Knowledge Maps and Their Use in Computer-Based Collaborative Learning Environments

By: William A. Kealy


Made available courtesy of Baywood Publishing: http://www.baywood.com/

*** Note: Figures may be missing from this format of the document

Abstract:
Conceptual learning of 13 graduate students at a university in the southeastern United States was studied during a six-week course that employed the Internet for collaborative learning of online articles. Five groups were formed with each creating a knowledge map—a diagram that graphically arranges and interconnects concepts to show their relationship--during the first and fifth week of the course. Multidimensional scaling (MDS) analysis was used to study the change in each group's mapping over time to determine the influence of a computer-based collaborative learning environment on conceptual understanding. MDS analysis was also used to compare a knowledge map and a similarity rating (of the same 10 concepts) completed by each student at the end of the course. Results of the first analysis indicated that, despite the collaborative nature of the learning environment, groups did not become more similar over time in their understanding of key concepts, However, the second analysis revealed common student perceptions about the dimensions that characterized the conceptual relationships involved. Importantly, this analysis provided evidence that knowledge maps are comparable to rating instruments, thereby supporting recent research claims that they are valid representations of conceptual knowledge. Prescriptions are made for the expanded role of MDS in collaborative knowledge mapping activities.

Article:
An important question that continues to interest educators is how to integrate sound instructional practices with educational technologies in ways that most effectively enhance student learning. Two responses to this question during the past decade have been the emergence of computer-based learning environments and computer-based learning tools. An example of the former is a computer-supported intentional learning environment (CSILE)—a network within which learners can collaboratively generate and share knowledge (Scardamalia & Bereiter, 1994).

Computer-based learning tools, on the other hand, are designed to assist learners in their thinking and reasoning processes. Examples include Inspiration (Anderson-Inman & Zeitz, 1993), SemNet (Fisher, 1990), DCS-4 (Aidman & Egan, 1998), and Learning Tool (Beyerbach & Smith, 1990), commercial computer programs that allow easy and rapid creation of a family of diagram-like graphic displays known variously as "concept maps," "semantic networks," and "knowledge maps." As an alternative to commercially available software for creating these displays, some organizations have developed their own mapmaking computer applications using authoring systems or languages such as, respectively, HyperCard and Java (Herl, O'Neil, Chung, & Schacter, 1999).

This article represents an initial exploration of the use of knowledge mapping within a computer-based learning environment and how these aspects of instruction influence one another.

Knowledge maps (the term used throughout this article) depict either concrete (e.g., "smog") or defined (e.g., "Greenhouse Effect") concepts as spatially distributed nodes that are interconnected by straight-line segments called links. The links are used to show a semantic association between two concepts. It is also commonplace to label either all or a selected number of these links with phrases like "is a part of" or "contributes to" to indicate the nature of the relationship between nodes (e.g., causal, hierarchical, temporal). Taken together, nodes, links, and labels yield prepositional statements such as "smog contributes to Greenhouse Effect" that efficiently depict
the important concepts within a domain and their semantic relationships (Jonassen, Beissner, & Yacci, 1993; Newbern, Dansereau, & Patterson, 1997; Novak, 1990).

Knowledge maps originated from two separate but related psychological perspectives: Ausubel's hierarchical memory theory and Deese's associationist memory theory (Jonassen et al., 1993; Shavelson, Lang, & Lewin, 1994). In the case of the former, the displays (i.e., "concept map") were intended to externalize the cognitive structure of a learner as a means for revealing what the learner knows. Within the latter tradition, similarly configured displays represented "semantic network" that exemplified the notion that a concept is defined by its relationship to other concepts (Shavelson et al., 1994). However, there is little structural difference between a concept map and a semantic network other than the requisite hierarchical organization of the Fowler (Lambiotte, Dansereau, Cross, & Reynolds, 1989; Shavelson et al., 1994).

Knowledge mapping, has been shown to influence cognitive outcomes by facilitating the analysis, elaboration, and integration of ideas (Beissner, Jonassen, & Grabowski, 1994; Wandersee, 1990), promoting reflection on the thought process (Beyerbach & Smith, 1990), and serving as a retrieval cue in STM for related text (Newbern et al., 1997). In the classroom these displays have been used to aid the study or thinking strategies of learners by graphically specifying the critical content in a text or lesson, how it is semantically organized, and how it relates to what they already know (Lambiotte et al., 1989; Novak, 1990).

Although knowledge maps can vary greatly in their surface features, Figure 1 exemplifies the basic components (e.g., nodes, links, labels) and spatial arrangement found in most displays of this type. Nodes on a knowledge map have been portrayed in various ways including as geometric shapes (e.g., circles, boxes) or as images (e.g.,

![Knowledge map diagram](image-url)
icons, pictures) enclosing a descriptive label. As mentioned earlier, links are used to graphically signal the semantic connection between two concepts. However, the relationship between two nodes may also be expressed by their relative spatial position; for example, placing one concept above another can suggest a hierarchical relationship. By the same token, connecting two nodes with an arrow versus a simple line implies that concepts are causally related (Lambiotte et al., 1989).

**KNOWLEDGE MAPPING AND CSILE**

Use of computer-based learning tools like knowledge maps in concert with computer-based learning environments is mutually beneficial: networked environments suggest innovative uses for existing tools while employment of these tools fosters a more effective use of CSELE-type instructional settings. The interplay between learning tool and learning environment is evident in the analysis of Perkins (1991) that identified five requisite components of learning environments: symbol pads, construction kits, phenomenaria, information banks, and task managers. Symbols pads, like notebooks and laptop computers, allow for the manipulation and storage of words, numbers, and images. Construction kits accomplish a similar cognitive outcome but with prefabricated materials such as Legos. Aquariums, simulations, and other "phenomenaria" facilitate direct learning experiences through real or virtual microcosms (Perkins, 1991).

Symbol pads, construction kits, and phenomenaria can be easily classified as learning tools that, like knowledge maps, act as a sort of "paintbrush" and "canvas" for adding substance to otherwise abstract ideas. On the other hand, information banks (e.g., encyclopedias, teachers) and task managers (e.g., text-book exercises, teachers) are a function of the learning environment itself, allowing for, respectively, the storage and management of knowledge products.

The distinction between components of a learning environment that act as tools and those that work as an integral part of the instructional milieu is especially pertinent for the types of learner-centered educational settings made possible with computer-based collaboration. The limitless information available through the Internet and the potential for peer management of the instructional process through computer networking enhances the attractiveness of a constructivist approach to education. At the same time, the open-ended nature of such a perspective would likely foster, as Perkins (1991) speculates, greater reliance on tools like phenomenaria and construction kits that emphasize learning through exploration and discovery.

As a collaborative activity, knowledge map construction can be a useful method for generating shared understanding about the meaning of concepts among members of the learning community. Such consensus building can be accomplished by either having students construct maps as a group or by providing the means for the exchange of individually constructed maps. When used within a community of learners (COL), knowledge maps essentially provide a "linguistic shorthand" for concepts that facilitates communication and the sharing of ideas.

Use of knowledge mapping within a COL raises two questions that focused this study. The first is whether knowledge maps created by groups of learners within a CSILE become increasingly similar over time as a reflection of the collaborative learning that occurs. This issue, in turn, raised a second question that pertains to the practice of knowledge mapping itself: is the knowledge map a valid representation of the knowledge structures (Shavelson et al., 1994) possessed by the map maker(s)? In particular, are knowledge maps suitable for making comparisons between collaborating groups of mapmakers to infer whether or not their knowledge structures have become similar over time?

**COLLABORATIVE KNOWLEDGE MAPPING**

Collaboratively generated and shared knowledge is the very hallmark of a community of learners. But could such collaboration lead to greater homogeneity in the appearance and structure of knowledge maps produced by its members over time? In other words, do knowledge maps made by course participants become increasingly similar in their implied cognitive structure over time as a function of ongoing collaboration? The main purpose of this preliminary investigation was to address these questions as an important first step in understanding how
learning context influences cognitive tools in general and knowledge maps in particular. Evidence suggesting reciprocity between CSILE and knowledge mapping would, among other things, point to the need for follow-on research to determine whether this was also true of other knowledge tools (with collaborative potential) such as on-line surveys.

Given the nature of knowledge mapping, it is not unreasonable to speculate that use of this tool could be influenced by the setting in which it is used. Knowledge mapping is, first and foremost, a skill that is easily (some would say intuitively) acquired. If not for the graphical idiosyncrasies that reflect individual differences in content knowledge, knowledge maps might all seem very similar in appearance. Hence, both the forms of knowledge maps and the mapping process itself are probably susceptible to aspects of the learning environment in which they are employed.

Ironically, the ease that individuals have in knowledge mapping may be a factor that undermines its usefulness as a collaborative activity since it can be accomplished without a team effort (Chung, O’Neil, & Berl, 1999). Though teams have successfully carried out knowledge mapping, with members collaborating remotely on networked computers, it has not been shown to be an inherently beneficial activity for team building. This was recently demonstrated in two one-year longitudinal studies concurrently carried out by Herl et al. (1999). In one study, teams of three students each worked over a computer network to collaboratively develop concept maps. In another study, students created individual concepts maps throughout the school year while consulting a simulated Internet Web site. Subjects in the latter study made significant knowledge gains based on four knowledge map assessment criteria (semantic content, organizational structure, number of terms, and number of links). In the other study, by contrast, teams that collaborated on the knowledge mapping task only negligibly improved in their content knowledge.

**KNOWLEDGE MAP VALIDITY**

Historically, knowledge maps have been used far more as tools for learning than as tools for learning assessment (Shavelson et al., 1994). With the advent of their use for assessment purposes, however, researchers began raising questions about the validity and reliability of knowledge mapping (see Ruiz-Primo and Shavelson (1996) for an extensive review on this topic). How well do concept terms on a knowledge map represent the overall subject domain? How generalizable axe map scores over raters and test occasions?

Interest in the validity of knowledge mapping was a second concern for the present study that naturally proceeded from its primary intent of assessing the similarity of cognitive structures among collaborating learners over time. Given the main purpose of the study, it was useful to think of validity as "the extent to which inferences to students' cognitive structures' on the basis of their concept map scores can be supported logically and empirically" (Shavelson et al., 1994, p. 22). It would be impossible, for example, to conclude that learners grow increasingly similar in their understanding of conceptual relationships based on changes in their knowledge maps if it could not also be shown that such displays are valid representations of the cognitive structures they portray.

Early attempts to establish the reliability and validity of knowledge mapping for assessment revealed poor correlations between mapping and other measures of learner achievement such as course grades and performance on standardized tests (Aidman & Egan, 1998; Shavelson et al., 1994). Data from recent studies, however, provide mounting evidence that knowledge maps are both reliable (Herl, Baker, & Niemi, 1996; Heil et al., 1999) and valid (McClure, Sonak, & Suen, 1999; Rice, Ryan, & Samson, 1998; Ruiz-Primo, Schultz, & Shavelson, 1997) assessment instruments.

This more favorable view toward map-based assessment emerges concomitant with a broader perspective of assessment as a system that includes a task, a response format (e.g., use of pencil-and-paper or computer), and a scoring system (Ruiz-Primo & Shavelson, 1996; Ruiz-Primo, Shavelson, & Schultz, 1997; Shavelson et al., 1994). Knowledge mapping tasks, for instance, may be "top-down" or "bottom-up," that is, either created with a predetermined framework or drawn from scratch, respectively (Lambiotte et al., 1989). Comparing the two
approaches, Ruiz-Primo, Schultz, Li, and Shavelson (1999) found the latter "low-directed" mapping task yielded scores that more validly represented learners' conceptual knowledge.

Yet, it is the scoring system itself, suggest Rice, Ryan, and Samson (1998)—and especially the criteria for allotting points—that may be the most influential factor for the validity of a knowledge mapping since the method of scoring "is a critical determinant of the strength of the relationship between map scores and scores on other assessments" (p. 1107). Scoring systems have generally employed one of three approaches: 1) scoring of map components (e.g., links, propositions), 2) use of a criterion map, and 3) a combination of both of these schemes (Ruiz-Primo, Shavelson, & Schultz, 1997). According to Shavelson et al. (1994), one of the most widely used scoring methods has involved judgments of the similarity between a learner's maps and one developed by an expert. Whatever the specific scoring system used, however, its selection should be, first and foremost, guided by the purpose for which the map scores are to be used (Rice et al., 1998).

The importance of considering the purpose map scores are to serve when determining the scoring system itself is illustrated by a research study of Osmundson, Chung, Herl, and Klein (1999) that examined students' development of conceptual understanding about the human circulatory, respiratory, and digestive systems over a six-week period by using knowledge mapping as a repeated measure. In this study, the researchers determined that a score based on an "expert map" was "somewhat limited in its capability to measure the nature and quality of change in students' ideas" (p. 4). Instead, they devised an assessment scheme consisting of 1) expert content scores, 2) the quality of propositions formed by student maps, and 3) the number and quality of connections that a proposition made between any two bodily systems. Results showed learners who collaboratively developed knowledge maps acquired significantly more scientific and principled information about the circulatory, respiratory, and digestive systems than students who collaboratively researched the human body using various instructional resources. More importantly for the present discussion, a significant and moderate correlation was observed between students' knowledge map content scores and ratings of essays completed during the last week of the study. This lends support to the position of Rice, Ryan, and Samson (1998) that knowledge map validity is partially controlled by whether or not the scoring system used reflects the intended purpose of the scores.

Within the context of the current study, the degree to which collaborating groups of learners become increasingly similar in their conceptual understanding was the paramount concern. By contrast, the issue of how much knowledge learners actually acquire through computer-based collaborative knowledge mapping was considered inconsequential. Conceivably, a situation could even arise whereby the cognitive structures of collaborating learners grow more similar over time but are composed of misconceptions and mistaken ideas. In any event, a map-scoring scheme based on comparisons to expert maps was thought to be of little value. Rather, the study's purpose directed the search for a scoring method that would reveal the similarities in cognitive structures among collaborating groups of students.

One promising strategy for observing the degree of homogeneity among learners in their cognitive structure is to have them rate the similarity or contrast between pairs of concepts. In the past few years, researchers have begun to combine the practice of concept ratings with knowledge mapping procedures in innovative ways for improving map construction as well as map assessment (Aidman & Egan, 1998; Herl et al., 1996; Trochim, Cook, & Setze, 1994). In some instances, ratings have been used to form similarity or proximity matrices that are then analyzed using multidimensional scaling (MDS) techniques. Results of MDS analyses have been used to create computer-generated, "implicit" knowledge maps (Aidman & Egan, 1998) that are compared to ones constructed by experts. Alternatively, MDS analyses of rating data have been directly compared against knowledge map scores (based on an expert map) to validate the assessment of the learners' cognitive structure (Herl et al., 1996). While the two approaches exemplify Trochim's (1999b) recommendation of contrasting knowledge mapping results with those obtained from another method to establish mapping validity, both involve the use of expert maps. Few, if any, studies to date, on the other hand, have employed MDS analysis as a common basis for judging the comparability of similarity ratings and knowledge maps for the same concepts. This was the strategy adopted by the present study.
METHOD

Participants
Thirteen graduate students (3 male, 10 female) enrolled in a six-week course on integrating technology into the classroom participated in the study. With the exception of one participant, all were education majors (six studying instructional technology). All were experienced computer users who were naive to the fact that they were partaking in a research study.

Learning Environment
Both the notion of CSILE (Scardamelia & Bereiter, 1994) and its use as a strategy for infusing technology into instruction was a theme that pervaded the course. During the entire course, students experienced learning through a CSILE approach using an interactive computer program named Construe (Lebow, Wager, Marks, & Gilbert, 1997). Accessed through a site on the World Wide Web, Construe allowed students to select and examine 33 full-text articles from academic journals covering a wide range of topics on learning and instructional technology.

Study participants read one article each week, writing a reaction to each reading using an on-line form in Construe that contained the response categories, "Gut Reactions," "Big Ideas," "Implications for Teaching," and "Nagging Questions." Reading reactions were automatically uploaded to a server where, through Construe, they could be accessed and read by anyone in the class. In addition to the reading reaction, students used an on-line form to define two concepts encountered in each article read. These concept definitions were posted on Construe, forming a pool of concepts from which items on the course's final examination were selected.

A key argument made in Scardamelia and Bereiter's (1994) seminal work on learning environments is that classrooms should operate as places where knowledge is produced, much like research "think tanks" found in business and industry. Yet, two aspects of corporate research teams that are usually absent in the classroom are 1) collaboration is indispensable because the problem faced is unsolvable by an individual, and 2) collaborators bring their unique subject matter expertise to bear on the group's problem.

In order to more closely mirror this corporate research model, thereby maximizing the potential offered by a CSILE environment, the class was divided into teams of "experts" faced with a problem too daunting for one person to undertake: the acquisition of knowledge from the 33 articles posted in Construe within a mere six-week period.

Teams were determined by having students complete a survey at the beginning of the course in which they rated, on a score of one to four, both their experience and interest in five areas: educational theories and concepts, instructional design and practice, educational technology and media, psychology of learning and cognition, and individual and societal characteristics of learners. Based on the scores obtained, two or three students were assigned to each group to serve as class "expert" for that particular subject area. At every class session, one member from each group was responsible for reporting to the entire class on six or seven works from the entire pool of articles that pertained to their field of experience and interest. Periodically, all the students viewed a matrix that documented which articles had been reported on by the different groups and, hence, were able to keep abreast of the progress the class as a whole was making toward covering the course material. In this manner a setting was created whereby learner collaboration was a necessary, viable, and potentially effective strategy for success.

PROCEDURES
During the first day of class, the researcher (i.e., course instructor) discussed the concept of COL and directed course participants to read the Scardamelia and Bereiter (1994) article (approximately 7,500 words) describing CSILE. The following week the class participated in a brainstorming session to identify 10 major concepts from the article: collaborative knowledge building, cognitive approach, community of learners, CSILE, discourse, intentional learning, process of expertise, progressive problem solving, school restructuring, and second-order environments.
Students were then introduced to the idea of knowledge mapping and were shown examples of maps created by previous classes. The class was also given a procedure for developing knowledge maps that included generating, linking, and labeling activities. The generation phase involved a brainstorming session to identify all the concepts within a given knowledge domain—such as the activity they had just completed dealing with the reading assignment. Each concept, it was explained, could be represented by a labeled shape or image, serving as a feature or "node" on the map. Linking entailed drawing a single line to connect two or more concepts that were associated in some way (e.g., semantically, causally, temporally). Students were also told they could place arrowheads on either or both ends of a link to indicate, respectively, a causal or reciprocal relationship between concepts. The class was urged to further articulate the relationships between concept nodes by labeling the links of their knowledge map. Such labeling of links is a common practice in the design of knowledge maps that, in part, distinguishes them from other types of spatial information displays like graphic organizers (Moore & Readence, 1984).

During this instruction students were also strongly encouraged to use proximity to signal degree of conceptual relationship: two concepts perceived to be highly related would be situated close to one another whereas relatively dissimilar concepts would be farther apart. This practice follows the notion of "semantic distance" proposed by Winn and Holliday (1982) in their prescriptions for the design of diagrams—the distance between elements on a display should correspond to either their degree of similarity in meaning or their functional association. There is also some empirical evidence that use of proximity in the design of knowledge maps enhances their performance as a learning tool (Wiegmann, Dansereau, McCagg, Rewey, & Pitre, 1992). Following the discussion on knowledge maps, students were shown how to use Inspiration, a popular computer program that facilitates knowledge mapping, brainstorming, and other tasks. Students then practiced using the program in their assigned teams to develop a knowledge map incorporating the 10 concepts generated from the earlier brainstorming session.

At each class period over the next three weeks, one member from each of the five teams discussed the Reading Reaction for an article read that week. During the fifth week of the course, the five teams again created knowledge maps of the same 10 concepts used in constructing their first map. A week later, students took a course exam in which they wrote an essay on how to apply the 10 concepts provided in integrating technology and instruction. The 10 concepts given were drawn from the pool of concept definitions developed by the class throughout the course: authenticity, cognitive apprenticeship, constructivism, cooperative learning, distributed knowledge, intentional learning, metacognition, multicultural education, reciprocal teaching, and zone of proximal development.

Upon completion of the essays, each student created a knowledge map of the 10 concepts on the test. During the mapmaking task, students were told to depict each concept node using just an ellipse. This constraint was placed on their mapmaking for two reasons. First, it focused student attention on the interrelatedness of concepts rather than the descriptive characteristics of each concept. Second, this facilitated map scoring by both eliminating pictorial aspects of map nodes and simplifying the task of measuring between concepts. As students constructed their maps, they were reminded to use linking, labeling, and spatial proximity to articulate the conceptual relationships between nodes. Figure 1 shows one of the completed knowledge maps that were typical of the displays constructed by students in this phase.

Four days later at the very end of the course, students completed a rating instrument to report the perceived similarity between the 10 concepts on the final exam. The instrument contained 45 items, each containing a different pair of concepts to reflect a balanced order of pairs for \( N = 10 \) (Dunn-Rankin, 1983). Each item listed a pair of concepts to the far left and four descriptive responses to the right: "totally related," "partially related," "fairly unrelated," and "completely unrelated." A rectangle enclosed each pair of concepts along with its set of responses. The 45 items were then separated from one another by a 5/8 inch (15.5 cm) space. Nine response items were arranged on a page with the five total pages shuffled and stapled in four different orders to form four versions of the instrument. Equal numbers of the versions were again shuffled and randomly
distributed to control for a possible response set by participants. Students were instructed to individually assess, for each item, the relatedness of the two concepts shown by circling one of the four responses provided.

**SCORING**

*Map Data*

All knowledge maps were reproduced in color on 8-1/2 x 11 inch (21.59 x 27.94 cm) white bond paper using a feature of *Inspiration* that scaled the printed displays to fit between the one-inch (2.54 cm) left and right margins of the page. For every map, the distances between the centers of all possible pairs of concept nodes were measured to the nearest tenth of a centimeter. Also, for each pair a "1" was assigned if a link (i.e., connecting line segment) had been drawn on the map and a "0" if no link was present. For each map scored, data on all node pairs were recorded on the triangular matrix formed by the intersection of the 10 concepts alphabetically listed as a column along the left margin and as a row across the top of the page.

Distance and link information were consolidated into one metric, reflecting the similarity between any given pair of nodes, by applying a modified version of Gower's Similarity Measure (as cited in Dunn-Rankin, 1983). In the current study, this involved a two-step procedure that first calculated an index based on the physical proximity of two nodes and then adjusted this value according to the presence or absence of a node link. Step One, or calculation of the Physical Proximity Index (PPI), was accomplished using the formula:

\[ PPI = 1 - \frac{\Delta_N}{(\Delta_L + \Delta_S)} \]

where, \(\Delta_N\) = distance between two selected nodes of a map, \(\Delta_L\) = largest inter-node distance on the map, \(\Delta_S\) = smallest inter-node distance on the map. In Step Two, a Dissimilarity Index (DI) was obtained for each node-pair by subtracting the mean of the PPI and corresponding Link Value (LV) from one:

\[ DI = 1 - \left( \frac{PPI + LV}{2} \right) \]

where, LV = 1 nodes are linked, 0 = if nodes are unlinked.

For purposes of analyzing knowledge maps, dissimilarities rather than similarities were calculated as the proper representation for inter-nodal distances. To illustrate this concept using one of the maps from the study (in which the largest and smallest inter-node distances were 16.9 cm and 2.8 cm, respectively), two linked nodes located 2.9 cm apart yielded a DI of .07. By contrast, two relatively distant and unlinked nodes that were 16.3 cm apart resulted in a DI of .93. In other words, the physical distance between two nodes on a concept map is directly proportional to the DI calculated.

*Rating Data*

Rating instruments completed by students were numerically coded, assigning values of 1, 2, 3, and 4, respectively, for responses of "totally related," "partially related," "fairly unrelated," and "completely unrelated." Hence, a higher number indicated greater perceived semantic distance (i.e., dissimilarity) between two concepts. The resulting scores for each student were then converted into a tri-angular matrix and analyzed.

**DATA ANALYSIS**

Multidimensional scaling (MDS) techniques have long been used to study the perceived relatedness of psychological objects by analyzing how they are physically arranged (e.g., clustered) and rated by subjects (Dunn-Rankin, 1983). Hence, MDS was considered to be an ideal analytical strategy for the current study since it could accommodate both spatial and numerical types of subjective judgments recorded, respectively, by the knowledge maps and rating instruments. Further, the thrust of the study was not hypothesis testing but rather hypothesis generation—a goal for which MDS is ideally suited.
One MD5 analysis contrasted the two knowledge maps developed by teams of students during the first and fifth weeks of the course. This was designed to gauge the effect that a collaborative learning environment has on the collective understanding of concepts taught in a class. The second MDS analysis compared knowledge mapping with an alternative method—in this instance, a rating instrument—for measuring the perceived relatedness of 10 concepts. Results from this analysis were intended to examine the relative validity of concept mapping as a tool for depicting the common understanding of concepts by participants in a collaborative learning environment.

Matrices derived from both knowledge maps and similarity ratings were analyzed using an individual differences multidimensional scaling (INDSCAL) analysis. Four analyses were performed, one for team maps created the first week, another for team maps constructed four weeks later, an analysis of individual maps made at the end of the course, and a fourth analysis on the similarity ratings of concepts. All analyses were performed at the interval level of measurement with the exception of similarity ratings that were treated as ordinal data.

INDSCAL is a weighted Euclidean model that tries to fit psychological objects—in this case, the 10 learning concepts—within an n-dimensional space. (Weightings are used to compensate for the biases of subjects on the importance they attach to a particular dimension of the stimulus space.) Accordingly, the "goodness of fit" in an MDS model, whether based on knowledge maps or similarity ratings, is a function of the consistency, both between and within subjects, in how concepts are represented and perceived.

RESULTS

Data from maps created by Team One (the group reporting on educational theories and concepts) and Team Four (the group reporting on psychology of learning and cognition) during the fifth week of the course was either lost or unreadable. Consequently, only three of the five teams (8 out of the original 13 participants) produced maps that were suitable for analysis. Additionally, the absence of a concept from one of the maps resulted in nine missing values in its corresponding triangular matrix.

While such a loss of data undoubtedly has a negative impact on a study's reliability and validity, in the present study the small number of participants was considered unfortunate but not catastrophic for a few reasons. First, each matrix analyzed represented a relatively large number of data points (i.e., the 10 map nodes yielded 45 conceptual relationships). Second, the study was not concerned with learning achievement but rather with the less critical aim of identifying trends of judgment (similar or dissimilar) among participants. Finally, and most importantly, is that the study represented a preliminary exploration in an area, namely, the interplay between learning tools and learning environments, that has received little research attention.

COMPARATIVE ANALYSIS OF TEAM MAPS

Results from an INDSCAL analysis of team maps created during the first and fifth weeks are reported in Table 1. This shows the "goodness of fit" by two-dimensional and three-dimensional MDS models that represent knowledge maps created by teams of students on the two occasions. Measure of fit is indicated by three indices: S-stress, Kruskal stress, and the squared correlation coefficient ($R^2$), averaged over the three matrices associated with a particular mapmaking session.

Of the two NMS models reported in Table 1, the three-dimensional model yielded the lowest S-stress and Kruskal stress values for both mapmaking sessions (.35 and .21 for Week One and .36 and .22 for Week Five, respectively). Hence, the three-dimensional model is preferred over the two-dimensional model, even though the latter is the more parsimonious MDS solution. Yet, the three-dimensional model does not provide an especially good fit for the data given that even a stress of less than 10 percent represents only a "fair" fit (Dunn-Rankin, 1983).

The INDSCAL analysis indicates that over a third of the mean variance ($R^2 = .376$) in data from knowledge maps in Week One was accounted for by the three-dimensional MDS model. By contrast, data from Week Five maps accounted for a much smaller proportion of variance ($R^2 = .247$).
These results address the first question of the current study, that is, whether CSILE-based learning yields greater homogeneity over time in the appearance and structure of knowledge maps produced by class participants. Increased agreement among students in their understanding of course concepts and how they are interrelated should have manifested itself as a better data fit in the MDS model for Week Five maps than the model based on maps constructed a month earlier. However, MDS analyses indicated stress, one indicator of fit, remained essentially unchanged from Week One and Week Five. Meanwhile, the other indicator of fit, the squared correlation coefficient, actually decreased between the first and second mapmaking session further suggesting students' conceptual understanding did not become more homogeneous over the five-week period.

COMPARISON OF KNOWLEDGE MAPPING AND SIMILARITY RATING

Individual Maps

Each course participant created a knowledge map on 10 learning concepts as part of an examination administered late in the course. Three maps of the 13 participants were discarded because they did not contain all of the 10 concepts on the test. Further, of the remaining maps only those of students who completed a similarity rating instrument two days later were analyzed. While yielding only six proximity matrices for MDS analysis, this was a necessary measure to avoid the confounding effects of comparing two datasets, one from knowledge maps and another from similarity ratings, which were generated by different groups.

An IN DSCAL analysis of individual maps, like that of team knowledge maps, produced indices of fit for both the two- and three-dimensional MDS models of transformed map data averaged across the six students. This data, shown in Table 2, indicate the two-dimensional model may be slightly preferred over the three-dimensional scheme since, despite the increased stress of the former, it accounts for about the same average variability as the latter (R2 values of .217 and .214, respectively) while offering a more parsimonious solution.

<table>
<thead>
<tr>
<th>Index of fit</th>
<th>2-dimensions</th>
<th>3-dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-stress</td>
<td>.459</td>
<td>.374</td>
</tr>
<tr>
<td>Kruskal stress</td>
<td>.334</td>
<td>.237</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.217</td>
<td>.214</td>
</tr>
<tr>
<td>Similarity ratings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-stress</td>
<td>.436</td>
<td>.342</td>
</tr>
<tr>
<td>Kruskal stress</td>
<td>.315</td>
<td>.233</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.306</td>
<td>.364</td>
</tr>
</tbody>
</table>

*a*Eight students comprising three teams, each producing one map. *b*Worst possible fit = 1; best possible fit = 0 for both S-stress and Kruskal stress.

Note: All analyses were at the interval level of measurement.

Table 1. Relative Fit of a Multidimensional Scaling (MDS) Model Based on Knowledge Maps by Students during Week One and Five of a Course

<table>
<thead>
<tr>
<th>MDS model</th>
<th>2-dimensions</th>
<th>3-dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week one maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-stress</td>
<td>.455</td>
<td>.342</td>
</tr>
<tr>
<td>Kruskal stress</td>
<td>.334</td>
<td>.212</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.231</td>
<td>.376</td>
</tr>
<tr>
<td>Week five maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-stress</td>
<td>.485</td>
<td>.357</td>
</tr>
<tr>
<td>Kruskal stress</td>
<td>.330</td>
<td>.223</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.176</td>
<td>.250</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Index of fit</th>
<th>2-dimensions</th>
<th>3-dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-stress</td>
<td>.459</td>
<td>.374</td>
</tr>
<tr>
<td>Kruskal stress</td>
<td>.334</td>
<td>.237</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.217</td>
<td>.214</td>
</tr>
<tr>
<td>Similarity ratings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-stress</td>
<td>.436</td>
<td>.342</td>
</tr>
<tr>
<td>Kruskal stress</td>
<td>.315</td>
<td>.233</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.306</td>
<td>.364</td>
</tr>
</tbody>
</table>

*a*Worst possible fit = 1, Best possible fit = 0 for both S-stress and Kruskal stress.

Note: Analysis of knowledge maps was at the interval level of measurement whereas similarity ratings were analyzed as ordinal data (ties were left tied).
**Similarity Ratings**

In the case of the similarity ratings, the two-dimensional model does not provide as good a fit of the data as the three-dimensional model given the decreased stress and improved $R^2$ of the latter ($R^2$ values of .306 and .364, respectively).

Based on indices of stress, both the two- and three-dimensional MDS models of the rating data were comparable to those derived from the knowledge maps. On the other hand, similarity ratings accounted for a relatively larger proportion of variance in the data than their knowledge map counterparts.

The two-dimensional MDS models derived from knowledge map and similarity rating data are shown superimposed on one another in Figure 2a. This strategy allows a direct comparison of the two schemes for representing conceptual relationship. In both cases, concepts are depicted by two-letter symbols; those
associated with rating data are in boldface while those related to knowledge map data are in a smaller italic font. Each pair of identical concepts from the two sets is connected by a thin dotted fine.

Upon initial inspection of Figure 2a, it would appear that knowledge maps and similarity ratings of the same concepts do not correspond very well in how they capture perceived relationships of the same concepts. For example, Figure 2a shows a close proximity only for members of the concept pairs, Constructivism (CN) and Authenticity (AU).

By contrast, members of all but one of the remaining concept pairs appear to be distantly distributed on opposite sides of an imaginary axis, shown as a dashed line in Figure 2a. Based on this observation, the knowledge map component of the display was flipped—rotated around this axis 180 degrees—while the components of the rating data display were kept in place. The rationale for doing this was that, like an orthogonal rotation performed during a factor analysis, an initial arbitrary frame of reference might be rotated into a more simplified arrangement to improve data analysis. In this instance, one data set was "flipped" into its mirror image and then rotated to maximize correlation with a second data set.

The reconfigured data display (Figure 2b) shows an overall improvement in the association between members of the concept pairs. For half the concepts (ME, IL, MT, DK, and CA) the fit between rating and map data improved greatly and moderate improvement occurred for two others (CL and RT). The fit for concept pairs CN and AU decreased slightly while ZD continued to show a poor fit between its corresponding rating and map data.

From examination of Figure 2b, four clusters of concepts emerge: Group I including Intentional Learning (IL), Cognitive Apprenticeship (CA), and Meta-cognition (MT); Group II consisting of Reciprocal Teaching (RT), Distributed Knowledge (DK), and Cooperative Learning (CL); Group III including the concepts Authenticity (AU) and Constructivism (CN); and Group IV that comprised the concept Multicultural Education (ME). This clustering suggests concepts within Group I deal with the self-regulation of mental processes while those in Group II are related to ideas about shared knowledge. Although the characterization of Groups III and IV is less obvious, the former involves relatively theoretical concepts while the latter deals with social matters.

Further interpretation of this display suggests that students may have perceptually organized the learning concepts they studied within a framework defined by notions of "socialization" and "utility," represented in Figure 2b by Dimensions 1 and 2, respectively. Dimension 1 could conceivably reflect a "personal-social" continuum in which the individual processes of Group I stand in opposition to the social concerns of Multicultural Education (Group IV). Though more tenuous, Dimension 2 seems to express the applicability, theoretical or practical, of the 10 concepts to educational issues. Concepts in Group II, for example, such as Reciprocal Teaching (RT) and Collaborative Learning (CL), are relatively easier to translate into classroom practice than Constructivism (Group III).

Although data from the similarity ratings suggest students believed Zone of Proximal Development (ZPD) to be the most abstract concept of all, this is at odds with the knowledge map data. Such variability raises the possibility that the class in general had a misconception about what ZPD meant. However, examination of concept definitions written by students reveals ZPD was correctly understood by at least those in the class who chose to define the term. Hence, the discrepancy between the mapping and rating data for ZPD remains an anomaly that cannot be explained within the scope of this study.

**DISCUSSION**

Measures of stress obtained in the MDS analysis of similarity ratings and knowledge mapping indicated both data sources were roughly equivalent in their goodness-of-fit to a MDS model. Additionally, graphic representations of the two-dimensional MDS models for map and rating data showed similar conceptual groupings. Both findings suggest that knowledge mapping and similarity rating are comparable means for understanding how students perceive conceptual relationships.
While rating strategies accounted for more variance in student responses than knowledge mapping, this may be attributed to differences in the respective levels of measurement—ordinal and interval of the two data sources. Because ratings are typically analyzed with the assumption (sometimes in error) that a linear relationship is present (Dunn-Rankin, 1983), they may inherently appear to have better explanatory power than knowledge mapping. Conceivably, interval rating data that incorporate a continuous range of numerical responses may fare no better than a comparable knowledge map in accounting for within-subjects variability.

The evidence that knowledge maps are nearly as good as similarity ratings for determining student perceptions about conceptual relationships was an important finding of the study—one that begins to address the issue of knowledge map validity mentioned earlier. By demonstrating, as Trochim (1999b) has proposed, comparability between knowledge mapping and an alternative, but established, assessment tool, this finding implies the former can exhibit an acceptable degree of content validity.

Given this evidence supporting knowledge mapping as a valid assessment tool, we can return to the issue of how maps drawn early in a computer-supported collaborative learning environment compare with those constructed several weeks later. Analysis of maps drawn in the first and fifth weeks of the course showed essentially the same goodness of fit to the MDS model (i.e., stress remained the same while $R^2$ decreased). Consequently, there is no indication that students within a computer-supported intentional learning environment become increasingly homogeneous in their perception and understanding of the concepts involved. If the opposite were the case, one would expect a growing similarity in the organization of knowledge maps created by students over time. Needless to say, such group consensus is neither necessarily indicative of a collaborative learning environment nor a requirement for it.

Unquestionably, results gleaned from this study are greatly limited by its small sample size. Nevertheless, this should be considered a preliminary investigation intended to guide future research. As stated earlier, the goal of this study was to examine the use of a knowledge tool mapping within a computer-based collaborative learning environment. The study also served as a pilot for exploring new methods for scoring and analyzing knowledge maps. To this end, the MDS analysis of knowledge maps created by students generated several useful suggestions for future classroom practice as well as for research related to this particular knowledge tool.

An important outcome of this study was that it begins to address questions regarding the validity and reliability of knowledge mapping by contrasting, as Trochim (1999b) has proposed, the results of knowledge mapping with those obtained from another method. Comparisons made in this study between knowledge maps and similarity ratings strongly suggest the two are roughly equivalent in their ability to portray the conceptual understanding of students. Hence, in terms of content validity, knowledge mapping appears to be comparable to traditional rating instruments.

A second result of the study was in its innovative use of multidimensional scaling for analyzing knowledge maps. The MDS techniques described herein may offer an alternative to other more subjective methods of evaluating a knowledge map, such as its comparison to an ideal or "expert" map (Bonk, Mulvaney, Dodzik, & Reynolds, 1994). Use of an ideal map as the template for judging student maps may be characterized as teacher-centered assessment. On the other hand, use of an MDS-generated map for the same purpose one that depicts the perceptions of the entire class—-holds the potential for a collaborative, student-centered assessment.

Perhaps the most important insight gained from this study is that a learning community may perceive a set of concepts to be globally organized along several characteristic dimensions: abstract-concrete, important-unimportant, theoretical-practical, and so forth. This type of organization clusters and separates concepts according to their gross similarities and differences to form a macrostructure like that shown earlier in Figure 26. Hypothetically, this consolidation of individual concepts into an overall knowledge structure is the inverse of the metacognitive strategy used in reading a general reference map. In the latter instance, research shows people process maps by first spatially dividing the display into major sectors (Thorndyke & Stasz, 1980), using
the organization of available features to do so (e.g., using the Mississippi River to divide the United States in half).

By contrast, students in the current study had the opposite task that of creating a map representing a given set of concepts and the associations between them. Considering the lifelong experiences most people have in using maps of various types (e.g., general reference maps, thematic maps, street maps) it is possible that some processing strategies used for these displays might be transferable to knowledge mapping tasks (Downs, 1981; Wandersee, 1990).

The aforementioned insights could have important implications for how students are instructed on the development of knowledge maps, particularly within a collaborative learning environment. Typically, students are taught a step-by-step procedure for developing knowledge maps that involves four steps 1) identifying concepts, 2) arranging the concepts, 3) linking the concepts, and 4) labeling the links (Jonassen et al., 1993 Lambiotte et al., 1989; West, Farmer, & Wolff, 1991).

An interesting alternative might be to perform MDS analyses on a preliminary set of knowledge maps for identifying the primary dimensions that characterize the entire set of concepts. Students would then use these dimensions to guide schemes for subsequent map development.

Suppose, for example, that MDS analysis suggests that two primary dimensions, Importance (important-unimportant) and Utility (theoretical vs. practical), characterize how students perceive the conceptual relationships for a given set of concepts. During follow-on mapping activities, students would be instructed to depict a concept's relative importance by the size of its node and the concept's relative practicality by how dark its node is shaded. Figure 3 shows a hypothetical map created by applying such guidelines to the map shown in Figure 1. In this illustration, size and value were specifically chosen to represent, respectively, the relative Importance and Utility of a concept since they are the most effective "retinal variables" (Bertin, 1983) for depicting information. Additional dimensions could be represented by the shape, color, and orientation of knowledge map nodes.

![Figure 3. Hypothetical knowledge map (from Figure 1) using the size, value, and occlusion of map nodes to depict predetermined conceptual dimensions.](image-url)
Conceivably, this approach to developing knowledge maps would facilitate learning in two ways. First, it would expand the collaborative contribution of this learning tool beyond that of a group activity whose products are shared. In the proposed use of knowledge mapping, the conceptual understanding of the class as a whole is used to guide the mapmaking activity of its members, either individually or as teams. Further, the single knowledge map generated by the class represents the perceptions of the entire community of learners. A second way that this knowledge mapping scheme may help learners is by providing them with additional variables to consider when creating their knowledge maps. The greater intellectual demand posed by considering one or more dimensions of conceptual relationship might promote greater mindfulness in determining the nature of each concept being mapped. Like the act of labeling links between concepts, the task of depicting conceptual relationships according to one or more descriptive dimensions may encourage deeper processing of the concepts to be learned.

Use of these descriptive dimensions to guide students in the mapmaking process may also enhance the value of the knowledge maps they produce as a means of assessing their learning. While research shows that knowledge maps can be successfully used to assess the knowledge structures of learners, their effectiveness hinges on the degree of guidance students receive on mapping procedures (West et al., 1991).

In this study knowledge maps were analyzed through a tedious process whereby, for each map created, physical distances were measured between all possible pairs of concept nodes. A less time-consuming alternative may be achieved through computer programs that not only permit students to create, arrange, and connect various concept labels on a screen but also calculate distances between concepts and produce a data matrix suitable for MDS analysis. Another approach, suggested by Trochim (1999a), involves sorting file cards (each containing a different concept) into piles according to their conceptual similarity, MDS analysis is then performed on the sort data to create the overall map follower’ by a cluster analysis to identify groupings of map elements.

Investigation of how to maximize the effectiveness of knowledge maps and other knowledge tools should be a potentially fruitful area for future research and development. Computer-based knowledge mapping programs, for example, should do more than simply provide the capability for representing information. Recent work in this area by researchers at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) has examined the use of Web-based computer technology that facilitates real-time construction and scoring of knowledge maps by students working alone and in collaboration with other learners (O’Neil & Klein, 1997; Herl et al., 1999; Osmundson et al., 1999). Besides these efforts, further study is needed to better understand the interaction between knowledge tools and learning environments. Some learning environments may benefit more than others through the use of a specific knowledge tool. On the other hand, knowledge tools yield products (e.g., knowledge maps) that may vary depending on the educational setting in which they are created. It is the challenge of future research then, to identify such tool-treatment interactions where they exist and, ultimately, to prescribe ways for enhancing the positive effects of knowledge tools on learning.

REFERENCE


