Process and Outcome Studies of Representational Guidance for Scientific Inquiry

Daniel D. Suthers
Department of Information and Computer Sciences
University of Hawai'i at Manoa
1680 East West Road, POST 303A
1.808-956.3890
suthers@hawaii.edu

Eva Erdosne Toth
Department of Psychology
Carnegie Mellon University
etoth@andrew.cmu.edu

Abstract. External (visual and textual) representations differ in the information they are capable of capturing, in the ease of recovering information that is captured, and in the ease of identifying information that is missing. Consequently, representations may differ in how they guide learning processes. We call these differences between representations "representational guidance." Effects of representational guidance on individual problem solving are well documented, yet there is a lack of comparative research on how representational effects extend to collaborative learning situations. This paper introduces the concept of representational guidance for collaborative learning discourse and summarized the results of two studies: a laboratory study investigating differences in discourse processes, and a classroom study investigating differences in products of the investigations. The laboratory study compared three representational tools for recording and reasoning about evidence during collaborative inquiry. The tools were unconstrained text (a word processor), graphical evidence mapping and a matrix or tabular representation. Participants were provided with hypertext-based materials pertinent to a problem in which multiple explanations were proposed for a disease. Analysis of discourse transcripts was consistent with predictions concerning the amount of talk about evidential relations. The classroom study focused on the effects of representational tools and of assessment rubrics on learning to evaluate diverse sources of empirical evidence and coordinate these data with domain theories. The representational tools were graphical evidence maps and text. The assessment rubrics provided specific criteria for scientific inquiry, encouraging self-reflection and peer coaching. The results indicate that evidence mapping provides better support than prose, as evidenced by amount of information and quality of inferences recorded in classroom artifacts. The assessment rubrics improved the quality of work of students using the evidence mapping tool but not those using prose.

Introduction

From early on in the history of science, "representing" (i.e., formulating various depictions of reality) and "knowing" (i.e., the cognitive processing of information) have been closely intertwined (Vandersee, 1990; Novak & Gowin, 1984). This is because external (visual and textual) representations differ in the information they are capable of capturing, in the ease of recovering information that is captured, and in the ease of identifying information that is missing (Larkin & Simon, 1987; Stenning & Oberlander, 1995). We call these differences between representations "representational guidance" because they constitute ways in which representations may differ in how they guide learning processes (Collins & Ferguson, 1993; Lajoie & Derry, 1993). Effects of representational guidance on individual problem solving and learning are well documented (e.g., Kotovsky & Simon, 1990; Larkin & Simon, 1987; Novak, 1990; Novick & Hmelo, 1994; Zhang, 1997). Yet there is a lack of comparative research on how representational effects extend to collaborative learning situations, in spite of the fact that a number of different representations are used to support collaborative inquiry (e.g., Bell, 1997; Guzdial et al., 1997; Hewitt & Scardamalia, 1998; O'Neill & Gomez, 1994; Suthers et al., 1997). External representations serve as resources for conversation (Roschelle, 1994), so we expect that differences in resources will translate into differences in discourse and learning.

This paper introduces the concept of representational guidance for collaborative learning discourse and summarizes the results of two studies: a laboratory study investigating differences in discourse processes, and a classroom study investigating differences in products of the investigations. This work was undertaken with a diagrammatic software environment ("Belvedere") intended to support secondary school children's learning of critical inquiry skills in the context of science. The diagrams were first designed to capture scientific argumentation, and later simplified to focus on evidential relations between data and hypotheses. This change was driven in part by a refocus on *collaborative* learning, which led to a major change in how we viewed the role of the interface representations. Rather than viewing the representations as medium of communication or a formal record of the argumentation process, we came to view them stimuli and guides for conversation.

Representational Guidance

The underlying hypothesis of this work is that variation in features of representational tools used by learners working in small groups can have a significant effect on the learners' knowledge-building discourse and on learning outcomes (Suthers, 1999). Representational *notations* provide a set of primitive elements out of which representations can be constructed. This work is concerned with symbolic as opposed to analogical notations. Representational *tools* are implementations of representational notations, including software interfaces. The users of the tool construct, examine, and manipulate representational *artifacts* in the tool.

Each given representational notation manifests a particular representational guidance**,** expressing certain aspects of one's knowledge better than others do. The concept of representational guidance is borrowed from artificial intelligence, where it is called *representational bias* (Utgoff 1986). The phrase *guidance* is adopted here to avoid the negative connotation of *bias*. The phrase *knowledge unit* will be used to refer generically to aspects of one's knowledge that one might wish to represent, such as hypotheses, statements of fact, concepts, relationships, rules, etc. The use of this phrase does not signify a commitment to the view that knowledge intrinsically consists of "units," but rather that users of a representational system may choose to denote some aspect of their thinking with a representational proxy. Representational guidance manifests in two major ways:

- *Constraints*: limits on expressiveness, e.g., the representational system may provide a limited ontology of objects and relations (Stenning and Oberlander 1995).
- *Salience:* how the representation facilitates processing of certain knowledge units, possibly at the expense of others (Larkin and Simon 1987).

Thesis

The core idea may now be stated as follows: Representational tools mediate collaborative learning interactions by providing learners with the means to express their emerging knowledge in a persistent medium, inspectable by all participants, where the knowledge then becomes part of the shared context. Representational guidance constrains which knowledge can be expressed in the shared context, and makes some of that knowledge more salient and hence a likely topic of discussion. Furthermore, even before a representation is created, the possibilities afforded by a representation may influence negotiations between learners.

External Representations in Collaborative Contexts

Substantial research has been conducted concerning the role of external representations (as opposed to mental representations) in individual problem solving. The kind of external representation used to depict a problem may determine the ease with which the problem is solved (Kotovsky and Simon, 1990; Larkin and Simon, 1987; Novick and Hmelo, 1994; Zhang, 1997). The constraints built into representations may restrict the problem solver's search space, to the possible detriment *or* enhancement of problem-solving success (Amarel, 1968; Hayes, 1989; Klahr and Robinson, 1981; Stenning and Oberlander 1995). One might ask whether this research is sufficient to predict the effects of representations in collaborative learning.

A related but distinct line of work should be undertaken in collaborative learning contexts for several reasons. The interaction of the cognitive processes of several agents differs from the reasoning of a single agent (Okada and Simon, 1997; Perkins, 1993; Salomon, 1993), and therefore may be affected by external representations in different ways. In particular, shared external representations can be used to coordinate distributed work, and will serve this function different ways according to their representational guidance. The act of constructing a shared representation may lead to negotiations of meaning that may not occur in the individual case. Also, the mere presence of representations in a shared context with collaborating agents may change each individual's cognitive processes. One person can ignore discrepancies between thought and external representations, but an individual working in a group must constantly refer back to the shared external representation while coordinating activities with others (Micki Chi, personal communication). Thus it is conceivable that external representations have a greater effect on individual cognition in a social context than they do when working alone. Finally, prior work on the role of external representations in individual problem solving has often used well-defined problems. Further study is needed on ill structured, open-ended problems such as those typical of scientific inquiry.

Two studies, reported in detail elsewhere but summarized here, provide initial evidence of representational guidance for collaborative inquiry. These are a laboratory study and a classroom study.

Comparing Three Representations in a Laboratory Setting

With the assistance of Christopher Hundhausen and Laura Girardeau, the first author conducted a study comparing three alternative notations for recording evidential relationships between data and hypotheses: *free text*, *graphs*, and *matrices,* with respect to participants' amount of talk about evidential relations (Suthers & Hundhausen, 2001).

We tested two specific hypotheses regarding the effects of three alternative representational environments on participants' collaborative discourse and learning outcomes. Our first hypothesis 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. Our second hypothesis predicted that these process differences would lead to significant differences in learning outcomes, with those who construct matrices remembering more data, hypotheses, and evidential relations than those who construct graphs, and those who construct graphs remembering more data, hypotheses, and evidential relations than those who construct plain text documents. This prediction was made because those representations that prompt for increased consideration of evidential relations are in effect prompting students to elaborate on the information being considered. This elaboration in turn should lead to increased memory for the information.

We employed a single-factor, between-subjects design with three participant groups defined by the software they used. Dependent measures included: (a) the percentage of utterances and participant actions in the software focused on evidential relations; (b) ability to recall the data, hypotheses, and evidential relations explored in a multiple-choice test; and (c) a written essay describing the data, hypotheses, and evidential relations in the problem.

Sixty students were recruited out of introductory science courses in self-selected, same gender pairs, Participant pairs were randomly assigned to the three treatment groups such that there were no significant differences between the groups' mean G.P.A. Participant pairs worked with software with two main windows, one containing a workspace for creating either text, graph, or matrix representations, and the other presenting a science problem (what causes a mysterious neurological disease in Guam?) as a fixed sequence of 15 information pages. (Initially participants worked on a shorter "warm-up" problem concerning mass extinctions.) Participants were instructed to visit each page in the sequence, and to record data, hypotheses, and evidential relations. Once finished, they were individually given a post-test, and then asked to work together on an essay summarizing their findings.

All 30 sessions were videotaped and transcribed, including both verbal utterances and actions performed with the software. Transcript segments were coded on the following dimensions (among others): Off Task vs. On Task, and Evidential Relations vs. other content categories. Evidential Relations were further subdivided into Representational (represented with the software) vs. Verbal (spoken). Two independent analysts achieved 89% overall agreement (0.86 kappa) in coding 20% of the sessions using these categories. A single analyst coded the remaining 80% of the sessions. Essays were scored according to the strength and inferential difficulty of the evidential relations they cited, as determined by a systematic expert analysis.

Using a non-parametric Kruskall-Wallis test, we found significant differences with respect to overall percentages of evidential relations segments ($df = 2$, $H = 8.712$, $p < 0.013$), and with respect to the percentages of verbal evidential relations segments ($df = 2$, $H = 12.56$, $p < 0.0019$). A post-hoc Fischer PLSD test determined that, in both cases, the significant differences were between Matrix and Graph ($p < 0.05$), and between Matrix and Text ($p < 0.05$). These results confirm our prediction that notation significantly impacts learners' discussion of evidential relations, although the predicted effect of Graph versus Text was not significant.

A more recent analysis addressed the question of whether the representations differ in the extent to which learners will elaborate on information they previously represented. We recorded when each of a canonical set of information items (data, hypotheses, and evidential relations found in the problem materials) was represented in a group's artifact (if at all), and identified all subsequent reintroductions of these items as a topic of discussion. Matrix showed significantly more reintroduction of all types of information (data, hypotheses, and evidential relations).

With respect to learning outcomes, analyses of variance found no significant differences between the groups' post-test scores ($df = 2$, $F = 0.046$, $p < 0.96$) and essay scores ($df = 2$, $F = 0.74$, $p < 0.49$), although trends for the essay scores were in the predicted direction. These results were disappointing, but not surprising. Participants spent less than an hour on task, and this may not have been enough time for learning outcomes to develop fully.

We further examined the essays to determine the extent to which essay content reflected the work done with the representations. There were no overall differences between representational notations on the percentage of essay content found in the representational artifacts. However, this overlap decreased as the inference required increased: 93% of data items, 70% of hypotheses, and 42% of evidential relations found in the essays were previously recorded in the representational artifacts. This suggests that learners are doing additional inferential work when they write the essays, which were written without the artifacts present.

Guidance for Inquiry in a Classroom Setting

Working with Arlene Weiner, the authors developed a comprehensive method for implementing collaborative inquiry in the classroom supported by the Belvedere software. The approach includes student activity plans worked out in collaboration with teachers. Students work in teams to investigate "science challenge problems," (http://lilt.ics.hawaii.edu/belvedere/) designed with attention to National Science Education Standards to match and enrich the curriculum. A science challenge problem presents a phenomenon to be explained, along with indices to relevant resources. The teams plan their investigation, perform hands-on experiments, analyze their results, and report their conclusions to others. Investigator roles are rotated between hands-on experiments, tabletop data analysis, and computer-based literature review and use of simulations and analytic tools as well as Belvedere. Assessment rubrics are given to the students at the beginning of their project as criteria to guide their activities. The rubrics guide peer review, and help the teacher assess nontraditional learning objectives. See Suthers et al. (1997) for further information on this integrated approach to classroom implementation.

We conducted a classroom study comparing two forms of guidance for inquiry with respect to quality of inquiry process and conclusions (Toth et al. 2001). The forms of guidance included Belvedere's graphical representations of evidential relations, and assessment rubrics. Version 2.1 of Belvedere was used. The assessment rubrics were paper-based charts that included detailed criteria, in Likert-scale format, for progress in each of four components of scientific inquiry: data collection, evaluation of information collected, quality of reports, and quality of peer presentations. The rubrics were provided to students at the outset of the study with explicit instructions concerning their use during the activity to guide inquiry. A 2x2 design crossed Graph (Belvedere) versus Text (Microsoft Word) conditions with Rubric versus No-rubric conditions across four 9^{th} grade science classes. Students spent about 2 weeks on each of three science challenge problems, including problem introduction, investigation, preparation of reports, and presentations to peers.

The data analysis was based primarily on artifacts produced by groups of students, namely their Belvedere graphs or Word documents, and their final report essays. (These studies were undertaken at an overseas location, in Department of Defense Dependent Schools. Hence we were unable to observe classroom activity.) The amount of information recorded did not differ significantly between groups. Significant results were obtained on the *categorization of information* and the *number of evidential relationships* recorded. Specifically, a factorial ANOVA indicated that the Graph groups recorded significantly more inferences than the Text groups (F (1, 18) = 27.3; p = .0001), and the Rubrics users recorded significantly more inferences than those groups who did not use the rubrics for explicit reflection (F (1, 18) = 7.2; p = .01). An interaction between the type of representational tool and the use of rubrics (F (1, 18) = 3.3; p = .088) prompted a closer look at differences in mean scores. A post-hoc paired comparison of the four treatment groups' performance using Tukey's HSD indicated that the combination of graphing and rubrics resulted in a larger number of inferences formulated and recorded compared to all other conditions; while the use of either graphing or rubrics alone did not result in a significantly higher performance compared to either of text groups. Further analysis showed that this interaction was primarily due to the Graph/Rubrics students having recorded significantly more inconsistency relations. These relations represent negative evidence; i.e., there seems to be less of a "confirmation bias." Thus there appears to be a synergistic effect between effective representations and guidelines for their use, particularly with respect to attending to discrepant evidence. The combination of rubrics encouraging students to look for and record disconfirming as well as confirming information and explicit representational devices for recording such inferences.

Summary

The first study compared three representational tools for recording and reasoning about evidence during collaborative inquiry in a laboratory setting. The tools were unconstrained text (a word processor), graphical evidence mapping and a matrix or tabular representation. Based on the representational features of each tool, the first author predicted that collaborating users of these respective tools would differ in how much they talked about evidential relations. Analysis of discourse transcripts was consistent with predictions. These studies were too short to obtain outcome differences. A longer term classroom study investigated the effects of representational tools and of assessment rubrics on learning to evaluate diverse sources of empirical evidence and coordinate these data with domain theories. The representational tools were graphical evidence maps (Belvedere) and text (MS Word). The results indicate that evidence mapping provides better support than prose, as evidenced by amount of information and quality of inferences recorded in classroom artifacts. The assessment rubrics improved the quality of work of students using the evidence mapping tool but not those using prose. The combined use of evidence maps and assessment rubrics appear to complement each other and provide students with an integrated cognitive and social support for inquiry.

Taken together, these studies demonstrate representational guidance with very different measures and settings. Yet many questions remain. Although we observed content differences between learner-constructed Graph and Text artifacts, we only observed significant process differences with Matrix. (but not Graph). We have not yet attempted to measure the quality of interaction between students beyond simple counting of discussion of evidence and reintroduction of previously represented information. For example, is it possible that the Matrix users are not seeing the forest for the trees, i.e., getting lost in the details? How exactly do learners utilize the representations as "resources for conversation"? Most importantly, we need to explore how to translate the observed process differences into learning outcome differences.

We believe that the comparative study of representational support for collaborative learning discourse is a critical area for continued and future study, because interest in the use of electronic media for learning is increasing. We have demonstrated effects worthy of study in proximal (face to face) learning situations, and expect to bring this work to distance and asynchronous computer mediated collaborative learning in the future.

References

Amarel, S.: "On representations of problems of reasoning about actions"; In Donald Michie (Ed.): "Machine Intelligence 3"; Amsterdam, Elsevier/North-Holland (1968), 131-171.

Bell, P. (1997, December). Using argument representations to make thinking visible for individuals and groups. In Proceedings of the Computer Supported Collaborative Learning Conference '97, 10-19. University of Toronto.

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

Guzdial, M., Hmelo, C., Hubscher, R., Nagel, K., Newstetter, W., Puntambekar, S., Shabo, A., Turns, J., & Kolodner, J. L. (1997, December). Integrating and guiding collaboration: Lessons learned in computer-supported collaborative learning research at Georgia Tech. In Proceedings of the 2nd International Conference on Computer Supported Collaborative Learning (CSCL'97) (pp. 91-100). Toronto: University of Toronto.

Hayes, J. R.: "The complete problem solver"; Lawrence Erlbaum Associates, Hillsdale, NJ (1989).

Hewitt, J., & Scardamalia, M. (1998). Design principles for the support of distributed processes. Educational Psychology Review, 10(1), 75-96.

Klahr, D. and Robinson, M.: "Formal assessment of problem-solving and planning processing in school children"; Cognitive Psychology 13 (1981), 113-148.

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.

Lajoie, S. P., & Derry, S. J. (Eds.). (1993). Computers as cognitive tools. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Novak, J.D. (1990). Concept mapping: A useful tool for science education. Journal of Research in Science Teaching 27(10), 937-49.

Novak, J.D. & Gowin, D.B. (1984). Learning how to learn. New York: Cambridge University Press.

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.

Okada, T. & Simon, H. A.: "Collaborative discovery in a scientific domain"; Cognitive Science 21, 2 (1997), 109-146.

O'Neill, D. K., & Gomez, L. M. (1994). The collaboratory notebook: A distributed knowledge-building environment for project-enhanced learning. In Proceedings of Ed-Media '94, Charlottesville, VA: Association for the Advancement of Computing in Education.

Perkins, D.N.: "Person-plus: A distributed view of thinking and learning"; In G. Salomon (Ed.): "Distributed Cognitions: Psychological and Educational Considerations"; Cambridge University Press. Cambridge (1993), 88-111.

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

Salomon, G.: "No distribution without individuals' cognition: A dynamic interactional view"; In G. Salomon (Ed): "Distributed Cognitions: Psychological and Educational Considerations"; Cambridge University Press, Cambridge (1993), 111-138.

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

Suthers, D. D. (1999, January). Representational support for collaborative inquiry. In Proceedings of the 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. & Hundhausen, C. (2001). Learning by Constructing Collaborative Representations: An Empirical Comparison of Three Alternatives. European Conference on Computer-Supported Collaborative Learning, Maastrict, the Netherlands, March 2001, 577-584.

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

Toth, E., Suthers, D., & Lesgold, A. (2001, to appear). Mapping to know: The effects of evidence maps and reflective assessment on scientific inquiry skills. Science Education.

Utgoff, P.: "Shift of bias for inductive concept learning"; In R. Michalski, J. Carbonell, T. Mitchell (Eds.): "Machine Learning: An Artificial Intelligence Approach, Volume II"; Morgan Kaufmann, Los Altos CA (1986), 107-148.

Vandersee, H. J. (1990). Concept mapping and the cartography of cognition. Journal of Research on Science Teaching, 27(10), 923-936.

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