ‘carbon dioxide’?” or “We should think how ‘producer,’ ‘oxygen,’ and ‘carbon dioxide’ relate to each other”). We expected (as in O’Neil et al., 1997b) that the more teams used these kinds of messages the higher their team performance.

One possible explanation is that our findings are attributable to the low reliability of the team process scales. While we did not do a reliability analyses of the scales in this study, we have evidence that the team process scales have low reliability (see O’Neil, Wang, Chung, & Herl, 1998).

Another possible explanation for these findings is that our task is highly knowledge dependent, and thus our participants lacked the requisite knowledge to be able to engage each other at a substantive level. The use of messages may reflect more the procedural aspects of constructing a knowledge map instead of any substantive discussion about the content. Support for this is seen in the total number of message usage by all groups. Decision making messages, which we designed to allow participant to discuss the content, accounted for 6 percent of messages use compared to 17 to 21 percent for the other categories.

Another explanation may be the split-attention effect (Sweller, 1994). While our task was fairly easy to carry out, the attentional demands were heavy. Selecting messages required participants to examine their message handout, drawing attention away from the map and messages that other members sent. Reading messages that other members sent draws attention away from focusing on the knowledge map, and constructing the map draws attention away from other members’ messages. If focusing on the knowledge map is the most important contributor to constructing good knowledge maps, then it may have been the team leader who engaged others the least to have contributed the most. This is a reasonable assumption given the highly cognitive and knowledge driven nature of knowledge mapping.

Support for the split-attention effect is seen in correlations (a) between the number of messages that suggested participants were focusing on the text box and performance, and (b) between overall knowledge map activity (i.e., the number of times a knowledge map node or link was added, moved, deleted, or revised) and performance. Low performing groups in general apparently spent more time focusing on the text box than on their knowledge map. Further, participants in low performing groups were not able to profit as much from the discussion related to their knowledge maps, possibly because the large amount of messages they were sending was interfering with their ability to read and follow the discussion. These findings suggest that the most important facet of group knowledge mapping is paying attention to the knowledge map. The demands of the system may have induced too heavy a cognitive load on participants.

These findings raise questions about our group knowledge mapping task relative to our teamwork process measures. In particular, what is a *team* and what is a teamwork task? A team is typically a group of individuals who collectively share a common goal, interact dynamically, interdependently, adaptively, and have specific roles or functions to perform. The absence of an individual severely impacts the functioning of a team (McIntyre & Salas, 1995). The purpose of a team is to perform a task that cannot be performed by individuals alone. While this sometimes may be the goal in educational settings, more often the function of small groups in education is to improve individual learning (Webb & Palincsar, 1996). In our current work we have cast a teamwork perspective (as characterized by Tannenbaum et al., 1996) on a weak teamwork task. Our collaborative knowledge mapping task may be more like a small group task than a teamwork task. A measurement system based on teamwork processes may be insensitive to small group processes, which may partially explain our findings.

The second issue raised by our findings is that our computer-based system may have to be revised, especially the user-interface, to lessen the cognitive load placed on participants. We are currently investigating the use of audio (computer generated speech) as a means to deliver the messages. The use of an audio channel for messages should lessen the attentional demands, thus allowing participants to focus more of their time on the knowledge map.

Despite these shortcomings we believe our approach to using networked computers to assess teamwork processes remains viable given the alternatives. The goal of large scale assessment of teamwork skills - let alone small scale ones - remains elusive. Existing approaches are labor and time intensive, often involving hundreds of hours of transcription and coding of video or audio data. These approaches are unrealistic as assessment options, being untimely and prohibitively expensive. Our software architecture is designed to be domain independent and should transfer to other computer-based team environments.

Such a measurement system would offer the capability of quickly assessing team processes and outcomes in educational environments (K-12) or industrial or military training environments.

References

Baker, D. P., & Salas, E. (1992). Principles for measuring teamwork skills. *Human Factors, 34*, 469-475.

Brannick, M. T., Roach, R. M., & Salas, E. (1993). Understanding team performance: A multimethod study. *Human Performance, 6*, 287-308.

Burke, C. S., Volpe, C., Cannon-Bowers, J. A., & Salas, E. (1993, March). *So what is teamwork anyway? A synthesis of the team process literature*. Paper presented at the 39th annual meeting of the Southeastern Psychological Association, Atlanta, GA.

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Dansereau, D. F. (1995). Derived structural schemas and the transfer of knowledge. In A. McKeough, J. Lupart, & A. Marini (Eds.), *Teaching for transfer: Fostering generalization in learning* (pp. 93-121). Mahwah, NJ: Erlbaum.

Herl, H. E., Baker, E. L., & Niemi, D. (1996). Construct validation of an approach to modeling cognitive structure of U.S. history knowledge. *Journal of Educational Research, 89*, 206-218.

HyperCard 2.2 [computer software]. (1993). Cupertino, CA: Apple Computer.

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, NJ: Erlbaum.

Macintosh [computer hardware]. (1993). Cupertino, CA: Apple Computer.

McIntyre, M., & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. A. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 9-45). San Francisco, CA: Jossey-Bass.

Morgan, B. B., Jr., Salas, E., & Glickman, A. S. (1993). An analysis of team evolution and maturation. *Journal of General Psychology, 120*, 277-291.

O'Neil, H. F., Jr., Allred, K., & Dennis, R. A. (1997a). Use of networked simulations as a context to measure team competencies. In H. F. O'Neil, Jr. (Ed.), *Workforce readiness: Competencies and assessment* (pp. 229-254). Mahwah, NJ: Erlbaum.

O'Neil, H. F., Jr., Baker, E. L., & Kazlauskas, E. J. (1992). Assessment of team performance. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 153-175). Norwood, NJ: Ablex.

O'Neil, H. F., Jr., Chung, G., & Brown, R. (1997b). Use of networked simulations as a context to measure team competencies. In H. F. O'Neil, Jr. (Ed.), *Workforce readiness: Competencies and assessment* (pp. 411-452). Mahwah, NJ: Erlbaum.

O'Neil, H. F. Jr., Wang, S., Chung, G. K. W. K, & Herl, H. E. (1998). *Final report for validation of teamwork skills questionnaire using computer-based teamwork simulations*. (Final deliverable to Statistics Canada). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST.)

Oser, R. L., McCallum, G. A., Salas, E., & Morgan, B. B. Jr. (1989, March). *Toward a definition of teamwork: An analysis of critical team behaviors* (NTSC Tech. Rep. 89-004). Orlando, FL: Naval Training Systems Center, Human Factors Division.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3-30). Norwood, NJ: Ablex.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.

System 7.5 [computer software]. (1992). Cupertino, CA: Apple Computer.

Webb, N. (1993). *Collaborative group versus individual assessment in mathematics: Processes and outcomes* (CSE Tech. Rep. No. 352). Los Angeles: Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Webb, N., & Palincsar, A. S. (1996). Group processes in the classroom. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 841-873). New York, NY: Prentice Hall.

Appendix A

Messages Grouped by Category and by Message Number

Items requiring input from participants are delimited with “C”, “L”, or other text between brackets. For example, in message number 6, “Let’s link [C] to [C],” C indicates the list of concepts. An example of the fully resolved message is, “Let’s link ‘oxygen’ to ‘photosynthesis’.” Other text delimited by brackets that are not “C” or “L” (e.g., message number 29, “I [agree / disagree] with [M1 / M2 / M3]”) are substituted directly in the fully resolved message. An example of message 29 fully resolved is “I agree with M2.”

Messages Grouped by Team Process Category

Message number	Message
	Adaptability
4[1]	What about [C]-[L]-[C]? (4)
5	How should we link [C] and [C]? (5)
7	Link [C] to what? (7)
9	Did we use all the concepts? (9)
29	I [agree / disagree] with [M1 / M2 / M3]. (29)
30	What do you think, [M1 / M2 / M3]? (30)
31	Any ideas? (31)
	Coordination
13	We should focus on one [concept / link] to complete this task. (13)
17	Can you hurry before time runs out? (17)
18	Keep track of the time. (18)
19	Time is almost up. (19)
20	We need to get organized to complete this task. (20)
21	We need to hurry to complete this map. (21)
22	We only have X minutes left. (22)
	Decision Making
4[1]	What if we add[C]-[L]-[C]? (4)
10	If we don’t link [C] to [C] with [L], then [...]. (10)
11	If we link [C] to [C] with [L], then [...]. (11)
14	We should think how [C], [C], and [C] relate to each other. (14)

(Appendix A continues)

(Appendix A continued)

Message number	Message
	Decision Making (continued)
15	What if we change [L] to [L]? (15)
16	What if we [add / erase / change / move] [C / L]? (16)
	Interpersonal
23	I need to hear from all of you. (23)
24	I think [M1 / M2 / M3] should contribute more. (24)
25	It's important that we all contribute. (25)
26	We all need to participate. (26)
27	We have to work together on this. (27)
28	You're doing great - keep going. (28)
32	Good idea. (32)
	Leadership
1	Let's add a [concept / link]. (1)
2	Let's try to include [C]. (2)
3	Let's [add / erase] [C]. (3)
4[1]	Let's add [C]-[L]-[C]? (4)
6	Let's link [C] to [C]. (6)
8	Make a link to something. (8)
12	Move some concepts around to make the map clearer. (12)
	Communication
33	Yes (33)
34	No (34)
35	I sent the wrong message (clicked on the wrong number). (35)

[1] Message 4 was a combination of three message types, adaptability, decision making, and leadership. The category of the message depended on the message option selected.

Messages Sorted by Message Number

Message number	Message	Category
1	Let's add a [concept / link]. (1)	Leadership
2	Let's try to include [C]. (2)	Leadership
3	Let's [add / erase] [C]. (3)	Leadership
4[1]	What about [C]-[L]-[C]? (4)	Adaptability
4[1]	What if we add [C]-[L]-[C]? (4)	Decision making
4[1]	Let's add [C]-[L]-[C]? (4)	Leadership
5	How should we link [C] and [C]? (5)	Adaptability
6	Let's link [C] to [C]. (6)	Leadership
7	Link [C] to what? (7)	Adaptability
8	Make a link to something. (8)	Leadership
9	Did we use all the concepts? (9)	Adaptability
10	If we don't link [C] to [C] with [L], then [...]. (10)	Decision making
11	If we link [C] to [C] with [L], then [...]. (11)	Decision making
12	Move some concepts around to make the map clearer. (12)	Leadership
13	We should focus on one [concept / link] to complete this task. (13)	Coordination
14	We should think how [C], [C], and [C] relate to each other. (14)	Decision making
15	What if we change [L] to [L]? (15)	Decision making
16	What if we [add / erase / change / move] [C / L]? (16)	Decision making
17	Can you hurry before time runs out? (17)	Coordination
18	Keep track of the time. (18)	Coordination
19	Time is almost up. (19)	Coordination
20	We need to get organized to complete this task. (20)	Coordination
21	We need to hurry to complete this map. (21)	Coordination
22	We only have X minutes left. (22)	Coordination
23	I need to hear from all of you. (23)	Interpersonal
24	I think [M1 / M2 / M3] should contribute more. (24)	Interpersonal
25	It's important that we all contribute. (25)	Interpersonal
26	We all need to participate. (26)	Interpersonal
27	We have to work together on this. (27)	Interpersonal
28	You're doing great - keep going. (28)	Interpersonal
29	I [agree / disagree] with [M1 / M2 / M3]. (29)	Adaptability

(table continues)

(table continued)

Message number	Message	Category
30	What do you think, [M1 / M2 / M3]? (30)	Adaptability
31	Any ideas? (31)	Adaptability
32	Good idea. (32)	Interpersonal
33	Yes (33)	Communication
34	No (34)	Communication
35	I sent the wrong message (clicked on the wrong number). (35)	Communication

[1]Message 4 was a combination of three message types, adaptability, decision making, and leadership. The category of the message depended on the message option selected. The format of the actual message was "[What about / What if we add / Let's add] [C]-[L]-[C]? (4)."

Appendix B

Message Handout

Adding Concepts and Links ...

Let's add a [concept / link]. *(1)*
Let's try to include [C]. *(2)*
Let's [add / erase] [C]. *(3)*
[What about / What if we add / Let's add] [C]-[L]-[C]? *(4)*

How should we link [C] and [C]? *(5)*
Let's link [C] to [C]. *(6)*
Link [C] to what? *(7)*
Make a link to something. *(8)*

Discussing Concepts and Links ...

Did we use all the concepts? *(9)*
If we don't link [C] to [C] with [L], then [...]. *(10)*
[...] *means type in your reason.*
If we link [C] to [C] with [L], then [...]. *(11)*
[...] *means type in your reason.*
Move some concepts around to make the map clearer. *(12)*
We should focus on one [concept / link] to complete this task. *(13)*
We should think how [C], [C], and [C] relate to each other. *(14)*
What if we change [L] to [L]? *(15)*
What if we [add / erase / change / move] [C / L]? *(16)*

Keeping Track of Progress ...

Can you hurry before time runs out? *(17)*
Keep track of the time. *(18)*
Time is almost up. *(19)*
We need to get organized to complete this task. *(20)*
We need to hurry to complete this map. *(21)*
We only have X minutes left. *(22)*

Messages about the Group ...

I need to hear from all of you. *(23)*
I think [M1 / M2 / M3] should contribute more. *(24)*
It's important that we all contribute. *(25)*
We all need to participate. *(26)*
We have to work together on this. *(27)*
You're doing great - keep going. *(28)*

Quick Responses ...

I [agree / disagree] with [M1 / M2 / M3]. *(29)*
What do you think, [M1 / M2 / M3]? *(30)*
Any ideas? *(31)*
Good idea. *(32)*
Yes *(33)*
No *(34)*
I sent the wrong message (clicked on the wrong number). *(35)*

Concepts ...

atmosphere
bacteria
carbon dioxide
climate
consumer
decomposition
evaporation
food chain
greenhouse gases
nutrients
oceans
oxygen
photosynthesis
producer
respiration
sunlight
waste
water cycle

Links ...

causes
influences
part of
produces
requires
used for
uses