The Effects of Representation on Students' Elaborations in Collaborative Inquiry

Daniel D. Suthers, Christopher D. Hundhausen

Laboratory for Interactive Learning Technologies
Department of Information and Computer Sciences
University of Hawai`i at Manoa
{suthers,hundhaus}@hawaii.edu

ABSTRACT

In order to better understand how software design choices may influence students' collaborative learning, we conducted a study of the influence of tools for constructing representations of evidential models on collaborative learning processes and outcomes. Pairs of participants worked with one of three representations (matrix, graph, text) while investigating a complex public health problem. Focusing on students' collaborative investigative processes and post-hoc essays, we present several analyses that assess the impact of representation type on students' elaborations of their emerging knowledge. Our analyses indicate significant impacts on the extent to which students revisit knowledge and the likelihood that they will use that knowledge later.

Keywords
Collaborative representations, representational guidance

INTRODUCTION

External representations have long been a subject of study in the context of learning and problem solving tasks, with research showing that the choice of representation can influence an individual's conception of a problem and hence the ease of finding a solution (e.g., Koedinger, 1991; Kotovsky & Simon, 1990; Larkin & Simon, 1987; Novick & Hmelo, 1994; Zhang, 1997). This line of work has focused on individual performance. Little work has specifically compared the influence of alternate representations on *collaborative* learning processes (but see Baker & Lund, 1997; Guzdial, 1997; and Suthers & Hundhausen, 2001). One might ask whether it is sufficient to extrapolate from what is known about the influence of representations on problem solving by individuals, inferring the effect of representational properties on the group by the aggregation of the effects on the individuals. While we believe that much can be gained from such reasoning, we also believe that external representations play additional roles in group learning situations. Researchers should not only empirically validate their extrapolations from the individual studies in collaborative contexts, but also consider the possibility that the shared use of representations by *distributed* cognitions (Salomon, 1993) might involve additional emergent phenomena.

External representations play at least three roles that are unique to situations in which a group is constructing and manipulating shared representations as part of a constructive activity: (1) initiating negotiations of meaning, (2) serving as a representational proxy for purposes of deixis, and (3) providing a foundation for implicitly shared awareness:

1. When an individual wishes to add to or modify a shared representation, there will be some level of obligation to obtain agreement or permission from one's group members. This obligation will lead to explication and negotiation of representational acts in advance of their commission. This discourse will include negotiations of meaning and shared belief that would not be necessary in the individual case, where one can simply change the representation as one wishes. Thus, the creative acts afforded by a given representational notation may affect which negotiations of meaning and belief take place.

2. The components of a collaboratively constructed representation, having arisen from negotiations just discussed, evoke in the minds of the participants rich meanings beyond that which external observers might be able to discern by inspection of the representations alone. Residing in the shared context of subsequent interaction, these components can serve as an easy way to refer to ideas previously developed, this reference being accomplished by deixis rather than complex verbal descriptions (Clark & Brennan, 1991). In this manner, collaboratively constructed external representations facilitate subsequent negotiations, increasing the conceptual complexity that can be handled in group interactions and facilitating elaboration on previously represented information.

3. The shared representation also serves as a group memory, reminding the participants of previous ideas (encouraging elaboration on them) and possibly serving as an agenda for further work. Individual work also benefits from an external memory, but in the group case there is an additional awareness that one's interlocutors may be reminded by the representation of prior ideas, prompting oneself to consider potential commentary that others will have on one's proposals. That is, it becomes harder to ignore implications of prior ideas if one is implicitly aware that one's interlocutors may also be reminded of them by the representations (Micki Chi, personal communication).

In summary, there is good reason to believe that representational effects will extend to collaborative learning situations in ways worthy of study in their own right.

Further study is needed because these effects may differ between notational systems, and designers of representational tools for collaborative learning need to be informed of the implications of their notational design choices. Representational notations can differ on what information they are capable of expressing (Stenning & Oberlander, 1995), what information they make salient (Larkin & Simon, 1987), and what epistemic processes they suggest (Collins & Ferguson, 1993). We claim that the ways in which a collaboratively constructed representational artifact can play the roles just discussed—initiating negotiations of meaning, representational proxy for purposes of deixis, implicitly shared awareness—is sensitive to the notation's expressiveness and salience of information. Suthers (1999, 2001) has termed this representation-specific influence *representational guidance*. See those publications for further discussion of the origins of representational guidance and a comparison of representations in CSCL systems. See also Toth et al. (in press) for a study of representational guidance in a classroom context.

To explore the ways in which representation impacts group learning, we have conducted an empirical study of the effects of representational tools on students' collaborative discourse and learning outcomes. In these studies, pairs of college science students investigated a problem in the area of public health. They used software based on one of three alternative representational notations (matrix, graph, or text) to compile data, hypotheses, and evidential relations, with the goal of coming to a conclusion about the cause of the problem. In our first analysis of the resulting data (Suthers & Hundhausen, 2001), we considered students' activity and talk surrounding evidential relations, as well as their learning outcomes as measured by a posttest and a post-hoc essay.

In this paper, we present new analyses that explore the influence of representational tool on participants' subsequent elaboration of the data items, hypotheses, and evidential relations that they represent. Elaboration may differ because the notations differ in salience of information (e.g., data and hypotheses are salient in graphs as visual shapes), and in whether they suggest consideration of relationships between new and previously represented information (e.g., the cells of a matrix prompt for consideration of *all* relationships between row and column items). From a pedagogical standpoint, representations that encourage elaboration of previously represented knowledge are beneficial in two important respects. First, they serve as *mediational resources* (Roschelle, 1994), facilitating collaborative interactions in which students elaborate on and refine the structure and content of their knowledge. Second, in encouraging elaboration of students' emerging domain knowledge, representations help students to integrate that knowledge with their existing knowledge, leading to better retention (Craik & Lockhart 1972; Stein & Bransford 1979; Chi *et al.* 1989).

The remainder of this paper is organized as follows. In Sections 2 and 3, we briefly review the design of the study and our prior results. For a more comprehensive description, see Suthers & Hundhausen, (2001). We present our new analyses in Section 4, and summarize our conclusions in Section 5.

STUDY DESIGN

Our study employed a single-factor, between-subjects design with three participant groups defined by the representational software they used: Matrix, Graph, and Text. All three groups were given the identical task of exploring an unsolved "challenge problem"—presented as a series of textual web pages—by recording data, hypotheses, and evidential relations as they encountered them.

We recruited 60 students (32 women, 28 men) in self-selected, same-gender pairs, out of introductory biology, chemistry, physics, and computer science courses at the University of Hawai`i. Participants were all under 25 years of age, and had a mean grade point average of 2.99 (on a 4-point scale). All but three participants were native English speakers. The three non-native speakers were fluent. Participants were paid a $25 honorarium.

Pairs of participants used one of three different versions of software for representing data, hypotheses, and evidential relations. All three versions of the software had two distinct windows (Figure 1). Participants used the right hand
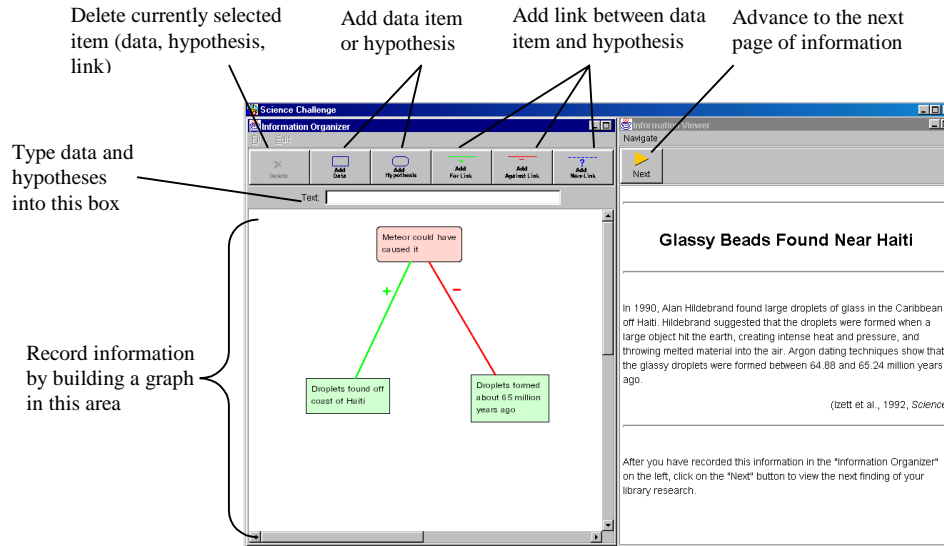
Figure 1. The Graph version of the software

window, identical in all three versions of the software, to move forwards but not backwards through a sequence of 15 pages that presented information relating to the cause of a mysterious neurological disease on the island of Guam. The left-hand window contained a tool for constructing representations of the data, hypotheses, and evidential relations participants gleaned from the information pages on the right. This window varied by condition. The Matrix version contained a spreadsheet-like tool that enabled participants to type in data items along the left-hand column and hypotheses along the top row, and to select evidential relations denoted by "+," "-," or "?" in the corresponding cells. In the Graph version, the left window contained a tool based on Belvedere (Suthers et al, 1997) that enabled one to build a graph of nodes expressing data items and hypotheses, and links labeled "+," "-," or "?" representing evidential relations. The Text version contained a simple word processor into which participants could type data, hypotheses, and evidential relations in any way they wished.

At the beginning of the learning session, participants were given a brief (10-minute) introduction to the software they would be using. So that they could become acquainted with the software and the information-recording process, participants then worked on a warm-up problem (on mass extinctions), which was completely unrelated to the main problem. After 15 minutes, participants were instructed to stop work on the warm-up problem, and to move on to the main problem (on the neurological disease). Participants were given as much time as they needed to explore all 15 pages on the main problem in linear order (one could not go back to previous pages). Following the learning session, participants were given 20 minutes to individually complete a multiple-choice post-test, and 30 minutes to collaboratively write an essay that discussed their hypotheses and the evidence for and against them.

PRIOR RESULTS

In previous analyses (Suthers & Hundhausen, 2001), we focused on participant talk and activities dedicated to evidential relations, as well as participants' learning outcomes. We predicted that participants who construct matrices would talk more about evidential relations than participants who construct graphs, and that both of these groups would talk more about evidential relations than participants who construct plain text documents. This prediction was made because the representation of evidential relations is no more salient than anything else in a textual representation; while graphs represent relations with an explicit object (a link) and carry with them the expectation that one construct such links; and matrices prompt for all possible relationships with empty fields. We also predicted that these process differences would lead to significant differences in learning outcomes. With respect to participant talk and activities, a content analysis revealed significant differences in the extent to which the three treatment groups tended to issues of evidence. Specifically, our analysis found that, as compared to the Graph and Text groups, a significantly higher percentage of the Matrix groups' total on-task activity was dedicated to evidential relations. This result held for both their verbal talk and their representational acts with the software. Although Graph users had higher numerical counts of evidence-focused activity than Text users, there was no significant difference between

these groups. However, these process differences did not translate into learning outcome differences. We found no significant differences between the groups with respect to both post-test scores and the quality and quantity of information discussed in participant essays, although essay scores trended in the predicted direction. The lack of significance of learning outcomes was disappointing but not surprising. The total amount of time spent working with the tool was less than an hour. We speculate that this is not enough time for learning outcomes to develop.

ANALYSES OF ELABORATION

The results reviewed in the previous section furnish evidence for our general hypothesis that the type of representation students' use in collaborative scientific investigations will impact the focus of their discourse. We now turn our attention to an important related question: To what extent do the alternative representations encourage students to *elaborate on* previously represented items? This section presents several analyses that explore this question from different angles. Throughout these analyses, we use the term *elaboration* in the sense of *revisitation*, or subsequent consideration. Specifically, in our session transcripts, we classified as an elaboration any subsequent reference to an item, where a subsequent reference could take any of the following four forms:

- An explicit verbal reference to the item;

- An implicit verbal reference to the item through the item's representational proxy;

- A verbal or representational formulation of, or reference to, an evidential relation that includes the item (in the case of data items and hypotheses); or

- A representational change (e.g., changing an evidential relation from "+" to "-," or changing the wording of an item).

In addition, in order to increase the likelihood that participants' elaboration of an item was prompted by the representation, and not by participants' working memory, we required that there be a reasonable delay between participants' initial representation of the item and their subsequent elaboration of the item. In particular, we counted only elaborations that took place while participants were viewing an information page that followed the page they were viewing when they initially represented the item.

Baseline: Representation of Items

To provide a baseline for our analyses of the impact of representation on elaboration, we begin this section by examining the extent to which participants represented information gleaned from the trail of web pages they encountered during the learning session. Table 1 summarizes the data items, hypotheses, and evidential relations that students in each treatment group represented, measured by both mean counts, and percentages of our *reference items*.

To interpret these data, we need to clarify two important questions: (1) what did we count as an "item?" and (2) what is a "reference item?" The answers to these questions are closely related. As one might have expected, participants chose to represent and relate information in different-sized semantic chunks. For example, upon reading the first information page, one pair created a single data item that read, "Northern Guam is a limestone plateau with high concentrations of calcium in the water." In contrast, another pair divided the same information into three separate data items: (a) "northern Guam," (b) "limestone plateau," (c) "high calcium in water." Clearly, both of these pairs represented the same information. In order to ensure that pairs who chose to divide information into smaller semantic chunks did not get credit for representing more information, we performed the same task as the participants in our study, using a matrix to create a set of 15 data items, 4 hypotheses, and 22 evidential relations that we believe a scientist exploring the materials would have created. These items, which we call *reference items*, served as normalized semantic units for our counts. Thus, in cases in which participants chose to represent smaller fragments of a given reference item, we collapsed all such fragments into a single item. Participants occasionally created items that were not in our set of reference items. (This happened most frequently in the case of evidential relations.) In these instances, we counted *each* such item, regardless of its chunk size.

Table 1. Data items, hypotheses, and evidential relations each treatment group represented in learning sessions, both as mean counts and as mean percentages of the sets of reference items (standard deviations in parentheses).

	Graph		Matrix		Text	
	Count	%	Count	%	Count	%
Data	14.7 (0.5)	98.0 (3.2)	14.8 (0.6)	98.7 (4.2)	14.7 (0.7)	98.0 (4.5)
Hypothesis	3.8 (1.4)	57.5 (16.9)	5.3 (3.0)	72.5 (7.9)	7.2 (2.7)	80.0 (23.0)
Evidential Relations	9.2 (5.1)	25.0 (18.0)	47.5 (40.2)	63.2 (22.8)	15.0 (11.4)	30.9 (20.1)
Total	27.7 (5.6)	54.9 (11.3)	67.6 (41.8)	77.1 (12.4)	36.9 (13.3)	60.2 (12.0)

Turning to the data themselves, we note several trends. While the three groups were identical in terms of number of represented data items (first row), the Text group represented more hypotheses than the other two groups, as reflected in the second row by both the count and percentage of reference hypotheses represented. An analysis of variance (ANOVA) indicates that this difference is statistically significant (df = 2, F = 4.80, p = 0.0165); a post-hoc Tukey test reveals that the difference is between Text and Graph (p < 0.05). The Matrix group represented substantially more evidential relations than the other two groups, as reflected in the third row by both the count and percentage of reference evidential relations represented. This difference, according to an ANOVA, is statistically significant (df = 2, F = 7.21, p = 0.031), with the differences lying between both Matrix and Graph (Tukey test, p < 0.05), and Matrix and Text (Tukey test, p < 0.5). This result echoes our prior results concerning discussion of evidence (Suthers & Hundhausen, 2001). The large difference in number of evidential relations represented translates into a statistically significant difference in total number of items represented (df = 2, F = 6.68, p = 0.0044). Post-hoc Tukey tests show the difference to be between both Matrix and Graph (p <0.05), and Matrix and Text (p < 0.05). Finally, consider the mean counts in relation to our reference items. On average, participants represented 14.7 data items, 98% of our 15 reference items. This indicates that participants are in high agreement with us concerning the 15 data items to be gleaned from the materials. In contrast, Matrix and Text had on average more hypotheses than we did (5.3 and 7.2 compared to 4), and Matrix had far more evidential relations (47.5 compared to 22). Clearly, Matrix users were not as discriminating as we were in creating evidential relations.

Elaboration of Data and Hypotheses in Session

We now turn to our first analysis of elaboration, which considers the extent to which students revisited, within their learning sessions, the data and hypotheses that they initially represented. (Revisitation of non-represented items was negligible.) In accordance with our general hypothesis that representation type affects elaboration, we hypothesized that the Matrix group would revisit represented data and hypotheses more consistently than the Graph group, and that the Graph group would revisit represented data and hypotheses more consistently than the Text group. Our reasoning was that the Matrix representation encourages elaboration of data and hypotheses because it explicitly represents all possible evidential relations between the two (by cells to be filled in), and hence encourages students to reconsider represented data and hypotheses as they explore possible evidential relationships. In contrast, since the Text representation does not explicitly represent evidential relations, we reasoned that it would not prompt students to reconsider the data and hypotheses that they write down. We speculated that the Graph representation would lie somewhere in the middle of these two representations. Graph should encourage elaboration because data and hypothesis statements are reified as visual objects (shapes) arranged on the screen. The salience of these objects was expected to encourage subsequent discussion of the corresponding statements through reminding and ease of deixis. However, revisitations would be less frequent than in Matrix, because although Graph explicitly represents evidential relations by links it does not explicitly represent their absence, so it does not encourage exploration of all possible relationships.

Table 2 presents the mean ratio and percentage of represented data and hypotheses that participants revisited in their learning sessions. (The denominators of the ratios are the sums of the counts of represented data and hypothesis items from Table 1. Relations will be considered in the next section.) As these numbers indicate, there exists a gap between both the Graph and Text groups, and the Matrix and Text groups. A non-parametric Kruskall-Wallis test of the mean percentages indicates that there does indeed exist a statistically significant difference (df = 2, H = 10.21, p = 0.0061), and post-hoc Fischer PLSD tests confirm that the difference lies between both Graph and Text (p < 0.05), and Matrix and Text (p < 0.05). These results confirm our hypothesis that Matrix and Graph are superior to Text for prompting elaboration on represented information.

Table 2. Mean ratios and percentages of represented data items and hypotheses that participants revisited within their learning sessions (standard deviations in parentheses).

	Graph		Matrix		Text	
	Mean Ratio	Mean %	Mean Ratio	Mean %	Mean Ratio	Mean %
revisited / total	13.3 (3.7) / 18.5 (1.4)	71.9 (18.8)	12.2 (4.0) / 20.1 (3.2)	61.6 (21.5)	8.6 (4.3) / 21.9 (2.7)	39.3 (18.8)

Having detected general differences, we now turn to a more detailed analysis of actual reintroduction events. Specific questions to be addressed by this analysis include

- How often do participants actually revisit each revisited item?

- Do they tend to revisit those items fairly recently after they represent them, or much later in the session, perhaps as their relevance becomes evident to a discussion?

We answered these questions by examining logs of revisitation events indexed by (a) the number in sequence of the information page that was visible when the event occurred (there were 15 total information pages); and (b) the number of the segment in which the event occurred. A *segment* is a verbal utterance or a representational change that expresses a single thought or idea (see Suthers & Hundhausen (2001) for details of coding). Summary data from these logs are presented in Table 3. A "span" is the number of pages or segments from the initial representation of the item to the revisitation event.

Table 3. Mean number of revisitations per data item/hypotheses, and the mean page and segment spans per revisitation

	Graph	Matrix	Text
Mean # revisitations per item	1.7 (0.3)	4.7 (4.2)	2.7 (1.5)
Mean page span per revisitation	5.2 (1.7)	6.0 (1.3)	4.9 (0.6)
Mean segment span per revisitation	275.0 (203.7)	326.2 (114.2)	224.4 (83.8)

A nonparametric Kruskall-Wallis test did not yield significant differences between the groups with respect to the mean number of revisitations per item (df = 2, H = 5.23, p = 0.0732), although a trend suggests that, while the Graph pairs revisited slightly more data and hypotheses than Matrix pairs (Table 2), they did not revisit those items as often as did Matrix pairs. With respect to the average page and segment span of each revisitation, a non-parametric Kruskall-Wallis test detects no significant differences (page span: df = 2, H = 4.51, p = 0.1047; segment span: df = 2, F = 3.42, p = 0.1805). We speculate that this lack of difference indicates that the sequencing of information in the pages, which dictates opportunities for elaboration, has more impact on the timing of revisitations than does representation type.

Elaboration of Evidential Relations in Session
We now consider the evidential relations that students represented and revisited in their learning sessions. Table 1 showed a significant difference in the percentage of reference relations represented, with Matrix representing more. However, the other groups may not have represented some of these reference relations because the data and hypotheses to be represented were not available. To rule out this explanation, we compared the extent to which participants actually represented relevant evidential relations upon representing the corresponding data item and hypothesis to be related. We focused on our set of 22 reference evidential relations because these are the only relations that we can reasonably expect participants to represent. Table 4 lists the mean percentage of those reference evidential relations for which both relata were available that were filled in by participants across treatment groups. An ANOVA of these percentages (we can apply an ANOVA because the denominator is fixed) detects a statistically significant difference (df = 2, F = 8.98, p = 0.001). Post hoc Tukey tests show the difference is between Matrix and Graph (p < 0.05), and Matrix and Text (p < 0.05). These results confirm our reasoning that Matrix users filled in significantly more evidential relations because of a property of the notation: the empty cells created when one represents a new data or hypothesis in the Matrix prompt users to fill in the available evidential relations. In the other two representations available evidential relations are not as obvious, so one is less likely to tend to them.

Table 4. Mean percentage of available reference evidential relations that were represented. By *available*, we mean evidential relations whose data and hypothesis components have already been represented.

Graph	Matrix	Text
33.2 (21.8)	72.5 (25.8)	34.2 (23.6)

We now consider participants' revisitation of previously represented evidential relations. Table 5 presents the mean ratio and percentage of revisited evidential relations. A comparison to Table 2 indicates that subsequent elaboration of evidential relations was much more rare than elaboration of data and hypotheses, However, a non-parametric Kruskall-Wallis test of the groups' mean percentage of revisited evidential relations yields a significant difference between the groups (df = 2, F = 6.85, p = 0.0325). A post-hoc Fischer PLSD test shows the difference to be between Matrix and Graph (p < 0.05).

Table 5. Mean ratio and percentage of represented evidential relations that participants revisited

	Graph		Matrix		Text	
	Mean Ratio	Mean %	Mean Ratio	Mean %	Mean Ratio	Mean %
revisited / total	0.2 (0.4) / 9.2 (5.1)	2.1 (4.4)	7.3 (9.2) / 47.5 (40.2)	14.8 (20.4)	0.8 (1.1) / 15.0 (11.4)	5.0 (7.3)

To explain the fact that participants revisited evidential relations less frequently than they revisited data and hypotheses results, we observe that evidential relations are already a syntheses of the domain information that participants encountered. Indeed, representing an evidential relation constitutes a more reflective activity than representing a data or hypothesis. We thus speculate that students tend not to see evidential relations as items that warrant further reflection. This is not to say that such reflection would not be valuable. For example, students might reflect on the warrants behind their inferences. However, getting students to reflect further on evidential relations appears to be a challenge for designers of collaborative representations.

We have two explanations for the difference in revisitations between Matrix and Graph. First, this difference may actually be symptomatic of a problem with the Matrix representation. While 46% of the revisitations of relations in Matrix were changes to the type of relation, there was only one change event in all of the Graph sessions and none in Text. We believe that Matrix users felt compelled to modify their relations much more often than other participants because they were prompted by the cells to invent relationships between items that were not particularly relevant to each other (as well as items that were). The video data includes many examples of participants changing each relationship several times while they attempted to resolve the ambiguity.

A second explanation requires understanding relevant details of the software tools. In the Graph tool, one creates a new relation by selecting the appropriate relation's icon ("+", "-", or "?"), then selecting the statements that form the start point and endpoint of the link in turn. The method of changing the type of an existing link is entirely different: one must either right-click to obtain a link editor, or delete and then recreate the link. In contrast, the method of changing a relation in Matrix is identical to the method of creating it in the first place: one selects the cell of the matrix to obtain a menu of options. We speculate that there would be more revisitations in Graph if the method of modifying the relation became obvious while creating it. This discussion illustrates the importance of considering one's instructional objectives even in the design of micro-level human-computer interactions.

Elaboration of Session Items in Essays
Our final analysis of elaboration considers the extent to which participants included represented items in their post hoc essays, which they wrote roughly 25 minutes after the learning session. In a sense, this is an analysis of *retention*: Do participants tend to remember and integrate into their own findings those items that they represented during the learning session? Consider this question with respect to two hypotheses. First, the null hypothesis is that there will be no relationship between representation and essay contents, and therefore no content differences between essays. Second, we might hypothesize that there is a relationship, but that it is independent of the particular representation being used: representing an item increases the likelihood that it will be remembered and included in the essay regardless of the representation. If this were indeed the case, then we would expect the Text group to include significantly more hypotheses in their essays than the Graph group, and the Matrix group to include significantly more evidential relations in their essays than the other two groups, because this is the pattern of representational counts found in Table 1.

Departures from this pattern may indicate influences of the representations on retention for reasons other than the mere fact that the items were represented.

Table 6 presents the mean number of data, hypotheses, and evidential relations participants included in their essays. ANOVAs suggest that half of the prediction of the second hypothesis holds: The Text group included significantly more hypotheses than the Graph group (df = 2, F = 4.79, p = 0.0166; Tukey test: p < 0.5); however, there exist no differences between the groups with respect to number of evidential relations included in their essays (df = 2, F = 0.19, p = 0.8318).

Table 6. Mean number of items included in essays

	Graph	Matrix	Text
Data	9.8 (3.2)	10.6 (3.0)	10.5 (3.4)
Hypothesis	3.7 (1.3)	4.8 (1.1)	5.3 (1.1)
Evidential Relations	10.1 (5.9)	11.2 (4.1)	9.9 (5.3)
Total	23.6 (9.8)	26.6 (6.7)	25.7 (7.9)

Thus, more items represented in the session did not necessarily translate into more items discussed in the essay. This result admits the possibility that there may be group differences with respect to the percentage of "carryover" items: those data items, hypotheses, and evidential relations that were represented in the session and subsequently included in the essay. To test this possibility, we computed each group's percentage of represented-in-session items that were also included in the essay (see Table 7). Inspecting these percentages, we find that the Graph condition had a higher percentage of carryover items than both Matrix and Text. A non-parametric Kruskall-Wallis test indicates that this difference is statistically significant (df = 2, H = 6.48, p = 0.0391). A post-hoc Fischer PLSD test shows that the difference is between Graph and Matrix (p < 0.05). This difference lies primarily in the use of evidential relations (third row of Table 7).

Table 7. Mean percentage of represented items included in essays

	Graph	Matrix	Text
Data	63.1 (16.0)	66.2 (22.1)	64.3 (20.2)
Hypotheses	71.8 (25.1)	80.4 (28.5)	56.1 (14.7)
Evidential Relations	36.4 (33.1)	20.9 (23.0)	35.4 (29.9)
Total	55.4 (17.6)	36.2 (21.0)	48.9 (16.3)

In interpreting this result, note that Graph users were more focused with respect to what they represented in their learning sessions. Table 1 tells us that they were more selective than users of other representations in both the hypotheses and the evidential relations that they represented. Graphs prompt users to identify and represent *some* relationship involving each new item, but does not specify which relationship, and (unlike Matrix) does not encourage representation of all possible relationships. Thus pairs are faced with the need to discuss which relationship to represent, so they engage in a discussion of the possible relationships and their significance. We therefore speculate that Graph pairs are encouraged to engage in higher-order thinking when faced with the choice of how to connect a newly added item. In contrast, Text users (who had the most hypotheses in both the session representations and the essays, yet the least overlap between the two) were less discriminating in the hypotheses they represented, and were not prompted to evaluate these hypotheses in any particular way, so apparently reinvented hypotheses as they wrote their essays. (There is a nonsignificant trend for hypotheses in Table 7 according to a Kruskall-Wallis test, df = 2, H = 5.27, p = 0.0716; 46% of the hypotheses in Text essays were new.) Matrix pairs may have filled in cells (47.5 cells on average) without being very discriminating of which relations are important, but then selected a smaller set of relations (11.2 on average) while writing their essays. This interpretation is corroborated by our coder's informal observation that some Matrix groups filled in the cells late in the session by systematically going down columns or across rows with minimal discussion, while Graph users usually linked items as they went, discussing each link.

SUMMARY AND CONCLUSIONS

Prior analysis of our data (Suthers & Hundhausen, 2001) showed that the exhaustive prompting of Matrix for consideration of all possible evidential relations leads participants to discuss issues of evidence more than users of other representations. The present analysis added the following results:

- Graph users represent the fewest items. Text and Matrix users represent more hypotheses than we derived from the materials and Matrix users represent far more evidential relations than were in our analysis.

- Users of visually structured representations (Graph, Matrix) revisit previously discussed ideas more often than users of Text. Matrix users revisit prior data and hypotheses mainly to fill in the matrix cells that relate them.

- Revisitation of relations is rare except for Matrix users, who often modify their relations.

- The representational work done by Graph users has a greater impact on the content of their essays than the representational work done by users of Text or Matrix.

We draw several general conclusions from these results. The choice of representational notation for collaborative learning applications does matter. Representational notations can have significant effects on learner's interactions, and may differ in their influence on subsequent collaborative use of the knowledge being manipulated. Specifically, visually structured and constrained representations can provide guidance for collaborative learning that is not afforded by plain text. However, not all guidance is equal, and more is not necessarily better. For example, it is possible to over-prompt for consideration of irrelevant relationships. Whether the increased talk about evidence prompted by Matrix is valuable is a pedagogical decision that must be carefully considered in light of the possibility that many of the evidential relationships considered may be irrelevant. A representation such as Graph may guide students to consider evidence without making them unfocused.

We believe that each representation has its own strengths and weaknesses, and each may be the best choice for different cognitive tasks, learning objectives, and populations. In fact, our current version of Belvedere integrates three representational "views" (Graph, Matrix, and a Hierarchy representation not discussed here) of evidence models in one tool, providing an interesting platform for future studies. We speculate that Graph will be most useful for gathering and relating information by the relationships that motivated its inclusion; Matrix for subsequently checking that no important relationships have been missed and for scanning for patterns of evidence; and Hierarchy for performing selective queries on a complex evidential model.

There is of course a great deal of future work suggested by the studies reported here, ranging from further analysis of existing data to new studies. The analyses presented here only assess the extent to which students revisited represented items; they say nothing about the quality or depth of the exchanges in which items were revisited. What we really want to know is whether students are deeply reflecting on domain concepts and relationships. Ongoing work is analyzing when relations were created and quality of the negotiations leading to each represented relation. An argumentation analysis of students' essays is also underway to determine whether there are any structural differences between the groups' essays. Subsequent analysis of our data will shift from comparison of group means to analysis of individual events, specifically to better understand how the different representations are appropriated as resources in support of collaborative discourse. Future studies currently being planned include attempts to replicate our results in distance learning contexts, with particular attention paid to the designed *integration* of discourse representations (chat and threaded discussion) with visual knowledge representations (Hoadley & Enyedy, 1999). We also plan to work with teachers in developing strategies for use of our multi-representational version of Belvedere.

REFERENCES

Baker, M. & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning, 13:* 175-193.

Chi, M. & Bassok, J. Lewis, M. Reimann, P. Glaser, R. (1989). Learning from examples via self-explanations. In *L. Resnick (Ed.) Knowing, learning and instruction: Essays in Honor of Robert Glaser,* pp. 251-282. Hillsdale, NJ: Lawrence Erlbaum Associates

Clark, H.H. & Brennan, S.E. (1991). Grounding in Communication. In L.B. Resnick, J.M. Levine and S.D. Teasley (eds.), *Perspectives on Socially Shared Cognition,* American Psychological Association, pp. 127-149.

Collins, A. & Ferguson, W. (1993). Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist, 28*(1): 25-42.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*: 671-684.

Guzdial, M. (1997, December). Information ecology of collaborations in educational settings: Influence of tool. In *Proc. 2nd International Conference on Computer Supported Collaborative Learning (CSCL'97) (pp. 91-100).* Toronto.

Hoadley, C. & Enyedy, N. (1999). Between Information and Communication: Middle Spaces in Computer Media for Learning. In *Proc. Computer Support for Collaborative Learning (CSCL) 1999 Conference,* C. Hoadley & J. Roschelle (Eds.) Dec. 12-15, Stanford University, Palo Alto, California. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 242-251.

Koedinger, K. (1991). On the design of novel notations and actions to facilitate thinking and learning. *Proc. Int. Conference on the Learning Sciences,* pp. 266-273. Charlottesville, VA: Association for the Advancement of Computing in Education.

Kotovsky, K. and H. A. Simon (1990). What makes some problems really hard: Explorations in the problem space of difficulty. *Cognitive Psychology, 22*: 143-183.

Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science 11*(1)*:* 65-99. 1987.

Novick, L. R. & Hmelo, C. E. (1994). Transferring symbolic representations across nonisomorphic problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6): 1296-1321.

Roschelle, J. (1994, May). Designing for cognitive communication: Epistemic fidelity or mediating collaborative inquiry? *The Arachnet Electronic Journal of Virtual Culture,2(2).*

Salomon, G. (1993). Distributed cognitions: psychological and educational considerations. Edited by G. Salomon. Cambridge, England; New York, NY: Cambridge University Press, 1993.

Stein, B. S., & Bransford, J. D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning and Verbal Behavior, 18*:769-777.

Stenning, K. & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science, 19*(1): 97-140. 1995.

Suthers, D. D. (1999, January). Representational support for collaborative inquiry. In *Proc 32nd Hawai`i International Conference on the System Sciences (HICSS-32).* (CD-ROM). Maui, Hawai`i: Institute of Electrical and Electronics Engineers, Inc. (IEEE). Available: http://lilt.ics.hawaii.edu/lilt/papers/hicss99.pdf.

Suthers, D. D. (2001). Towards a Systematic Study of Representational Guidance for Collaborative Learning Discourse. *Journal of Universal Computer Science* 7(3), 2001. Electronic publication: http://www.jucs.org/jucs_7_3/towards_a_systematic_study

Suthers, D. & Hundhausen, C. (2001, March). Learning by Constructing Collaborative Representations: An Empirical Comparison of Three Alternatives. In P. Dillenbourg, A. Eurelings, K. Hakkarainen (Eds.) *European Perspectives on Computer-Supported Collaborative Learning, Proc. First European Conference on Computer-Supported Collaborative Learning* (pp. 577-584). Universiteit Maastricht, Maastrict, the Netherlands.

Suthers, D., Toth, E., and Weiner, A. (1997, December). An integrated approach to implementing collaborative inquiry in the classroom. In *Proc. 2nd International Conference on Computer Supported Collaborative Learning (CSCL'97)* (pp. 272-279). Toronto.

Toth, E., Suthers, D. & Lesgold, A. (in press). Mapping to know: The effects of representational guidance and reflective assessment on scientific inquiry skills. Accepted for publication in the journal *Science Education.*

Zhang, J. (1997). The nature of external representations in problem solving. Cognitive Science, 21(2): 179-217.