Representations for Scaffolding Collaborative Inquiry on III-Structured Problems Dan Suthers

Learning Research & Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260
suthers+@pitt.edu
(412) 624-7036 voice
(412) 624-9149 fax
http://advlearn.lrdc.pitt.edu/

After August 1st, 1998:

Dept. of Information and Computer Sciences
University of Hawai'i

1680 East West Road, POST 317
Honolulu, HI 96822
suthers@hawaii.edu
(808) 956-7420 voice
(808) 956-3548 fax

Portions presented at the 1998 AERA Annual Meeting, San Diego, California.

This is a work in progress: comments welcome.

Abstract. The Belvedere software environment was designed to support students engaged in collaborative learning while solving ill-structured problems requiring integration of multiple sources of data to evaluate competing hypotheses or solutions. Students are posed with web-based "science challenge problems," which present a recent or current debate in science along with on-line articles, data, and suggestions for hands-on data-gathering activities. Students use the Belvedere inquiry-diagramming facility to record hypotheses under consideration, information gathered, and the evidential relations between them. Preliminary studies with Belvedere suggest that the design of representational tools can have a significant effect on the learners' knowledge-building discourse. However, these effects are insufficiently studied. After several years of laying the groundwork by building and deploying such software, the author and colleagues have begun such an in-depth investigation, examining the effects of textual, diagrammatic and tabular representational tools on the quality of knowledge-building discourse between learners. The paper describes the Belvedere work that led to this position, lays out a research agenda, and describes pilot studies now underway.

1. Introduction

Decades of research into cognitive and social aspects of learning has developed a clear picture of the importance of learners' active involvement in the expression, examination, and manipulation of their own knowledge [e.g., Chi & Bassock 1989; Perkins et al. 1995; Scardamalia et al. 1994], as well as the equal importance of guidance provided by social processes and mentorship [Brown & Campione 1994; Lave & Wenger 1991; Slavin 1990; Webb & Palincsar 1996]. Recently these findings have been reflected in software technology for learning: systems are now providing learners with the means to construct and manipulate their own solutions while being guided by the software and interacting with other learners.

The present work is concerned with representational tools in support of "computer supported collaborative learning" (CSCL, [Koschmann 1994; Pea 1994]). In the K-12 environment, collaborative learning with computers is not only a necessity due to resource limitations, but is also a way to engage learners more actively in their learning [Johnson & Johnson 1989; Rysavy & Sales 1991; Slavin 1990; Webb & Palincsar 1996]. In postsecondary distance education, electronic forms of collaborative learning help reduce the isolation of telecommuting learners and increase the interactivity of the distance learning experience [Abrami & Bures 1996; Jonassen et al. 1995].

The Belvedere software environment was designed to provide representational and coaching support to students engaged in collaborative learning while solving ill-structured problems in science and other areas (such as public policy), problems requiring integration of multiple sources of data to evaluate competing hypotheses or solutions. The educational goals are to facilitate student's acquisition of subject matter knowledge, of an understanding of how scientists' data gathering activities are motivated by or have

implications for theoretical issues, and of critical inquiry and collaborative problem solving skills. Students are posed with web-based "science challenge problems," which present a recent or current debate in science along with on-line articles, data, and suggestions for hands-on, data-gathering activities. Students use the Belvedere inquiry-diagramming facility to record hypotheses under consideration, information gathered, and the evidential relations between these. They also use standard computer tools such as spreadsheets. A computer-based coach provides advice based on students' diagramming activity.

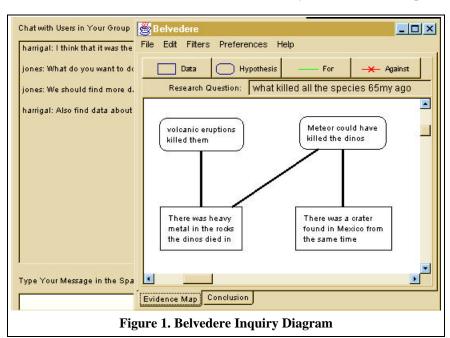
During classroom and laboratory use of "Belvedere," the author observed that the categorical choices required by the software influenced distinctions attended to by learners, and that learners' interactions appear to be further guided by the objects and relationships (expressed or potential) that their representations make salient. Based on these observations, the author is undertaking a systematic study of how variation in features of the representational tools provided by these environments can have a significant effect on the learners' knowledge-building discourse and on learning outcomes. Representational tools can help learners see patterns, express abstractions in concrete form, and discover new relationships [Goldenberg 1995]. These tools can function as "epistemic forms" [Collins & Ferguson 1993]: cognitive tools that lead learners into certain knowledge-building interactions [Jonassen & Reeves 1996]. The present work is based on the hypothesis that properly designed representational tools can function as "epistemic forms" for collaborative as well as cognitive learning interactions. As learner-constructed external representations become part of the shared context, the distinctions and relationships that are made salient by these representations may influence knowledge-building discourse in certain predictable ways to be discussed.

This paper begins with a description of the Belvedere software environment. Then, examples of the kinds of interactions that led to the representational bias hypothesis are provided, followed by a theoretical account and an outline of the research program that is presently underway.

2. Software for Collaborative Inquiry

The "Belvedere" software is a networked software system that provides learners with shared workspaces for coordinating and recording their collaboration in scientific inquiry. The version described in this paper, Belvedere 2.0 and 2.1, is a complete redesign and reimplementation of Belvedere 1.0, previously reported in Suthers & Weiner [1995] and Suthers *et al.* [1995]. Belvedere's core functionality is a shared workspace

constructing "inquiry diagrams," which relate data and hypotheses by evidential relations (consistency inconsistency). The software includes artificial also intelligence coaches that provide advice, "chat" a facility for unstructured discussions, and facilities for integrated use with Web browsers. The diagramming window is shown in Figure 1, with an additional window (left side) for a "chat" facility. The default "palette" (the horizontal row of icons) makes salient the most crucial distinctions we want learners



Suthers - 2

to acquire in order to conduct scientific inquiry. Left to right, the icons are "data" for empirical statements, "hypothesis" for theoretical statements, then links representing "for" and "against" evidential relations. Learners use the palette by clicking on an icon, typing some text (in the case of statements) and optionally setting other attributes, and then clicking in the diagram to place the statement or create the link. The palette is configurable: other categories and relations can be added, such as "unspecified" statements about which learners disagree or are uncertain, "principle" for law-like statements, and a link for conjunction. An icon for an automated "coach" can also be added. Extensions underway include alternate views on the workspace (e.g., evidence *tables*), as well as alternate workspace types (e.g., concept maps and causal loop diagrams).

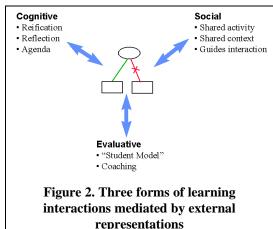
We use a diagrammatic interface for cognitive, collaborative, and evaluative reasons (Figure 2). These reasons, which apply more generally to any visually inspectable and manipulable external representations of one's emerging knowledge, are discussed further below.

2.1 Cognitive Support

Concrete representations of abstractions turn conceptual tasks into perceptual tasks. Thus, shared symbolic representations such as diagrams can help learners "see" and internalize abstractions and keep track of them while working on complex issues. The inquiry diagram serves both as a record of what the learners have done, and an agenda of further work (especially with the help of coaching, discussed below). The representations help guide learners' thinking and activity.

2.2 Collaborative Support

Shared learner-constructed representations such a



diagrams support collaboration by providing shared objects of perception that coordinate group work by serving as referential objects and status reminders, as well as the shared task of constructing the representations. These advantages manifest in different ways depending on whether learners are co-present or collaborating over the network. When they are co-present, diagrams support collaboration by helping learners keep track of and refer to ideas under discussion, whether using a single display, or two displays near each other. In these situations learners often use gestures on the display to indicate prior statements and relationships. In some group configurations we have seen learners work independently, then use gesturing on the display to re-coordinate their collaboration when one learner finds relevant information [Suthers & Weiner 1995]. This can occur when information is brought to the group from off-line sources, such as hands-on experiments. Distally, learners can work in parallel on the same workspace, as long as they are not editing the same object at the same time. On networked computers, all changes are propagated to others working with the same diagram in a "what you see is what I see" manner. In addition to the diagram, a "chat" facility and a remote pointing mechanism support unstructured natural-language conversations, needed to coordinate the inquiry diagramming activity when collaborating at a distance.

2.3 Evaluative Support

Shared learner-constructed representations such as diagrams provide mentors such as the teacher and the computer with *a basis for assessing learners' understanding of scientific inquiry, as well as of subject matter knowledge.* This assessment can support computer coaching of the inquiry process. As described in Paolucci *et al.* [1996] and Toth *et al.* [1997], we have constructed two types of coaches. One provides general advice on the structure of the inquiry diagram from the standpoint of scientific argumentation. It helps the learners understand principles of inquiry such as: hypotheses are meant to explain data, and are not accepted

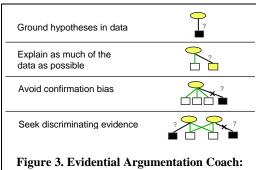


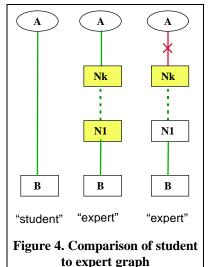
Figure 3. Evidential Argumentation Coach: Selected Advice Patterns

merely by being stated; multiple lines of evidence converging on a hypothesis is better than one consistent datum; hypotheses should try to explain all of the data; one should seek disconfirming evidence as well as confirming evidence (addressing the confirmation bias); discriminating evidence is needed when two hypotheses have identical support; etc. As illustrated in Figure 3, these principles are expressed as "advice patterns" that are matched to student diagrams. When the darkened portions and links marked with "?" in Figure 3 are missing, the corresponding advice can be given.

The other coach performs comparisons between the learners' diagrams and an inquiry diagram provided by a subject matter expert. This coach can provide learners with feedback concerning correctness, or confront learners with new information (found in the expert's diagram) that challenges learners in some way. As shown in Figure 4, when a learner adds an evidential link to the diagram, a search is done in the expert diagram between the corresponding nodes. Information found along the way that is not in the student diagram can be presented as potential elaborations on the learner's link (middle example) or potential contradictions (right hand example).

2.4 Other Features of Belvedere

Other features, briefly noted, include the following. Users can set different "belief levels" for the statements and relations, and display these as line thickness with a "filter." Java applets have been embedded in the Web-based curricular materials, enabling learners to send

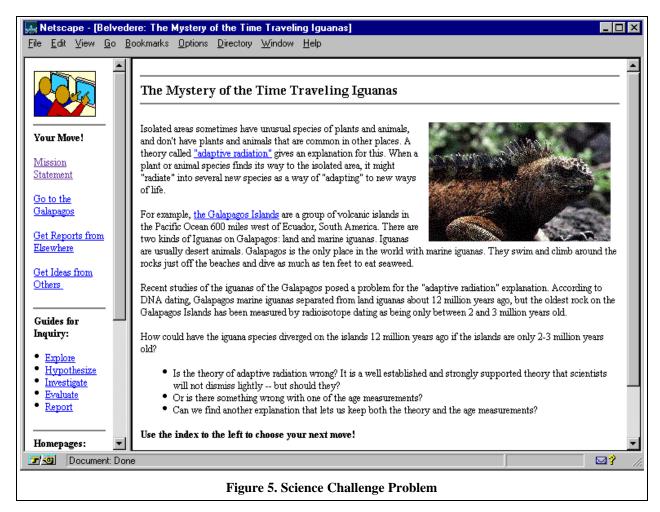


references to these pages into the workspace with a click of a button. References to external objects can also be sent from other applications directly into the Belvedere workspace. For example, in one demonstration of this facility reported in Koedinger, Suthers & Forbus [1998], we enabled an Active Illustration simulation to send summaries of simulation runs as "data" objects into Belvedere. The feasibility of embedding other kinds of documents in Belvedere, such as MS Word and Excel documents, has been demonstrated. It is possible to reinvoke these applications in a platform independent manner. Thus Belvedere can be used as a conceptual organizer for use of various tools during an inquiry.

2.5 Software Implementation

The Belvedere application is written in Java, and is available on request¹ for MacOS, Windows '95, NT, and Solaris. It is deployed as a client within a networked architecture that is designed to provide affordable widespread access to intelligent collaborative educational functionality on a variety of desktop platforms. See Suthers & Jones [1997] for a discussion of the architecture and other design aspects.

¹ Send Email to advlearn+@pitt.edu with platform requirements and a description of your intended use.



2.6 Classroom Implementation

The software was delivered along with a comprehensive methodology for implementing computer-supported collaborative inquiry in the classroom, developed largely by Eva Toth and Arlene Weiner. The approach includes student activity plans worked out in collaboration with teachers. Students work in teams to investigate real-world "science challenge problems," designed to match and enrich the curriculum with attention to National Science Education Standards. A science challenge problem presents a phenomenon to be explained, along with indices to relevant resources (e.g., Figure 5). The teams plan their investigation, perform hands-on experiments, analyze their results, and report their conclusions to others. The classroom activity plans provide teachers with specific guidance on how to manage these activities with different levels of computer resources. Teachers and students are provided with assessment instruments designed as an integral part of the curriculum. Assessment rubrics are given to the students at the beginning of their project as criteria to guide their activities. They guide peer review, as well as helping the teacher assess nontraditional learning objectives. See Suthers, Toth & Weiner [1997] for further information on this integrated approach to classroom implementation, as well as discussion of a third-party evaluation. The present paper focuses on representational issues that arose from informal observations during use in the classroom and laboratory sessions, which have resulted in a new line of work.

_

² The classroom version of the materials are available at http://advlearn.lrdc.pitt.edu/belvedere/materials/.

3. Observations Concerning Representations and Discourse

Belvedere 1.0 was initially used with students aged 12-15 working alone or in pairs in our lab, as well as by students working in small groups in a 10th grade biology classroom. Belvedere 2.x is under use by 9th and 10th grade science classes in Department of Defense Dependent Schools overseas. During this time we learned some important lessons about the role of external representations in collaborative learning.



3.1 Locus of Discourse

Belvedere 1.0 was designed under the assumptions that students can learn the nuances of scientific argumentation if provided with a visual representation language that is *capable of capturing all of these nuances, in which they construct their arguments*, possibly assisted by automated coaching. Under these assumptions, Belvedere 1.0 was provided with a rich palette of statement types and relationships. An example is shown to the left. We expected students to express all of their significant argumentation in the diagrams using primitives such as these.

However, we found that much relevant argumentation was "external," arguing *from* the representations rather than arguing *in* the representations. Faced with a decision concerning some manipulation of the representations, students would begin to discuss substantial issues until they reached tentative agreement concerning how to change the representation. In the process, argumentive statements and relations we would have liked students to represent went unexpressed. Our initial frustration soon gave way to an understanding that this is an opportunity: proper design of manipulable representations can guide students into useful learning interactions. This led to reconsideration of interactions we had seen. Several specific interactions, although subtle, were particularly influential in our thinking.

3.2 Discussion Initiated by Categorical Choices

We often observed that learners who were using Belvedere, which requires all knowledge units to be categorized at the time of creation, initiated discussion of the appropriate categorical primitive for a given knowledge unit when they are about to represent that unit [Suthers 1995]. Although

this is not surprising, it is a potentially powerful guide to learning, provided that it happens at the right time, and that discussion focuses on the underlying concepts rather than the interface widget to select.

Example. Two students are working with Belvedere, which requires all statements to be categorized as either "data" or "claim." (The example is from a videotape of students in a 10th grade science class using an early version of Belvedere.)

- S1: So data, right? This would be data.
- S2: I think so.
- S1: Or a claim. I don't know if it would be claim or data.
- S2: Claim. They have no real hard evidence. Go ahead, claim. I mean who cares? who cares what they say? Claim.

The choice forced by the tool led to a peer coaching interaction on an distinction that was critically important for how they subsequently handled the statement. The last comment of S2 shows that the relevant epistemological concepts were being discussed, not merely which toolbar icon to press or which representational shape to use.

It is not always useful to confront learners with choices, even if they may become important at some point in the development of expertise. For example, in other interactions with a more complex version of this tool that provided more categories, we saw students' discussion sidetrack on subtle distinctions that were not important at their stage of learning, nor for the task at hand.

Example: Two different students are working with a version of the Belvedere evidence mapping tool that included categories for "theory," "hypothesis," "claim," "warrant," "observation" (later called "data"), and "law":

S_M: "So what would that be..."

S E: "Uhh..."

S_M: "An ob--"

S E: "A claim?"

S_E consults sheet of paper in front of her; [pause] "How about a law? scientific color?"

S M: "Do you want to say a warran-- uhh, no."

S_E?: "Wait, what's a warrant? I just read that; why some things..."

S_M: "[sigh] Oh dear."

S_E: "Kind of like a law, like ..." [pause]

S_M: "Yeah, but there are exceptions, I think, because it can't travel, like, complete distances."

It was not necessary for these students to be struggling with all of these concepts at the outset of their learning experience, although S_M's final utterance was a nice touch. These observations led us to simplify the representational tool.

To summarize we found suggestive evidence for the hypotheses that (1) by manipulating the primitive concepts and relations provided by a representational formalism, it is possible to manipulate *what* distinctions are attended to by learners; (2) by manipulating when the software requires that a category be chosen, it is possible to manipulate *when* these distinctions are attended to. Combining this control with a cognitive analysis of a given learning task, we believe it is possible to design interface interaction sequences that draw learners' attention to the right distinctions at the right time [Kaput 1995; Koedinger 1991].

3.3 Discussion Guided by Salience and Task

Example: Three statements are clustered near each other in a two dimensional display that uses arcs for relationships. There are as of yet no arcs drawn between the statements. The student points to two statements simultaneously with two fingers of one hand, and draws them together as she gestures towards the third statement, saying "Like, I think that these two things, right here, um, together sort of support that" (Figure 6, from a videotape of an early laboratory study of Belvedere).

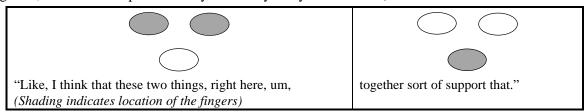


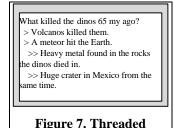
Figure 6. Gesturing to express a relationship between adjacent units.

This event was originally taken as merely an example of how external representations facilitate the expression of complex ideas. However, this observation applies to any external representation. Reconsideration of this example led to the hypotheses that several features of the representational system in use made the student's utterance more likely. First, elaboration on these particular statements is more likely because they (instead of others) are expressed as objects of perception in the representation. Second, this event is more likely to occur in a representational environment that provides a primitive for connecting statements with a support relation than in one that did not -- the students perceive their task as one of linking things together. Third, it may have been easier to recognize the relationship between the three statements because they happened to be spatially nearby each other. In this example, proximity was

determined by the users rather than intrinsic to the representational toolkit. However, a representational tool could constrain proximity based on potential relationships between knowledge units.

3.4 Representations in CSCL Systems for Collaborative Critical Inquiry

Observations such as these led the author to reexamine other software systems in use for the collaborative learning of "critical inquiry" and "scientific argumentation" skills, and identify the need for a series of systematic studies. Several major representational approaches to CSCL for critical inquiry are summarized below, as background for a discussion of their implications for discourse.

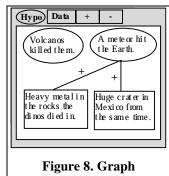


Discussion

Hypertext/hypermedia systems include CLARE [Wan & Johnson 1994]; CSILE [Scardamalia & Bereiter 1991; Scardamalia *et al.* 1992], the Collaboratory Notebook [Edelson & O'Neill 1994 O'Neill & Gomez 1994], Web-Camile and Web-SMILE [Guzdial *et al.* 1997]. (Seminal systems include Apple's HyperCard, as well as gIBIS [Conklin & Begeman 1987] and NoteCards [Harp & Neches 1988], which were not developed for educational applications.) These systems all have in common a hyperlinking of different comments relevant to an issue, usually with categorization of the hyperlinks or their targets with labels such as "answer, argument, problem, solution,

comment," etc. There is wide variation in this category: some take the form of a threaded discussion or other tree structure that may be viewed in summary form (see Figure 7 for a characterization), while others support construction of graphs of "nodes" or "cards" through which one navigates, viewing one card at a time. For the purposes illustration, threaded discussions (Figure 7) will stand for this class of representations. This choice was made to provide a simple reference point for comparison, and is not meant to detract from the richness of mature systems such as CSILE.

Argument mapping environments, a variation on concept mapping [Novak 1990], include Belvedere [Suthers *et al.* 1995; Suthers & Weiner 1995; Suthers *et al.* 1997], ConvinceMe [Ranney *et al.* 1995], and Euclid [Smolensky *et al.* 1987]. All of these utilize node-link graphs representing argumentation or evidential relationships between assertions (usually categorized as "hypothesis" versus "data" or "evidence"). As characterized in Figure 8, the entire graph is viewed and manipulated at once, distinguishing these systems from hypermedia environments in which one normally views and manipulates one node of the graph at a time.



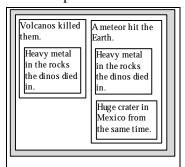


Figure 9. Containment

SenseMaker [Bell 1997], a component of the KIE system [Bell et al. 1995], exemplifies an intermediate approach. Statements are organized in a 2-dimensional space and viewed all at once, as in argument graphs (see Figure 9). However, SenseMaker uses containment rather than links to represent the relationship of evidential support: an empirical statement is placed inside the box of the theory it supports. Logically this is similar to keeping a list for each hypothesis of the empirical observations that support it. SenseMaker also uses containment to represent decomposition of a theory into hypotheses, a facility that was also available in early versions of Belvedere, but not used spontaneously by students.

Finally, another representation is an evidence or criteria matrix (Figure 10). Such matrices organize hypotheses (or solutions) along one axis, and empirical evidence (or criteria) along another, with matches between the two being expressed symbolically in the cells of the matrix. Puntambekar *et al.* [1997] experimented with such a representation in a paper-based collaboration tool.

+ -			
Data \ Hypo	Volcani c	Me te or	
Heavy metal in the rocks.	+	+	
Huge crater in Mexico.		+	

Figure 10. Matrix

Examining the insets above, the differences in representational formalisms provided by existing CSCL software for critical inquiry is

striking. Yet more striking is the fact that there appear to be no systematic studies *comparing* the effects of external representations on collaborative learning discourse,³ although a number of valuable studies have been conducted on individual representational formalisms. Given that these representations define the fundamental character of software intended to guide collaborative learning, a systematic comparison is overdue. The question is *not* "who's system is better?" but rather "what kinds of interactions, and therefore learning, does each representational formalism encourage?" It may well be the case that all of the above representations are useful, albeit for different learning and problem solving phases or task domains.

3.5 External Representations in Individual and Collaborative Contexts

Substantial research has been and continues to be conducted concerning the role of external representations in individual problem solving [e.g., Chandrasekaran et al. 1993; Larkin & Simon 1987; Novick & Hmelo 1994; Stenning & Oberlander 1995; Zhang 1997]. One might ask whether this research is sufficient to predict the effects of representations in collaborative learning. A related but distinct line of work undertaken in collaborative learning contexts is needed for several reasons. The interaction of the cognitive processes of several agents is different than the reasoning of a single agent [Schoen 1992; Okada & Simon 1997; Perkins 1993; Salomon 1993], so 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 biases. 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. 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, much prior work on the role of external representations in individual problem solving have used well defined problems. Further study is needed on ill structured, open ended problems such as those typical of scientific inquiry.

4. Analysis of Representational Bias

The open question concerns the hypothesis 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 thereby on learning outcomes. The claim is not merely that learners will talk about features of the software tool being used. Rather, with proper design of representational tools, this effect will be observable in terms of learners' talk about and use of *subject matter concepts and skills*. In order to unpack "proper design," we need to investigate the specific effects of particular features of representational tools. This section sketches an initial theory of how representations may guide learning interactions, and applies this analysis to make predictions concerning the effects of selected features of representational tools. The discussion begins with some definitions.

³ At a recent conference on computer supported collaborative learning, private communications with several designers of systems exemplified above corroborated this need. The author and other designers were not aware of such a study, and had all chosen designs based on informed intuition. Guzdial [1997] is an exception.

4.1 Definitions

Representational tools are software interfaces in which users construct, examine, and manipulate external representations of their knowledge. The present work is concerned with symbolic as opposed to analogical representations. A formalism/artifact distinction [Stenning & Yule, 1997] is critical to the present work: A representational tool is a software implementation of a **representational** *formalism* that provides a set of primitive elements out of which representations can be constructed. (For example, in Figure 8 the representational formalism is the collection of primitives for making hypothesis and data statements and "+" and "-" links, along with rules for their use.) The software developer chooses the representational formalism and instantiates it as a representational tool, while the user of the tool constructs particular **representational** *artifacts* in the tool. (For example, in Figure 8 the representational artifact is the particular diagram of evidence for competing explanations of mass extinctions.)

Learning interactions include interactions between learners and the representations, between learners and other learners, and between learners and mentors such as teachers or pedagogical software agents. Our present concern is with interactions between learners and other learners, specifically verbal and gestural interactions termed **collaborative learning discourse**.

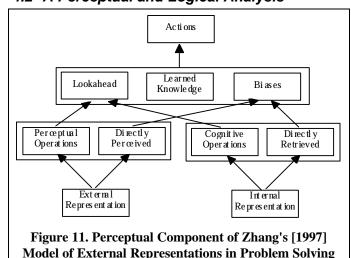
Each given representational formalism manifests a particular **representational bias**, expressing certain aspects of one's knowledge better than others [Utgoff 1986]. The phrase **knowledge unit** will be used to refer generically to components of knowledge one might wish to represent, such as hypotheses, statements of fact, concepts, relationships, rules, etc. Representational bias manifests in two major ways:

Constraints: limits on expressiveness, and constraints on the sequence in which knowledge units can be expressed [Stenning & Oberlander, 1995; Reader unpublished].

Salience: how the representation facilitates processing of certain knowledge units, possibly at the expense of others [Larkin & Simon 1987].

Representational tools mediate collaborative learning interactions by providing learners with the means to articulate emerging knowledge in a persistent medium inspectable by all participants, where the knowledge then becomes part of the shared context. Representational bias *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.

4.2 A Perceptual and Logical Analysis



Zhang [1997] distinguishes *cognitive* and perceptual operators in reasoning with representations (Figure 11). Cognitive operations operate on internal representations; while perceptual operations operate on external representations. Importantly. perceptual operations take place without making an internal copy of the representation (although internal representations may change as a result of these operations). The author's theoretical outlook is highly sympathetic with Zhang's account. Expressed in terms of Zhang's framework, the present analysis is concerned primarily with perceptual rather than cognitive operations. This is because the

proposed work is concerned with how representations that reside in learners' perceptually shared context mediate collaborative learning interactions. While it is the case that cognitive operations on internal representations will influence interactions in the social realm, CSCL system builders do not design internal representations -- they design tools for constructing external representations. These external representations are accessed by perceptual operations, so it is the perceptual features of a representational formalism that are of greatest interest for CSCL systems.

Stenning and Oberlander [1995] distinguish constraints inherent in the logical properties of a representational formalism from constraints arising from the architecture of the agent using the representational formalism. This corresponds roughly to the present distinction between constraints and salience, if one considers primarily the perceptual architecture. "Constraints" are logical and semantic features; while "salience" is not, being better understood in terms of Zhang's distinction between obtaining information by direct perception versus application of perceptual operators (Figure 11): information that is recoverable from a representation is salient to the extent to which it is recoverable by direct perception rather than through perceptual operators.

The discussion now turns to the identification of dimensions along which different representational formalisms vary, and predictions that a given kind of learning interaction will increase along that same dimension.

4.3 Representational Formalisms Bias Learners Towards Particular Ontologies

The first major hypothesis claims that important guidance for learning interactions comes from ways in which a representational formalism *limits* what can be represented [Reader, unpublished; Stenning & Oberlander, 1995]. A representational formalism provides a set of primitive elements out of which representational artifacts are constructed. These primitive elements constitute an "ontology" of categories and structures for organizing the task domain. Learners will see their task in part as one of making acceptable representational artifacts out of these primitives. Thus, they will search for possible new instances of the primitive elements, and hence (according to this hypothesis) will be biased to think about the task domain in terms of the underlying ontology. This point is illustrated by the examples from the author's Belvedere experience, given in section 3.2.

4.4 Salient Knowledge Units Receive More Elaboration

This hypothesis states that learners will be more likely to attend to, and hence elaborate on, the knowledge units that are perceptually salient in their shared representational workspace than those that are either not salient or for which a representational proxy has not been created. This is for two reasons:

Reminding: The visual presence of the knowledge unit in the shared representational context serves as a reminder of its existence and any work that may need to be done with it.

Ease of Reference: It is easier to refer to a knowledge unit that has a visual manifestation, so learners will find it easier to express their subsequent thoughts about this unit than about those that require complex verbal descriptions [Clark & Brennan 1991].

These claims apply to any visually shared representations. However, to the extent that two representational formalisms differ in kinds of knowledge units they make salient, these functions of reminding and ease of reference will encourage elaboration on different kinds of knowledge units. The ability to manipulate learners' elaborations is important because substantial psychological research shows that elaboration leads to positive learning outcomes, including memory for the knowledge unit and understanding of its significance [e.g., Craik & Lockhart 1972; Chi et al. 1989].

Example. Consider the three representations of a relationship between four statements shown in Figure 12. The relationship is one of evidential support. The middle formalism uses containment to represent evidential support, while the right-hand formalism uses an arc. It becomes easier to perceive and refer to the *relationship* as an object in its own right as one moves from left to right in the figure. Hence the present hypothesis claims that relationships will receive more elaboration in the rightmost representational formalism. (The opposite prediction could be made in situations where learners see their task as one of

putting knowledge units "in their place" in the representational environment. Once a unit is put in its place or connected to other units, learners may feel it can be safely ignored as they move on to other units not yet placed or connected.

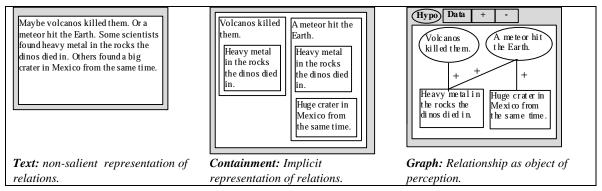


Figure 12. Example of Elaboration Hypothesis

4.5 Salience of Missing Knowledge Units Guides Search for New Knowledge

Some representational formalisms provide structures for organizing knowledge units, in addition to primitives for construction of individual knowledge units. *Unfilled "fields"* in these organizing structures, if perceptually salient, can make *missing* knowledge units as salient as those that are present. If the representational formalism provides structures with predetermined fields that need to be filled with knowledge units, the present hypothesis predicts that learners will try to fill these fields. For example, a two dimensional matrix has cells that are intrinsic to the structure of the matrix: they are there whether or not they are filled with content. Learners using a matrix will look for knowledge units to fill the cells. The ability to manipulate learners' awareness of missing knowledge could be a useful form of scaffolding for metacognitive competence.

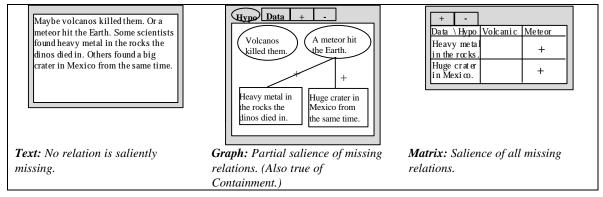


Figure 13. Example of Salient Absence Hypothesis

Example. Figure 13 shows artifacts from three representational formalisms that differ in salience of missing evidential relationships. In the textual representation, no particular relationships are salient as missing: no particular prediction about search for new knowledge units can be made. In the graph representation, the lack of connectivity of the volcanic hypothesis to the rest of the graph is salient. However, once some connection is made to one data item, the hypothesis will appear connected, so one might predict that only one relationship involving each object will be sought. In the matrix representation, all undetermined relationships are salient as empty cells. The present hypothesis predicts that learners will be more likely to discuss all possible relationships between objects when using matrices.

Many other variations on representational formalisms with corresponding predictions are possible; however the discussion must end here for space considerations. The discussion now turns to studies now underway designed to test these predictions.

5. A Systematic Investigation of the Effects of Representational Bias on Discourse

The author is undertaking studies that test the effects of these formalisms on collaborative discourse and learning. At this writing, pilot studies have begun, and a proposal for in-depth study is under review. Four representational formalisms are being compared in a proximal collaborative learning configuration. The four formalisms, which are characterizations of systems being deployed today, were chosen to maximize predicted differences along certain dimensions. These formalisms are **threaded discussions** (Figure 7), **graphs** (Figure 8), **containment** (Figure 9), and **matrices** (Figure 10). These formalisms intentionally differ on more than one feature (see Table 1). The research strategy is to maximize the opportunity to observe significantly different effects on learners' knowledge building discourse and on learning outcomes. These results will then inform well-motivated selection of studies that vary one feature at a time as needed to disambiguate alternate representational explanations for the results. This expedient approach is necessary in order to explore the large space of experimental comparisons within the time scale on which collaborative technology is being adapted.

Based on the features of representational formalisms shown in Table 1 and the discussion of section 4, the following predictions are made. The symbols ">" and ">>" indicate that the discourse phenomenon at the beginning of the list (concept use, elaboration, or search) will occur at a significantly greater rate in the treatment condition(s) on the left of the symbol than in those on the right. The double symbol ">>" indicates greater confidence in the prediction.

Concept Use (section 4.3): Graph, Matrix >> Container, Threaded Discussion. The Graph and Matrix representations require that one categorize statements and relations. This will initiate discussion of the proper choice, possibly including peer coaching on the underlying concepts. The Container and Threaded Discussion representations provide only implicit categorization. There may be some talk of where to put things, but this talk is less likely to be expressed in terms of the underlying concepts.

Elaboration on Relations (section 4.4): Graph > Matrix >> Container > Threaded Discussion. Graphs and Matrices make relations explicit as objects that can be pointed to and perceived, while this is not the case in the other two representations. More tenuously, the ability to link relations to other relations in Graphs may increase elaboration on relations relative to Matrices, if this facility is used. The appearance of one statement inside another's container constitutes a more specific assertion than contiguity of statements in a Threaded Discussion. Hence subjects are more likely to talk about whether a statement has been placed correctly in the Container representation.

Search for Missing Relations (section 4.5): Matrix >> Graph, Container >> Threaded Discussion. In the Matrix representation, there is an empty field for every undetermined relationship, prompting subjects to consider all of them. In Graphs or the Container representations, salience of the lack of some relationship goes away as soon as a link is drawn to the statement in question or another is placed in its container, respectively. Threaded Discussion does not specifically direct search towards missing relationships.

	Threaded	Containers	Graphs	Matrices
Organization of Inquiry Activity	Discussion topics are posted, followed by chronologically organized replies.	Hypotheses are first recorded as boxes in the workspace. Empirical observations are sorted by placing them in boxes.	Hypotheses and empirical observations are recorded at any time as shapes placed in the workspace. Evidential relations are recorded by linking shapes together.	Hypotheses and empirical observations are recorded at any time by creating new columns & rows (respectively). Evidential relations are recorded by placing symbols in empty cells.
Ontology	Implicit: ◆ statements ◆ reply chronology	Implicit: ◆ hypothesis ◆ empirical observation ◆ consistency	Explicit: hypothesis empirical observation consistency inconsistency	Explicit: hypothesis empirical observation consistency inconsistency
Salience of Known Relations	Implicit in Context: ◆ reply chronology	Implicit in Context: ◆ consistency	Explicit Object: • consistency • inconsistency	Explicit Object: • consistency • inconsistency
Salience of Missing Relations	No salience.	Lack of <i>some</i> consistency relation for a hypothesis.	Lack of <i>some</i> consistency or inconsistency relation for a statement.	Relations for <i>all combinations</i> of hypothesis and empirical observation.

Table 1. Features of Selected Representational Formalisms

Subjects are middle school students recruited from area schools. This age group was selected because the "knowledge mapping" software is expected to be most beneficial to them, and because proximal collaborative learning is of special interest in K-12 settings. Subjects are using a version of Belvedere modified for the experiments. They are presented with a "science challenge problem" in a web-browser, and asked to resolve the problem, recording their progress in a representational tool displayed next to the browser. It is important that these are relatively ill-structured problems: at any given point many possible knowledge units may reasonably be considered. This provides the necessary degrees of freedom within which representational bias can work.

Follow-up studies will test the generality of selected results of the first study in a distance learning environment. The method will be similar except that learners will be working in different rooms using a "chat" facility as a medium of discourse instead of speech. The second set of studies will draw upon an undergraduate population, being more representative of learners likely to be using distance collaboration tools.

6. Conclusions

Educational technology is at a crossroads in terms of opportunity for making an impact. Schools K-12 are making larger investments in hardware and software, and colleges and universities are increasingly turning to distance education technology in response to the need to reach a broader customer base. With these opportunities, it is critical that the research and development community maximize the effectiveness of technology for learning. Prior experience with Belvedere suggests that variation in features of the representational tools provided by such technology can have a significant effect on the learners' knowledge-building discourse and on learning outcomes. The paper outlined a systematic investigation being undertaken by the author, one which will inform the design of future software learning environments and provide a better theoretical understanding of the role of representational bias in guiding learning processes.

7. Acknowledgments

The author is grateful to Alan Lesgold for his mentorship, Eva Toth and Arlene Weiner for substantial contributions to Belvedere's curriculum and classroom implementation, Dan Jones for work on the client-server architecture and collaborative database, Kim Harrigal for work on the client, Massimo Paolucci and Joe Toth for work on the coach, Cynthia Liefeld for assistance on the ongoing pilot studies, and Micki Chi for discussions concerning the role of representations in learning. Work was funded by DoDEA's Presidential Technology Initiative and by DARPA's Computer Aided Education and Training Initiative.

8. References

- Abrami, P.C. & Bures, E.M. (1996). Computer-supported collaborative learning and distance education, Reviews of lead article. The American Journal of Distance Education 10(2): 37-42.
- Bell, P. (1997). Using argument representations to make thinking visible for individuals and groups. In *Proceedings of the Computer Supported Collaborative Learning Conference* '97, pp. 10-19. University of Toronto, December 10-14, 1997.
- Bell, P., Davis, E. A., & Linn, M. C. (1995). The knowledge integration environment: Theory and design. In *Proceedings of the Computer Supported Collaborative Learning Conference* '95, pp. 14-21. Mahwah, NJ: LEA.
- Brown, A. L. & Campione, J. C. (1994). Buided discovery in a community of learners. In *K. McGilly (Ed), Classroom Lessons: Integrating Cognitive Theory and Practice*. Cambridge: MIT Press, 1994. pp. 229-270.
- Chandrasekaran, Bl., Narayanan, N., & Iwasaki, Y. (1993). Reasoning with diagrammatic representations: A report on the spring symposium. *AI Magazine 14*(2): 49-56. 1993.
- Chi, M. & Bassok, J. (1989). Learing 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.
- Chi, M.T.H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13:145-182, 1989.
- 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, 1991, 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.
- Conklin, J. & Begeman, M.L. (1987). gIBIS: A hypertext tool for team design deliberation. In *Hypertext'97 Proceedings*, Chapel Hill, NC, pp 247-252. New York: ACM.
- 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.
- Edelson, D. & O'Neill, D. (1994). The CoVis Collaboratory Notebook: Computer support for scientific inquiry. Paper presented at the *Annual Meeting of the American Educational Research Association*, New Orleans.
- Goldenberg, E. P. (1995). Multiple Representations: A Vehicle for Understanding Understanding. In *D. Perkins, J. Schwartz, M. West, & M. Wiske (Eds) Software Goes to School: Teaching for Understanding with New Technologies*, pp. 155-171. New York: Oxford University Press 1995.
- Guzdial, M. (1997). Information ecology of collaborations in educational settings: Influence of tool. *Proc.* 2nd Int. Conf. on Computer Supported Collaborative Learning (CSCL'97), Toronto, December 10-14, 1997. pp. 91-100.
- Guzdial, M., Hmelo, C., Hubscher, R., Nagel, K., Newstetter, W., Puntambekar, S., Shabo, A., Turns, J., & Kolodner, J. L. (1997). Integrating and guiding collaboration: Lessons learned in Computer-

- Supported Collaborative Learning research at Georgia Tech. *Proc. 2nd Int. Conf. on Computer Supported Collaborative Learning (CSCL'97)*, Toronto, December 10-14, 1997. pp. 91-100.
- Harp, B. & Neches, R. (1988). Notecards: An everyday tool for aiding in complex tasks. *Proc. Architectures for Intelligent Interfaces Workshop*, Monterey, California, Marcy 1988. ACM/SIGCHI.
- Johnson, D. & Johnson, R. (1989). *Cooperation and Competition: Theory and Research*. Interaction Book Company, 1989.
- Johnson, R.T., Johnson, D.W., & Stanne, M.B. (1985). Effects of cooperative, competitive, and individualistic goal structures on computer-assisted instruction. *Journal of Educational Psychology*, 77(6), 668-677.
- Jonassen, D.H., Davidson, M., Collins, M., Campbell, J., & Bannan Haag, B. (1995). Constructivism and computer-mediated communication in distance education. *Journal of Distance Education* 9(2): 7-26.
- Jonassen, D.H. & Reeves, T.C. (1996). Learning with technology: Using computers as cognitive tools. In D. H. Jonassen (Ed.), Handbook of Research on Educational Communications and Technology. pp. 693-719. New York: Simon and Schuster.
- Kaput, J. (1995). Creating Cybernetic and Psychological Ramps from the Concrete to the Abstract: Examples from Multiplicative Structures. In *D. Perkins, J. Schwartz, M. West, & M. Wiske (Eds) Software Goes to School: Teaching for Understanding with New Technologies*, pp. 130-154. New York: Oxford University Press 1995.
- Koedinger, K. (1991). On the design of novel notations and actions to facilitate thinking and learning. *Proceedings of the International Conference on the Learning Sciences*, pp. 266-273. Charlottesville, VA: Association for the Advancement of Computing in Education. 1991.
- Koedinger, K., Suthers, D, & Forbus, K. (1998). Component-Based Construction of a Science Learning Space. To appear in Proceedings of ITS'98, San Antonio, Texas, August 1998.
- Koschmann, T. D. (1994). Toward a theory of computer support for collaborative learning. *The Journal of the Learning Sciences*, 1994 **3**(3): 219-225.
- Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Paricipation*. Cambridge: Cambridge University Press.
- 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. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching* **27**(10): 937-49.
- 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. (1997). Collaborative discovery in a scientific domain. *Cognitive Science* 21(2): 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*, Vancouver, BC.
- Paolucci, M., Suthers, D., & Weiner, A. (1996). Automated advice-giving strategies for scientific inquiry. *Intelligent Tutoring Systems, 3rd International Conference, Montreal, June 12-14, 1996.*
- Pea, R. (1994). Seeing what we build together: Distributed multimedia learning environments for transformative communications. *Journal of the Learning Sciences*, 3(3): 285-299.
- Perkins, D.N. (1993). Person-plus: A distributed view of thinking and learning. In *G. Salomon (Ed)*. *Distributed cognitions: Psychological and Educational Considerations* pp. 88-111. Cambridge: Cambridge University Press.
- Perkins, D. N., Crismond, D, Simmons, R., & Unger, C. (1995). Inside Understanding. In D. Perkins, J. Schwartz, M. West, & M. Wiske (Eds) Software Goes to School: Teaching for Understanding with New Technologies, pp. 70-87. New York: Oxford University Press 1995.
- Puntambekar, S., Nagel, K., Hübscher, R., Guzdial, M., & Kolodner, J. (1997). Intra-group and Intergroup: An exploration of Learning with Complementary Collaboration Tools. In *Proceedings of the*

- Computer Supported Collaborative Learning Conference '97, pp. 207-214. University of Toronto, December 10-14, 1997.
- Ranney, M., Schank, P., & Diehl, C. (1995). Competence versus performance in critical reasoning: Reducing the gap by using Convince Me. *Psychology Teaching Review*, 1995, 4(2).
- Roschelle., J. (1994). Designing for cognitive communication: Epistemic fidelity or mediating collaborative inquiry? *The Arachnet Electronic Journal on Virtual Culture*, 2(2), 1994.
- Rysavy, D.M. & Sales, G.C. (1991). Cooperative learning in computer-based instruction. *Educational Technology Research & Development*, 39(2), 70-79.
- Salomon, G. (1993). No distribution without individuals' cognition: A dynamic interactional view. In *G. Salomon (Ed). Distributed cognitions: Psychological and Educational Considerations* pp. 111-138. Cambridge: Cambridge University Press.
- Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences* 1(1), 37--68.
- Scardamalia, M., Bereiter, C., Brett, C., Burtis, P.J., Calhoun, C., & Smith Lea, N. (1992). Educational applications of a networked communal database. *Interactive Learning Environments*, 2(1), 45-71.
- Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: Trying to bring the classroom into World 3. In *K. McGilly (Ed), Classroom Lessons: Integrating Cognitive Theory and Practice*. Cambridge: MIT Press, 1994. pp. 201-228.
- Schoen, D. (1992). Designing as reflective conversation with the materials of a design situation. Knowledge-Based Systems Journal - Special Issue on AI and Design, Vol. 5, No. 1, 1992, pp. 3-14.
- Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice*. Englewood Cliffs, NJ: Prentice-Hall.
- Smolensky, P., Fox, B., King, R., & Lewis, C. (1987). Computer-aided reasoned discourse, or, how to argue with a computer. In R. Guindon (Ed.), Cognitive science and its applications for human-computer interaction (pp. 109-162). Hillsdale, NJ: Erlbaum.
- Stenning, K. & Oberlander, J. (1995). A cognitive theory of grahical and linguistic reasoning: Logic and implementation. *Cognitive Science* 19(1): 97-140. 1995.
- Stenning, K. & Yule, P. (1997). Image and language in human reasoning: A syllogistic illustration. *Cognitive Psychology* 34: 109-159.
- Suthers, D. & Jones, D. (1997). An architecture for intelligent collaborative educational systems. *AI-Ed 97, the 8th World Conference on Artificial Intelligence in Education*, Kobe Japan, August 20-22, 1997.
- Suthers, D., Toth, E., and Weiner, A. (1997). An Integrated Approach to Implementing Collaborative Inquiry in the Classroom. *Proc. 2nd Int. Conf. on Computer Supported Collaborative Learning (CSCL'97)*, Toronto, December 10-14, 1997. pp. 272-279.
- Suthers, D. (1995). Designing for internal vs. external discourse in groupware for developing critical discussion skills. *CHI'95 Research Symposium*. Denver, May 1995.
- Suthers, D., Weiner, A., Connelly, A. and Paolucci, M. (1995). Belvedere: Engaging students in critical discussion of science and public policy issues. AI-Ed 95, the 7th World Conference on Artificial Intelligence in Education., August 16-19, 1995, Washington DC
- Suthers, D. and Weiner, A. (1995). Groupware for developing critical discussion skills. *CSCL* '95, *Computer Supported Cooperative Learning*, Bloomington, Indiana, October 17-20, 1995.
- Toth, J., Suthers, D., and Weiner, A. (1997). Providing expert advice in the domain of collaborative scientific inquiry. 8th World Conference on Artificial Intelligence in Education (AIED'97), Kobe, August, 1997.
- Utgoff, P. (1986). Shift of bias for inductive concept learning. In R. Michalski, J. Carbonell, T. Mitchell (Eds.) Machine Learning: An Artitificial Intelligence Approach, Volume II, Los Altos: Morgan Kaufmann 1986, pp. 107-148.
- Wan, D., & Johnson, P. M. (1994). Experiences with CLARE: a Computer-Supported Collaborative Learning Environment. *International Journal of Human-Computer Studies*, October, 1994.

- Webb, N. (1989). Peer interaction and learning in small groups. *International Journal of Education Research*, 13:21-40, 1989.
- Webb, N. & Palincsar, A. (1996). Group proxesses in the classroom. In *Handbook of Educational Psychology*, D. Berlmer & R. Calfee, Eds. Simon & Schuster Macmillian, New York 1996.
- Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, 21(2): 179-217, 1997.