DOCUMENT RESUME

ED 392 842 TM 024 671

AUTHOR Diehl, Christine; And Others
TITLE Multiple Representations for Improving Scientific Thinking.
SPONS AGENCY National Academy of Education, Washington, D.C.; National Science Foundation, Washington, D.C.; Spencer Foundation, Chicago, Ill.
PUB DATE Apr 95
CONTRACT GER-9355034
NOTE 36p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995). Tables contain handwriting that may not reproduce well.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Beliefs; *Computer Assisted Instruction; *Diagrams; Evaluation Methods; Higher Education; *Scientific Attitudes; *Thinking Skills; *Undergraduate Students
IDENTIFIERS Convince Me Computer Software; *Representational Modeling

ABSTRACT

The "Convince Me" "reasoner's workbench" software provides a means of explicating and revising arguments and a "reasoning engine" for assessing one's beliefs. A proposed modification to the Convince Me interface, the addition of an online diagrammatic representation to an argument's structure, is premised on the belief that multiple representations can lead to a more robust understanding of a concept and a belief in the utility of diagrams as tools for reflection. This study investigated how students made sense of their arguments, as structured with the Convince Me interface, through the use of diagrams. The impact of re-representing an argument was assessed through the Convince Me interface, an argument diagram, and a more standard argument listing. Results with 24 undergraduates indicated that by itself and with another form of representation, Convince Me appeared to be a good tool for representing subjects' beliefs. Multiple representations appeared useful for enhancing reasoning skills, allowing each student to use a representational form that was comfortable, while providing access to linked representations that offered a different perspective. (Contains 7 figures, 3 tables, and 46 references.) (Author/SLD)

MULTIPLE REPRESENTATIONS FOR IMPROVING SCIENTIFIC THINKING

Christine Diehl, Michael Ranney, and Patricia Schank

University of California, Berkeley

Address correspondence (for any and all authors) to:

EMST Division
4533 Tolman Hall
University of California
Berkeley, CA 94720-1670

(510) 643-5866
cdiehl@violet.berkeley.edu
FAX: (510) 642-3729

ABSTRACT

In considering evidence that individuals appear to be less than perfect reasoners, many researchers suggest that domain-specific and domain-general reasoning skills can and should be taught. The *Convince Me* "reasoner's workbench" software provides both a means of explicating and revising arguments and a "reasoning engine" for assessing one's beliefs. A proposed modification to the *Convince Me* interface, the addition of an on-line diagrammatic representation to an argument's structure, is premised on the belief that multiple representations can lead to a more robust understanding of a concept and on the utility of diagrams as tools for reflection. The present study investigates how students make sense of their arguments, as structured with the *Convince Me* interface, through the use of diagrams. The impact of re-representing an argument in an alternative form is assessed with exercises involving graphical and propositional tools: (a) the *Convince Me* interface, (b) an argument diagram, and (c) a more standard, propositional, argument listing. By itself, and in connection with another form of representation, *Convince Me* appears to be the better tool for representing a subjects' beliefs. Analyses of argument revision and subjects' evaluative comments indicate that, though the three argument representations all influence subjects' reasoning to some degree, the effect of a second representation is particular to its form: The diagram serves as a "visualization" tool and highlights complete/connected argument structures; the propositional listing serves as both a "bookkeeping" and "brainstorming" tool and focuses attention on the details of argument composition; the *Convince Me* system serves as an "analysis" tool and encourages reflection on reasoning strategies. It is concluded that multiple representations can be useful for enhancing reasoning skills, allowing each student the opportunity to express himself or herself with a representational form that is comfortable, while providing access to linked representations that offer a different evaluative perspective.

MULTIPLE REPRESENTATIONS FOR IMPROVING SCIENTIFIC THINKING

Life in today's media-rich information age exposes us on all sides to arguments and persuasion that necessitate the ability to reason effectively in order to decide what to believe. Stepping into the local bar may involve one in the middle of a heated discussion concerning the guilt or innocence of O.J. Simpson; lunch hour conversation may revolve around the pro's and con's of single-payer healthcare reform packages; and pre-election candidate debates could confuse even the most rational of minds. In these situations, one must be able to evaluate arguments, weigh the plausibility of assertions, appraise the rationality of inferences, and consider alternative hypotheses. How well do people meet these requirements for critical thinking? Ranney (in press) suggests that individuals often exhibit the ability to reason coherently and draw valid inferences: "We may often have the impression that people are remarkably adept at maintaining coherent mental models and representations. For example, if you had just told a friend that the city marathon is to be held today, you would not be terribly surprised if she quickly concludes, 'Well, in that case, I should move my car.' Indeed, her inference (and our understanding of it) would be a testament to the everyday power of human coherence-seeking." Although this scenario supports our intuitions of ourselves as rational human beings, researchers describe many difficulties that individuals have with formal and informal reasoning tasks (Linn & Songer, 1993; Perkins, 1986; Schank & Ranney, 1991).

Wason (1968) reports that individuals rarely pursue disconfirming evidence for their beliefs about a rule governing a series of numbers. Kuhn (1993) observed similar phenomena as well as evidence that people are often unlikely to generate alternative theories or evaluate their theory in the light of counter-arguments. Chinn and Brewer (1993) summarize research indicating that science students often ignore, reject, exclude or reinterpret anomalous data rather than re-evaluating their pre-instructional theories. One might ask whether poor performance on experimental tasks in fact indicates a lack of ability, or, perhaps, indicates a gap between competence and performance (cf. Ranney, Schank, & Diehl, in press). Indeed, Chinn and Brewer caution against inviting blind

theory change that ignores context influences (cf. Linn & Siegel, 1984), and advise fostering rational and reflective theory change. In a recent essay, Stanovich (1994, p. 17) contends that this "reflective, skeptical judgement is not something that is encouraged in any of the social settings in which children develop" and argues that "schools are unique settings for the teaching of rational thinking dispositions." Many researchers agree that domain-specific and domain-general reasoning skills can and should be taught (Giere, 1991; Nickerson, Perkins, & Smith, 1985; Perkins, 1985; Ranney, in press; Ruggiero, 1988).

TEC and ECHO

The ECHO Educational Project (EEP) at the University of California, Berkeley, has been investigating reasoning with a computer model based on the Theory of Explanatory Coherence (TEC; see Thagard, 1992). TEC attempts to account for how people decide the plausibility of beliefs asserted in an explanation or argument. The theory is based on a few "hall of fame" principles of reasoning, such as: (1) The believability of an idea generally increases with increasing simplicity. In other words, making lots of (that is, joint) assumptions is often counterproductive, compared to making fewer assumptions; (2) People tend to believe statements when there is more evidence to support them; and (3) We are more likely to believe something that doesn't conflict or compete with other things we strongly believe.

ECHO is a connectionist computer model based on TEC (e.g., Ranney & Thagard, 1988; Thagard, 1989). In ECHO, arguments are represented as networks of nodes. A hypothesis or piece of evidence is represented by a node, and explanatory or contradictory relations are represented by links between nodes. Hypothesis evaluation is treated as the satisfaction of constraints determined from the explanatory relations (that is, explanations and/or contradictions), TEC's principles, and from a few numerical parameters. Given a network of statements and relations between them, node activations are updated in parallel using a simple connectionist settling scheme. When the network of statements settles (or stabilizes), the nodes representing the most mutually coherent hypotheses and evidence exhibit high activation and may be regarded as accepted, and ill-supported nodes representing inconsistent rivals are deactivated and may be

considered rejected. By itself, ECHO neither "learns" connection weights nor infers new propositional relationships; these are provided by the student.

Research findings indicate that ECHO usefully models and predicts students' reasoning. Ranney and Thagard (1988) obtained verbal protocols of students reasoning about ballistics and represented the individual belief statements and explanatory and contradictory links using the ECHO model. ECHO's resulting activations reasonably and temporally replicated which beliefs students accepted or rejected as they considered an increasing body of information. Later studies used ECHO predictively, contrasting a priori activations of statements modeled using ECHO with students' explicit "believability" ratings for those statements (e.g., Schank & Ranney, 1991, 1992). The statements were either embedded in a textual controversy (similar to those presented in Table 2 below) or contributed from students' pre-existing knowledge. This predictive modeling of belief-coherence and revision yielded high activation-vs-rating correlations.

Convince Me

The EEP research group hypothesized that if ECHO helped model and predict human reasoning, it might also be used to teach students the reasoning skills for coherent argumentation. They developed the *Convince Me* Hypercard program to serve as a versatile and user-friendly interface for using ECHO as an instructional tool (Schank & Ranney, 1993; Schank, Ranney, & Hoadley, 1994). *Convince Me* provides both a means of explicating and revising arguments and a "reasoning engine" for coherence-based assessments of one's beliefs. The interface structures an argument by breaking down the process of building an argument into steps that identify hypotheses and evidence, as well as the explanatory and contradictory relations that join them (see Figure 1). The "simulation" and "model's fit" options provide feedback in the form of a correlation between (a) an individual's believability ratings for an argument's propositions and (b) ECHO's activations.

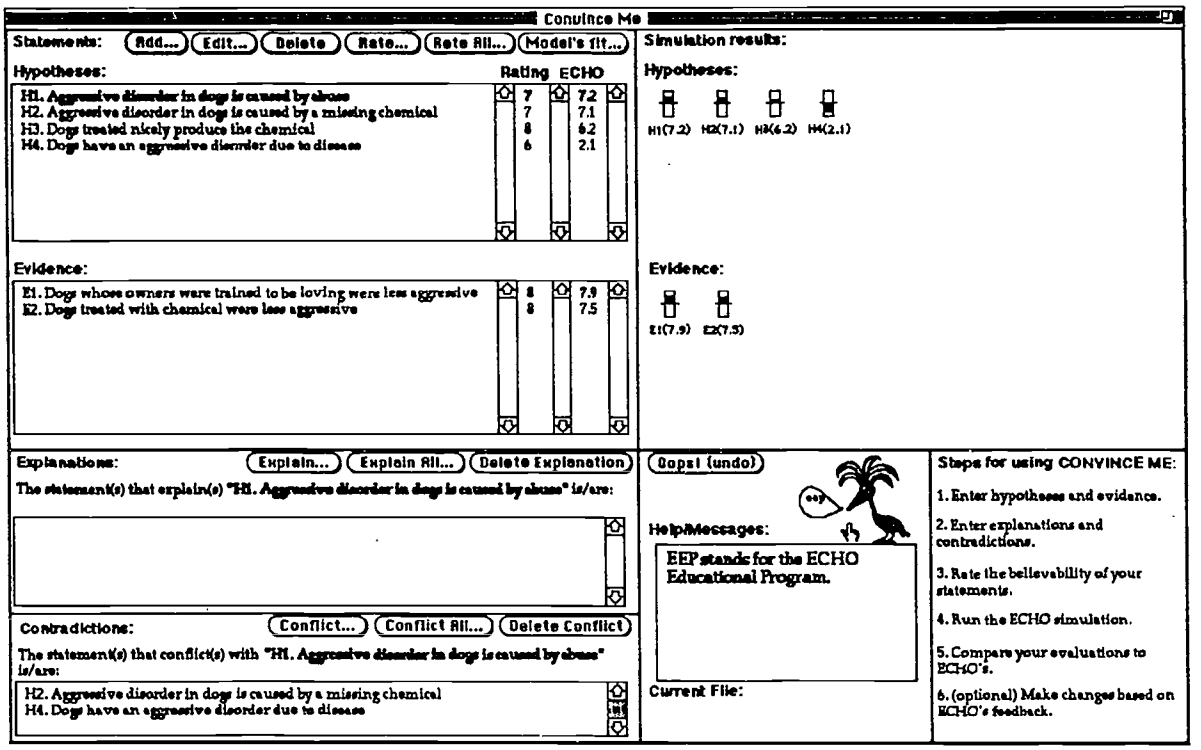


Figure 1. A student's argument in *Convince Me* .

Convince Me asks students to: (1) input their own situational beliefs (see Figure 2); (2) categorize them as hypotheses or evidence (see Figure 2); (3) indicate which beliefs explain or contradict which others (see Figure 2); (4) rate their beliefs' plausibilities (see Figure 2); (5) run the ECHO simulation, which predicts which of their beliefs "should" be accepted or rejected, based on the structure of their argument (see Figure 3); and (6) contrast their ratings with ECHO's predictions (see Figure 3). After comparing their ratings with ECHO, students can modify their argument or ratings or change ECHO's parameters to better model their individual reasoning styles (see Figure 3). Students review their argument to see if they left some explanations or contradictions out, if some independent explanations should be a joint explanation or vice versa, if they want to add or delete some statements, or if they want to change some of their ratings. They are advised not to say that they believe something if they don't, just because ECHO "believes" it. If a student thinks ECHO is being too "tolerant", she might reset the model's numerical parameters

by lowering the *explanation weight* and/or raising the *contradiction weight.* Alternatively, if she thinks ECHO is being too "skeptical", she could lower the *skepticism* weight. She could then re-run the simulation and see how ECHO "reasoning" differs.

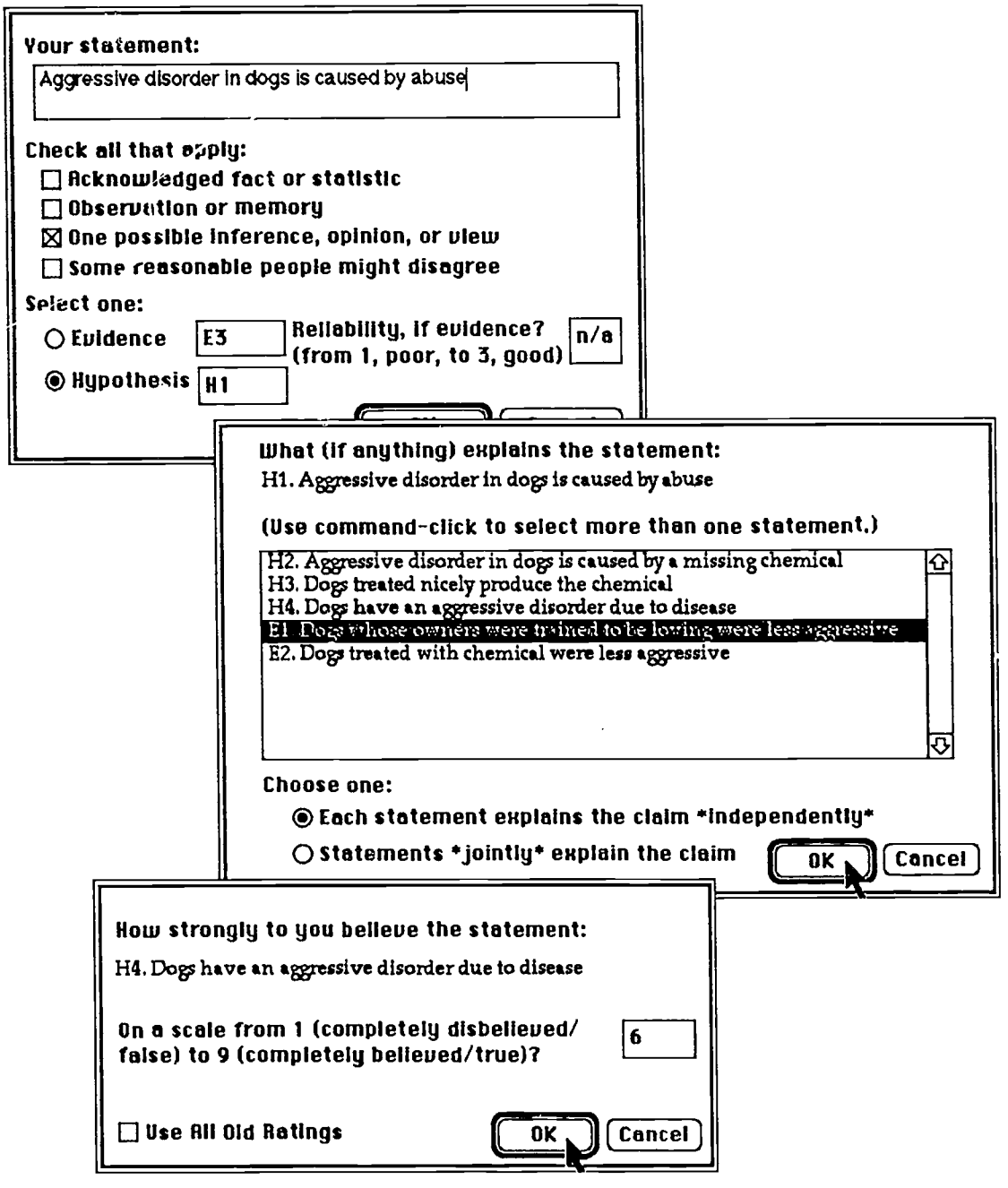


Figure 2. A student adds a hypothesis, an explanation, and rates her beliefs.

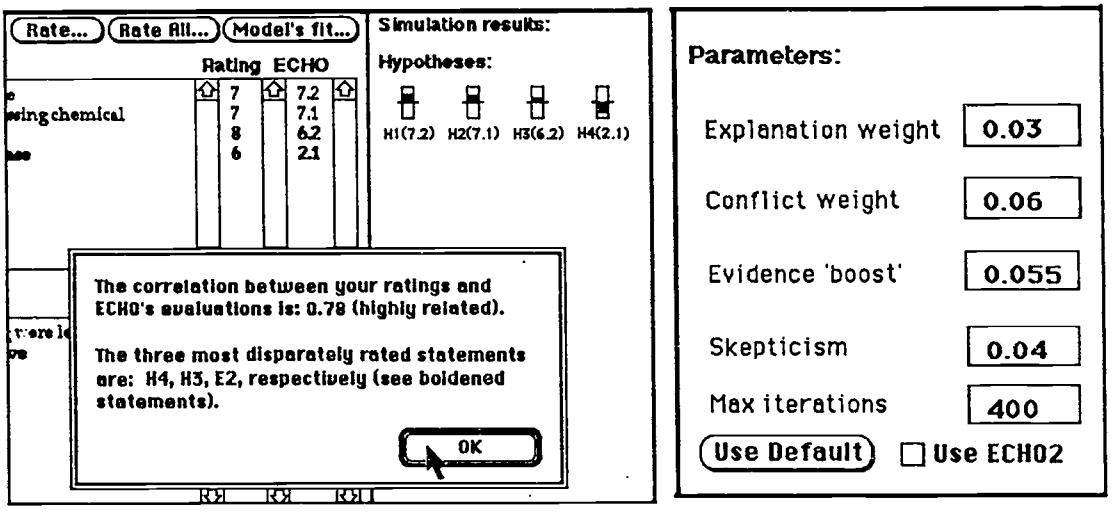


Figure 3. ECHO's ratings from the simulation (the ratings are represented graphically with the "thermometer" icons to the right.) and ECHO's default parameter settings.

Studies investigating the prescriptive utility of *Convince Me* are assessing the impact that this system has on students' ability to generate coherent arguments (Schank, Ranney, Hoadley, Diehl, & Neff, 1994). A recent study shows that the *Convince Me* system seems to make novice reasoners more like experts, even though the intervention employed lasted only a few hours (Ranney, Schank, Hoadley, & Neff, 1994). Another study contrasted students working with *Convince Me* versus students using paper-and-pencil to construct arguments. The results indicate that the interface and feedback enhanced the students' learning such that students were better able to articulate and assess their beliefs by virtue of their experiences with the system (Ranney, Schank, & Diehl, in press). The study also demonstrates that the curriculum itself does not account for the full performance gains or positive transfer available via *Convince Me*.

Design Issues: Multiple Representations

Research in instructional design has also influenced the development and refinement of the *Convince Me* system. Cognitive researchers in many disciplines assert the importance of representation in problem solving and conceptual change (Greeno, 1989; Larkin, 1983; Newell & Simon,

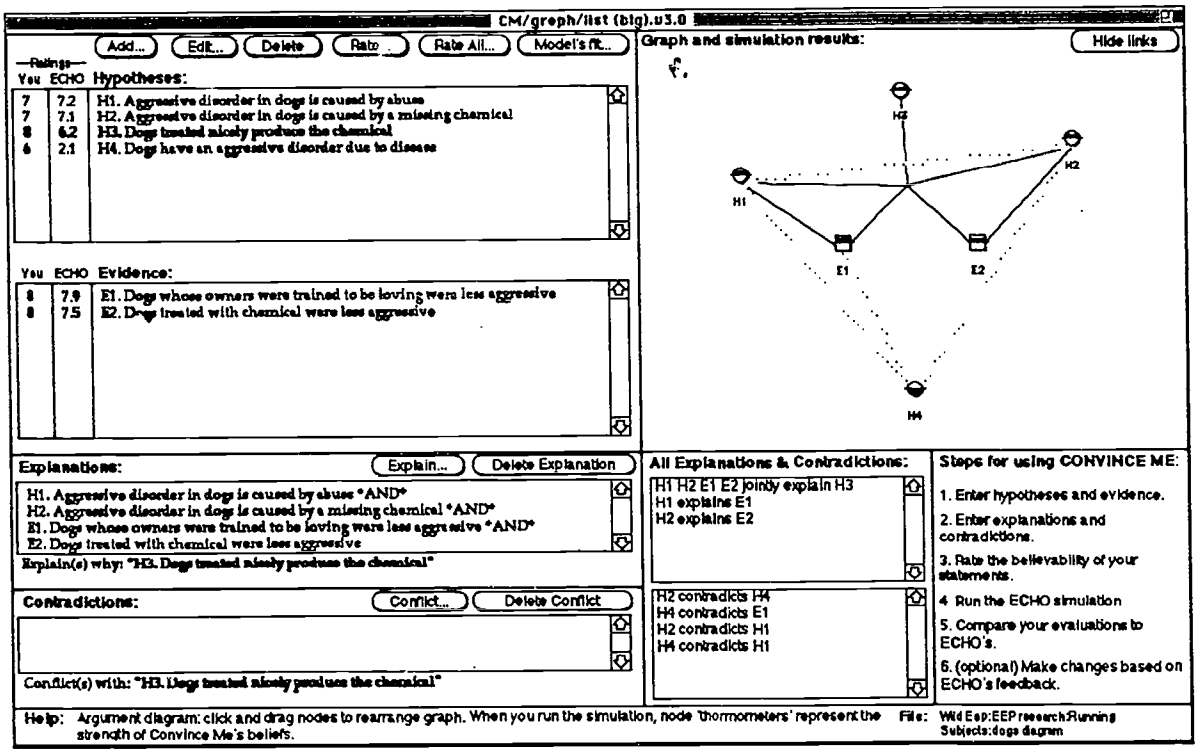
1972). To address the implications of these research findings, researchers have incorporated multiple representations in instruction. Confrey's Epistemology of Multiple Representations (Confrey, Smith, Piliero, & Rizzuti, 1991, p. 18) maintains that the instructional value of multiple representational forms is their potential to: "(1) highlight different aspects of the concept; (2) lead to a convergence across representations that may improve or strengthen the depth of understanding; (3) promote examination of the potential conflict among forms of representations; and (4) allow for assessing how changes in one representation affect another." Moschkovich and her associates (Moschkovich, Schoenfeld, & Arcavi, 1993) maintain that competence in a content domain can be measured by the ability to move flexibly across representations and perspectives, and they suggest revising the curriculum to make connections across multiple representations.

The belief that multiple representations can lead to a more robust understanding of a concept drives a proposed modification to the *Convince Me* interface. Premised on the utility of diagrams as tools for reflection, the modification will add an on-line graphical representation of an argument's structure. Many researchers have championed the computer as a tool for incorporating multiple representations in instruction (Collins & Brown, 1988; Noble, Flerlage, & Confrey, 1993; Schank, Linn, & Clancy, 1993; Snir, Smith, & Grosslight, 1993; White, 1993). Instructional software can provide the necessary environment for access to new forms of representations and simultaneous access to multiple, linked representations. Research in divers disciplines such as physics (Larkin, McDermott, Simon & Simon, 1980) and biology (Kindfield, 1994) indicates that the use of diagrammatic representations can enhance problem solving skills. Diagrams serve to make implicit knowledge explicit and are often used to organize knowledge, re-represent a problem, or direct learning activities.

The computer can serve as a powerful tool for providing students access to diagrammatic representations that parallel an underlying knowledge structure. Students using the GIL tutor for LISP programming (Ranney & Reiser, 1989) build a program "graph" by connecting objects that represent programming constructs. The graph uses a structure that parallels the planning of a program. The Map-SCHOLAR program (Collins, Adams & Pew, 1978) incorporates a graphic

structure that mirrors its semantic network of interrelated geography facts and concepts. More recent work with semantic concept networks has led to the development of graphical mapping tools that allow student to visualize concepts and relations between them (Jonassen, 1992). The *SemNet* program provides a diagrammatic representation of a network of concepts and their interconnections which aids in personal knowledge construction, meaning negotiation, knowledge integration, and social construction of knowledge (Fisher, 1990). Several computer environments for teaching scientific reasoning have also incorporated a diagrammatic representation. Cavalli-Sforza, Moore and Suthers (1993) have designed a computer environment to teach argumentation and scientific reasoning skills. A graphical interface serves as an extension of a computer "coach" by providing diagrammatic visualizations of the relationships among theories and evidence which enables the display of a scientific controversy from multiple perspectives. The Belvedere system (Paolucci, Suthers, & Weiner, 1994) also provides a computer coach and graphical interface for diagramming the structure of scientific theories and arguments. The representation makes abstract ideas and relationships concrete and highlights the logical structure of a debate.

Like the Cavalli-Sforza, etal. and Belvedere systems, the proposed *Convince Me* diagram, in the form of a network of units connected by links, will provide an isomorphic representation of the argument's conceptual structure. This spatial representation should enable a holistic, qualitative, evaluation by the student—one that augments the quantitative evaluation provided by the ECHO simulation (see Figure 4 for a preview of the proposed diagram interface). The benefits of a diagrammatic representation may, however, be offset by the ability of an individual to take advantage of the implicit information it contains (Larkin & Simon, 1987). It is possible that a diagram could have an unintended influence on a student's ability to reason coherently. The graphical representation may cause students to connect propositions for the wrong reasons (e.g., because part of it looks "visually sparse") or unduly influence the size of an argument (e.g., because the diagram gets unwieldy). Thus, while such a change has long been considered, it seems wise to first determine whether the addition of a diagrammatic argument representation would truly be an advantage to students using the *Convince Me* system.



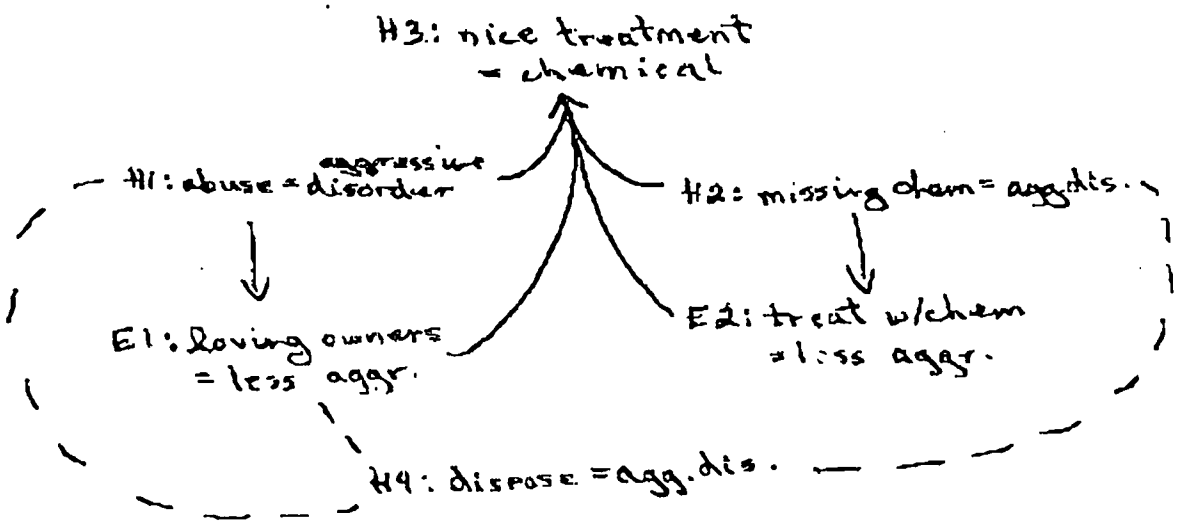


Figure 4. Example *Convince Me* "dogs" argument with trial on-line diagrammatic representation added to the interface (compare with the instantiation in Figure 1). A sample hand-drawn diagram of the argument appears below.

Design Evaluation

The present study investigates how students make sense of their arguments, as structured with the *Convince Me* interface, through the use of their own hand-drawn diagrams. The results will inform future design modifications, such as the addition of the on-line diagrammatic representation of an argument's structure. The impact of re-representing an argument in an alternative form is assessed with exercises involving graphical and propositional tools: (a) the *Convince Me* interface, (b) an argument diagram, and (c) a more standard, propositional, argument listing. Our experiences have suggested that both diagrammatic and textual/propositional/listing (hence "TPL") representations (see Figure 4, above, and Table 1) have both attractive and unattractive aspects. A proper description of the differences among these three representational systems is difficult to generate. Such differences are certainly better captured in a multidimensional fashion, rather than via a linear continuum (cf. Merrill & Reiser, 1994). For instance, *Convince Me*'s representational system, with its (only partial) node-centric highlighting, does not offer the at-a-glance information offered by the (more directly topologically isomorphic) node-link and TPL representations. On the other hand, its forms of belief-wise and argument-wise feedback are offered by neither of the other two representations. Further, the graph-link structure often loses the content of the nodes (replacing the propositions with short labels), unlike the TPL and *Convince Me* representations.

Table 1. A student's TPL representation for the dogs argument.

H1 Aggressive disorder in dogs is caused by abuse	
H2 Aggressive disorder in dogs is caused by a missing chemical	
H3 Dogs treated nicely produce the chemical	
H4 Dogs have an aggressive disorder due to disease	
E1 Dogs whose owners were trained to be loving were less aggressive	
E2 Dogs treated with chemical were less aggressive	
H1 jointly explains H2 E1 E2 H3	H1: 7
H2 contradicts H4	H2: 7
H1 explains E1	H3: 8
H4 contradicts E1	H4: 6
H2 explains E2	E1: 8
H2 contradicts H1	E2: 8
H4 contradicts H1	

It is hypothesized that the *Convince Me* simulation, which provides feedback for subjects to evaluate, will prompt subjects to usefully revise an argument originally represented with a diagram, while the use of alternative forms of argument representation (diagram or listing) will help subjects evaluate and revise arguments originally simulated with *Convince Me*. It is further predicted that a diagram representation will be more beneficial than a listing of the propositions—and their relationships—in helping students see "missing relations" in their *Convince Me* argument and make more complete, connected revised arguments.

METHODS

Subjects

Twenty-four undergraduates (12 men and 12 women) from the University of California, Berkeley, volunteered to receive five dollars per hour to serve as subjects in this experiment. The subjects ranged from 18 to 23 years of age. The subjects had varied backgrounds, though none had received formal instruction in the philosophy of science or logic.

Design and Procedure

All participants completed a curriculum booklet, engaged in three integrative exercises, and completed a questionnaire asking for reflections on their activities. The order of the exercises was completely counterbalanced, and the arguments' textual bases, and gender were counterbalanced. Subjects spent about six hours within a two-week period on the activities, as described below.

Materials and Specific Procedures

Curriculum Booklet (approximately three hours). The curriculum booklet contains three units: Unit 1, "Hypotheses, Evidence, and Theories," discusses distinctions among evidence, hypotheses and theories. In this unit, subjects complete exercises involving diagramming arguments. Unit 2, "Reasoning with Arguments," introduces the need for alternative hypotheses and addresses common biases in reasoning. The exercises in this unit involve listing the propositions

and relations in arguments. Unit 3, "Using *Convince Me*," describes how one could use *Convince Me* to evaluate arguments, and includes exercises using the *Convince Me* interface.

Integrative Exercises (approximately three hours). After completing Unit 3, subjects were given a set of integrative exercises. Each subject completed each of three exercise conditions: In D/CM exercises, subjects (a) generated an argument diagram using paper and pencil, (b) entered this argument into *Convince Me* and evaluated the argument, (c) made any desired changes to the argument with *Convince Me*, and (d) added any changes to the diagram. In CM/D exercises, subjects (a) generated and evaluated an argument with *Convince Me*, (b) diagrammed this argument using paper and pencil, (c) made any desired changes to the argument diagram, and (d) added any changes to the *Convince Me* argument. During CM/L exercises, subjects (a) generated and evaluated an argument with *Convince Me*, (b) made a listing of this argument using paper and pencil, (c) made any desired changes to the argument listing, and (d) added any changes to the *Convince Me* argument.

Argument Texts. The three argument texts used in the integrative exercises are reproduced in Table 2. The texts vary in length, that is, in the number of hypotheses and evidence presented. One text deals with an everyday physiological action—yawning; subjects are likely to draw on personal experience to elaborate the arguments presented. Another text introduces a nature-nurture controversy surrounding the treatment for a canine behavioral disorder. The last text presents a moral dilemma concerning abortion; presumably a visceral issue drawing on very personal beliefs.

Exit Questionnaire. Subjects rated how much they learned from each exercise (based on a seven-point scale) and commented on what they learned. They also indicated (and commented upon) which of the three exercises they believed to be the most useful and which exercise they believed to be the least useful.

Table 2. The three argument texts used in the integrative exercises.

DOGS Some dogs have an aggressive disorder. They bark more than other dogs, growl at strangers, and sometimes even bite. They also tend to have higher blood pressure and heart rate than other dogs. Some researchers think that these dogs get the aggressive disorder when their owners treat them poorly, that is, when the owner neglects the dog, doesn't give it enough love, or hits it. These researchers trained one group of aggressive-disorder dog owners to treat their dogs firmly yet lovingly. They found that all dogs whose owners were trained barked much less, were much friendlier to strangers, never bit a stranger, and had lower heart rate and blood pressure than dogs whose owners had not been trained. These researchers said that their experiment proved that abuse causes dogs to have the disorder. Other researchers disagree. They think that dogs with the disorder are born without a certain chemical in their body. They think that the lack of this chemical elevates their blood pressure and causes the disorder. These researchers gave one group of aggressive-disorder dogs a medicine that contained the chemical. They found that the dogs had a much lower heart rate and blood pressure, were friendlier to strangers, did not bark as much, and never bit anyone. These researchers said that their experiment proved that the missing chemical causes dogs to have the disorder.
ABORTION Smith believes that abortion is wrong because fetuses are alive. Jones disagrees, saying the abortion is fine, because we as a society kill living things (e.g., for food) all the time.
YAWN Wanda and Dave are walking through Pinetown one night, and both notice that an approaching teenager yawns when passing them. Dave thinks that the teenager's yawn was an subconcious aggressive display. He learned in biology that humans are genetically close to apes, and ape studies suggest that apes engage in "threat yawns." In a group, dominant male apes yawn more—an action that shows off their long canine teeth—while subordinate apes more often cover their yawning mouths with their paws. He says that since Pinetown is a dangerous area, this would explain why the teenager yawned when passing them. Wanda disagrees with Dave. She notes that people, as well as non-primates such as dogs, yawn when they are alone as well as in groups. She has read that yawning provides more oxygen to the brain and that the more oxygen, the more glucose we can burn for energy. She thinks that since it is late, the teenager is probably tired and yawned to get more oxygen to stay alert. She claims that the hypothesis that yawning is to increase oxygen also explains why it *seems* contagious—people in the same room are all just breathing the same stuffy air, and all need more oxygen.

DATA SOURCES AND STATISTICS

The primary data analysis is based upon a description of the initial argument structure and changes made to the argument after representing it in a second form. The analysis primarily contrasts the three sorts of exercises: (1) Diagram-*Convince Me* (D/CM), (2) *Convince Me*-Diagram (CM/D), and (3) *Convince Me*-Listing (CM/L). In addition, questionnaire data from subjects on how much they learned from each exercise and how useful the exercises were are also analyzed to determine the perceived utility of each representation.

The comparisons of the original argument focused on the *Convince Me* and diagram representations and are, therefore, based on the two primary exercises (CM/D and D/CM). *Convince Me* served as the initial representation for the CM/L exercise also; however, we were prompted to restrict the analysis of the original argument due to the loss of statistical power that results from an unbalanced comparison between the combined CM/D and CM/L exercises (n=48) and the D/CM exercise (n=24).

For comparisons of two means from groups with an equal number of measures, a related measures *t*-test (two-tailed) of a difference in means was used (Bruning & Kintz, 1987). Differences were typically tested against an alpha of .05 with forty-six degrees of freedom ($t_{.025,46} = |2.01|$). For comparisons of two means from groups with a dissimilar number of measures, a small-sample *t*-test (two-tailed) for independent measures was used. Differences were tested against an alpha of .05 with varying degrees of freedom. A *z*-test was used to assess the significance of a difference between two proportions.

RESULTS AND DISCUSSION

Argument Revision and Representation

Argument revision. More subjects revised their argument ($p < .05$) after re-representing it with *Convince Me* (D/CM, 96%) or as a diagram (CM/D, 88%) than with an argument listing (CM/L, 58%); however, in all exercises more than half the subjects revised their original argument based upon the second form of representation. Thus, all three representational tools—*Convince Me*, diagram and listing—were used in evaluating and revising an argument.

Using argument representations. In the exercises, subjects were asked to copy their original argument (from a diagram or *Convince Me*) to a second representation (diagram, *Convince Me*, or listing). They were then given the opportunity to revise their argument using the new representation (e.g., if a subject copied an argument from *Convince Me* to a diagram, she could make changes to the argument diagram—using a red pen to highlight the revisions). Subjects had little difficulty translating an argument from *Convince Me* to a paper-and-pencil listing, but during the

exercises in which subjects were required to enter an argument from a diagram into the *Convince Me* interface or vice versa, approximately half the subjects misrepresented the links between propositions. The most common mistakes included dividing a joint explanation in *Convince Me* into independent explanations in the diagram, merging independent explanations in a diagram into a joint explanation in *Convince Me*, not representing links from *Convince Me* in the diagram, and adding links to *Convince Me* that were not in the diagram. Although some of the translation errors may actually represent desired argument revisions; the analysis was based solely on the changes that subjects made to the argument *after* reviewing the new representation. In their questionnaire comments, subjects often reported difficulties with drawing and updating the diagrams. While the diagram appears to be useful in evaluating an argument, adding such a representation to the *Convince Me* interface should decrease translation difficulties associated with accurately representing and maintaining an argument's structure in a diagrammatic form.

Model's fit. A model's fit measure was obtained for each subjects' original and revised arguments (revisions made to a diagram or listing were updated in the original *Convince Me* argument and the model's fit function was then run). This correlation between the subjects' believability ratings and ECHO's predictions of the propositions' plausibility (as derived from the model's final activations) reflects how well the subjects' argument structure seems to match their beliefs. The higher the overall correlation, the more ECHO agrees with the subjects' belief ratings—based on their argument. Subjects achieved a better correlational fit ($p < .05$) for arguments entered initially in *Convince Me* (CM/D) over arguments initially represented with a diagram (D/CM); however, there was no significant difference between the correlations attained after an argument had been revised based upon a second representation (see Table 3). The model's fit increased when subjects revised an argument with *Convince Me* (D/CM) and decreased when subjects revised an argument with a diagram (CM/D) or listing (CM/L) ($p < .10$; see "Change in model's fit," Table 3). By itself and in connection with another form of representation, *Convince Me* appears to be the better tool for representing a subject's beliefs.

Table 3. Mean "model's fit" values for initial and revised arguments and mean changes in "model's fit" values.

	Convince Me		Diagram		Listing	
	mean	(exercise)	mean	(exercise)	mean	(exercise)
Model's fit: original argument	.689	(CM/D)	.480	(D/CM)	N/A	N/A
Model's fit: revised argument	.585	(D/CM)	.610	(CM/D)	.494	(CM/L)
Change in model's fit	.104	(D/CM)	-.078	(CM/D)	-.054	(CM/L)

Original argument structure. Arguments were originally structured through the *Convince Me* interface (CM/D) or with a diagrammatic representation (D/CM). For these two exercises, there are no significant differences in the original argument's structure as measured by metrics involving the mean number of components making up the argument—including hypotheses, evidence, explanations, contradictions, instances (any explanatory or contradictory link between two propositions; e.g., joint explanations with three explanatory propositions represents three instances), and total components. However, there is a tendency for *Convince Me* arguments to be more interconnected; they appear to have fewer statements/nodes but more relational links. One reason for this could be the fact that students often use multiple explanations (e.g., some independent and some joint) to connect two statements in the *Convince Me* interface because they can see only a select number of a statement's explanatory links at any time. That is, at any given time students can see which propositions explain a certain piece of evidence/hypothesis, but not all the propositions are explained by that particular piece of evidence/hypothesis.

Revised argument structure. Arguments were revised based upon a re-representation with the *Convince Me* interface, a diagrammatic representation, or a propositional listing. As in the original argument, aspects of the revised argument structures are quite similar across the exercises.

Revisions to argument structure. Overall, the *Convince Me* interface and the diagrammatic representation encouraged more subjects to make changes to their original arguments than did the propositional listing. Revisions to an argument are defined as: (1) for evidence and hypotheses, a revision is an addition of a new proposition, deletion of an existing proposition, or a significant rewording of an existing proposition such that it essentially becomes a new statement (computed as

an addition and deletion); (2) for contradictions, a revision is an addition of a new contradiction or deletion of an existing contradiction; (3) for explanations, a revision is an addition of a new explanation, deletion of an existing explanation, or change to an existing joint explanation (computed as an addition, deletion or both depending on the change); and (4) for belief ratings, a revision represents a change to an existing belief rating.

Most of the differences in revisions between the number of subjects who were using the *Convince Me* interface and subjects using diagrams were of marginal significance. Slightly more subjects added explanations using a diagram and slightly more added contradictions using *Convince Me*. Also, slightly more subjects deleted evidential propositions using *Convince Me*. However, the use of *Convince Me* for argument revision did result in more subjects changing contradictions overall (that is, including additions and deletions).

The differences in revisions between the number of subjects who were using either the *Convince Me* interface or the diagrammatic representation and those using the propositional listing were more notable. The use of a listing for argument revision resulted in fewer subjects changing propositions and links overall. More subjects added hypotheses or links to their argument when using either the *Convince Me* interface or a diagram—and more deleted evidence or links from their argument when revising it using the *Convince Me* interface.

All the representations resulted in an equivalent number of subjects changing belief ratings for evidential propositions; however, differences did appear in changes to belief ratings for hypothetical propositions. When using the *Convince Me* interface or diagrammatic representation to revise an argument, more subjects re-evaluated the belief ratings they assigned to the argument's original hypotheses. This is likely related to the addition of new hypotheses, as well as explanatory and contradictory links among the arguments' propositions (as indicated above).

Time On Task

There was no significant difference in the total time that subjects spent on the three exercises or the amount of time they spent working on the original argument structure. However, subjects did spend more time revising their arguments when using the *Convince Me* interface (16

mins.) than with the diagrammatic representation (9 mins., $p < .05$) or propositional listing (9 mins., $p < .05$). This prompts the question: Did subjects make more revisions with the *Convince Me* system because they enjoyed spending time with it? Or, did subjects take more time when revising with the *Convince Me* interface because it encouraged them to make more revisions? As subjects did not spend predominately more time with the *Convince Me* system while originally structuring their argument, it seems apparent that subjects did indeed gain some benefit from the interface, which led them to work longer on their argument revisions.

Gender Effects

Essentially, no gender effects were observed. There is no effect for gender differences found in the time spent on task, measures of the model's fit, original argument structure, revised argument structure, or argument revisions. The number of men and women who chose to revise their arguments in each exercise was practically equal: (1) D/CM - 11 men and 12 women revised; (2) CM/D - 10 men and 11 women revised; (3) CM/L - 7 men and 7 women revised.

Text Effects

It is not surprising that the heterogeneous basis of the three text passages would result in some differences in argument structure. Although all three texts were relatively brief, subjects constructed arguments that ranged in size from six propositions with three links to 24 propositions with 43 links (interestingly, the largest argument was constructed for the shortest passage). Most subjects appealed to personal experience and previous knowledge to argue the issues, stretching the exercises beyond the "canned" text we provided. Surprisingly, the only significant differences in argument structure were regarding the number of evidential propositions that subjects proposed for the "yawn" text and the number of hypotheses that subjects added to the "dogs" text. Subjects incorporated significantly more evidence in the original and revised argument structures for the "yawn" text than they did for the "dogs" and "abortion" texts. This may reflect the more experiential aspect of the text (i.e., referencing an everyday behavior) versus the more hypothetical situations presented in the other two texts (e.g., a nature-nurture controversy and a moral

dilemma). Subjects added significantly more hypotheses to their revised argument based on the "dogs" text than they did for the "abortion" text (mean values of .58 and .24; $p < .05$). This may also reflect the rather theoretical character of the "dogs" text's presentation of a nature-nurture controversy and the difficulty of addressing two ostensibly equally reliable bodies of evidence.

Reports from the Exit Questionnaire

Subjects reported that they learned more from the diagramming exercises and *Convince Me* than from the listing exercises (mean values of 5.15 and 4.94 vs. 3.85; $p < .05$). They also reported that working with *Convince Me* was apparently the most useful of the exercises (45.8% of the subjects responded "most") and that, overall, the listing exercises were the least useful (54.2% of the subjects responded "least") ($p < .05$ in both instances). It is interesting that subjects reported learning the least from the listing exercises, yet ranked listing as at least as useful as diagramming. The low utility rating for diagrams is likely related to the difficulty that subjects experienced in drawing neat and clear graphical representations, as mentioned above in the discussion on argument representation. A quarter of the subjects ranked the listing exercises as the most useful and, as previously indicated, over half the subjects revised their *Convince Me* argument based on a propositional listing representation. This suggests that even though subjects reported they learned more from the diagram and *Convince Me* exercises, the listing representation provided some utility for subjects in addition to the *Convince Me* interface.

General Discussion and Conclusions

Convince Me aids students in generating and analyzing arguments, providing feedback from a general computational model that yields predictions about the plausibility of an argument's propositions. Among others, this experiment sought to answer the question: Will incorporating a diagrammatic representation into the *Convince Me* interface increase the effectiveness of the software in helping students to more globally and rationally explicate and evaluate arguments? The best answer so far appears to be "yes and no."

It was predicted that the three argument representations contrasted in this experiment would encourage subjects to re-evaluate and revise an argument originally constructed using a different representation, and, as evinced in the results, the three representational forms all influenced subjects' reasoning to some degree. It was further hypothesized that feedback from the *Convince Me* system (especially from the simulation and model's fit—but also from the structural organization imposed by the graphical/iconic interface) would cause subjects to reflect on their reasoning strategies to better represent their beliefs in an argument originally structured with a diagram. The graphical configuration of an argument constructed (or re-represented) with a diagram was hypothesized to help subjects more fully visualize the relationships among propositions—beyond that provided by the *Convince Me* interface—thereby, enabling them to form a more complete and connected argument. It was also expected that a propositional listing would provide subjects with a complete overview of an argument's propositions and relations, which would direct the re-evaluation of an argument structured with the *Convince Me* interface. However, the implicit information afforded by a diagram's spatial representation was hypothesized to provide a more beneficial overview than the propositional listing.

The results show that more than half the subjects revised an argument when re-represented with either *Convince Me*, a diagram, or a propositional listing; the effect of a second representation, though, was particular to its form. Excepting a difference in the model's fit, the original representation of an argument appears structurally similar whether a subject uses *Convince Me* or a diagram. However, the higher model-fit attained through the interactive construction of an argument with *Convince Me* indicates that compiling and comparing the number and types of components in an argument (e.g., hypotheses, contradictions, etc.) does not fully illuminate the strengths and weaknesses, and the soundness or coherence, of the underlying reasoning. The diagrammatic representation and *Convince Me* encourage more revisions to the structure of the argument, yet more is not necessarily better. Arguments revised with either a diagram or listing are less reflective of a subject's beliefs—as evidenced by the decrease in the model's fit. It appears that subjects may require more training in the use of these alternative representations to take full

advantage of their strengths. A more sensitive coding scheme is necessary to ascertain the reasoning strategies that subjects used during their argument revisions.

Representational Reasoning "Tools"

A preliminary analysis of the reasoning underlying subjects' argument revisions and insights gained from their evaluative comments on the exit questionnaire suggest that each argument representation functions as a different type of "reasoning tool." A diagram appears to serve the purpose of a "visualization" tool—focusing subjects' attention on the interrelationships in an argument. Many subjects commented in the exit questionnaires that the diagram helped them visualize the argument better and see "*what* exactly explains *what*":

> Subject 1: Visually representing the arguments and hypotheses [with the diagram] helped me see the overall picture.
>
> Subject 3: With the diagram it was easy to find the interrelationships - very visually based.
>
> Subject 14: The diagram was extremely helpful...by connecting different statements I visually saw the inadequacies and biases in my hypotheses/evidence.
>
> Subject 20: The diagram was the best because it was the most ... visually structured. I liked the fact that I could see where my argument stemmed from and where it was heading.

In arguments originally represented or re-represented with a diagram, subjects often proposed alternative hypotheses (not derived from the text passage; cf., Schank & Ranney, 1991 & 1992) that attempted to "tie together" dichotomous propositions presented in each argument text—suggesting a reasoning strategy sensitive to an argument's relational structure. Two examples illustrate this function of the diagrammatic representation: Subject 15 included an alternative hypothesis in her original diagram for the "dogs" argument which proposed that "*Both* lack of drug [and] treatment by owner play a role in disorder." This hypothesis serves to explain both original hypotheses and both bodies of evidence. Figure 5 shows a "dogs" argument (originally entered in *Convince Me*, then re-represented with a diagram) revised by subject 4 who added the same alternative hypothesis, namely, that "AD [aggressive disorder] is caused by a *combination of* envir. and bio." Both subjects saw that the evidence explained by conflicting hypotheses could be connected

by one comprehensive hypothesis that subsumed the others (cf., Read & Marcus-Newhall, 1993). Moreover, the diagram drawn by subject 4 makes explicit the implicit function of this new "overarching" hypothesis by representing it as physically "arching over" the original argument.

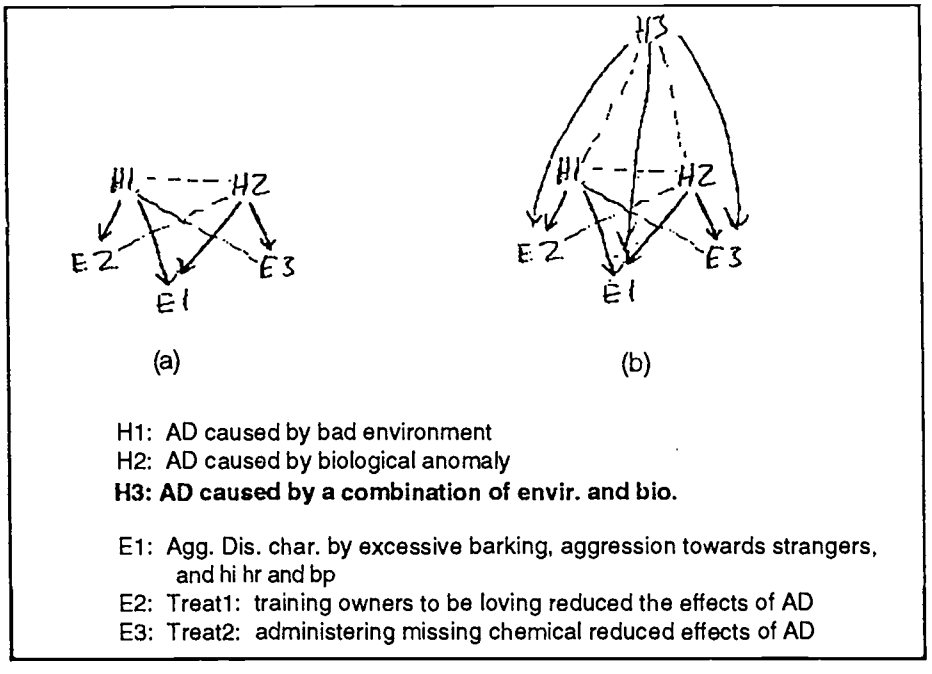


Figure 5. Original "dogs" argument as constructed by subject 4 with *Convince Me* (a), argument revisions after re-representing with a diagram (b), and text of the argument propositions.

Diagrams also prompt/remind subjects to add relations between propositions originally entered in a *Convince Me* argument. For example, Figure 6a shows the diagram representation of the original "abortion" argument that subject 13 constructed using the *Convince Me* interface. The graph clearly depicts two pieces of evidence that were (doubtless inadvertently) left completely disconnected in the original argument. The revised argument (reproduced in Figure 6b) better represents the relations among the propositions. The quantitative results reported in the results (see "Revisions to argument structure") also reflect these outcomes of diagrammatic representation: When re-representing a *Convince Me* argument with a diagram (versus listing), subjects are more likely to add hypotheses and relational explanations, and more likely to re-evaluate original hypothesis belief ratings.

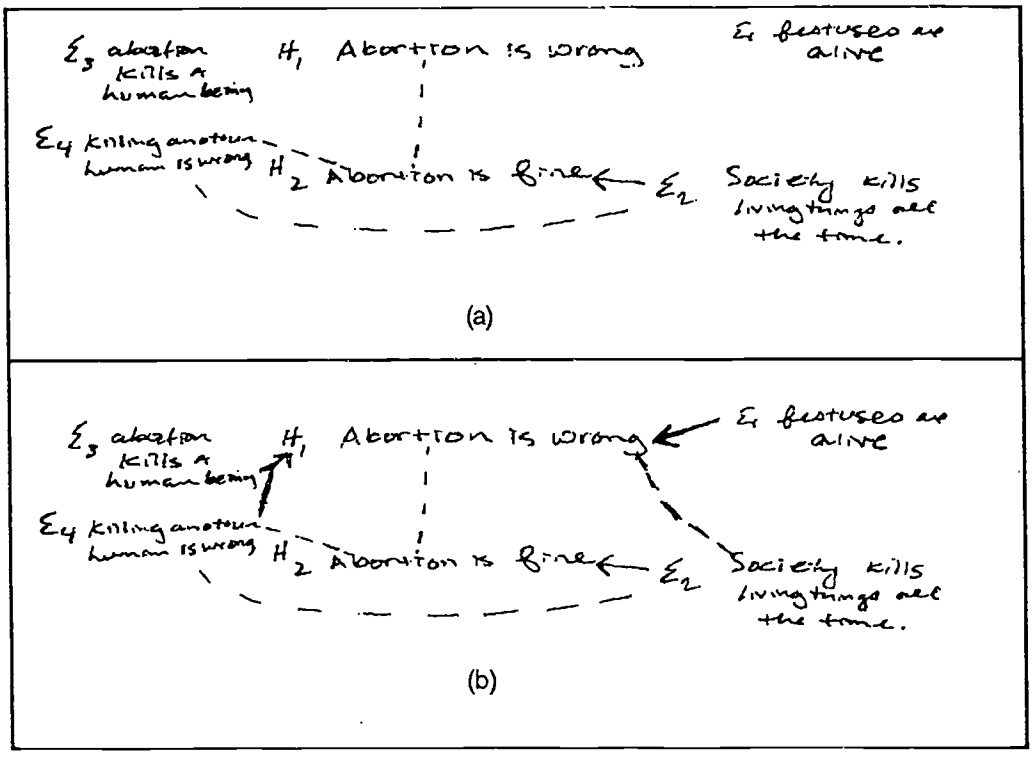


Figure 6. Original "abortion" argument as constructed by subject 13 with *Convince Me* and translated to diagram (a) and argument revisions after re-representing with a diagram (b).

The TPL representation of the relations between propositions also appears to aid students in evaluating the overall argument as structured with *Convince Me*, though apparently to a lesser degree. A propositional listing appears to serve the dual purpose of a "bookkeeping" and "brainstorming" tool—primarily focusing subjects' attention on the details of an argument's composition. Subjects reported that the listing exercises helped them to "put things in order" and "list a wide variety of ideas":

Subject 1: Sometimes my mind couldn't decide and the listing helped put things in order.

Subject 4: [I liked] brainstorming ideas [with the listing].

Subject 14: The listing enabled me to write down options, to delete similar statements, and to list a wide variety of statements.

Subject 17: [The listing] helped me not to forget any key pieces in the [text] passage.

The "bookkeeping" function of the propositional listing often results in minor editing changes to an argument's representation, and, occasionally, in changes to an argument's structure. For example, the only change that subject 8 made to his *Convince Me* "abortion" argument was to combine two pieces of evidence to form one proposition. He noted on the listing sheet that the change was "for clarity." Subject 18 reversed the direction of explanation between his hypotheses and evidence so that the evidential propositions consistently explained the hypothetical propositions. And, after reviewing the re-representation of his Convince Me "abortion" argument, subject 23 realized that he forgot the most important contradiction in the argument—"H1" (Abortion is wrong) contradicts "H2" (Abortion is okay)—noting in the margin, "I left this [contradiction] out in the first trial."

Sometimes, the listing prompts a more substantial change to an argument's structure or to a subject's believability ratings. After re-representing his *Convince Me* "dogs" argument with a listing, subject 12 decided that his hypothesis "a chemical elevates their blood pressure and heart rate" was actually an evidential proposition. He substituted "E6" (the new proposition) for "H5" (the old proposition), but he did not re-evaluate his belief ratings or make any other changes to the argument's relations that reflected the propositional revision.

The "brainstorming" function of the propositional listing results in subjects' proposing alternative hypotheses and related evidential propositions similar to those generated using a diagram representation; however, the added propositions are more often related to "fringe" ideas in an argument and less likely to be well connected to the rest of an argument. For example, subject 9 added the hypothesis "Yawning is not contagious" and two pieces of explanatory evidence to her "yawn" argument to contradict the hypothesis that yawning is contagious; however, these changes did not cause her to re-evaluate the two main hypotheses—that the teenager yawned as an aggressive display or because he was tired. The quantitative results reflect these outcomes of the propositional listing representation as well: When re-representing an argument with a listing, more than half the subjects did make some revision to the original argument, but they were less likely to add hypotheses and explanatory links, and less likely to extend belief revision to hypotheses.

The *Convince Me* "reasoner's workbench" serves the purpose of an "analysis" tool—focusing subjects' attention on their reasoning strategies. It encourages bookkeeping and brainstorming, like the listing representation, and helps to visualize, organize and tie together an argument, like the diagrammatic representation. Subjects' comments reflecting on their use of the *Convince Me* system indicate that it's feedback challenges them to justify the means by which they attained their beliefs:

Subject 6: [With *Convince Me*] I learned how to be more objective in my analysis.

Subject 8: "Convince Me" was the most useful exercise because when it disagreed with you, that showed that there were some aspects of the argument that you failed to perceive.

Subject 15: [*Convince Me*] made me justify my thought processes and many times I had to question them and find other ways to approach a problem.

Subject 17: I enjoyed the way [*Convince Me*] forces you to organize and forces you to address inconsistencies.

Revisions to an argument induced by *Convince Me* include "brainstorming" additional, supporting evidence (see Figure 7), addressing confirmation bias in an argument by adding alternative hypotheses, and "fine-tuning" an argument by revising (adding or deleting) the relationships among propositions and re-evaluating the argument's hypotheses. This "fine-tuning" contribution of the *Convince Me* interface is reflected quantitatively (see "Revisions to argument structure" in results) in the greater likelihood of subjects to revise contradictions (versus diagram or listing) and explanations (versus listing) and in the marginally greater likelihood of subjects to change hypothesis belief ratings (versus listing). The focus on reasoning strategies is supported quantitatively (see "Model's fit' in results) by the mean increase in model's fit for arguments re-represented and revised with *Convince Me*.

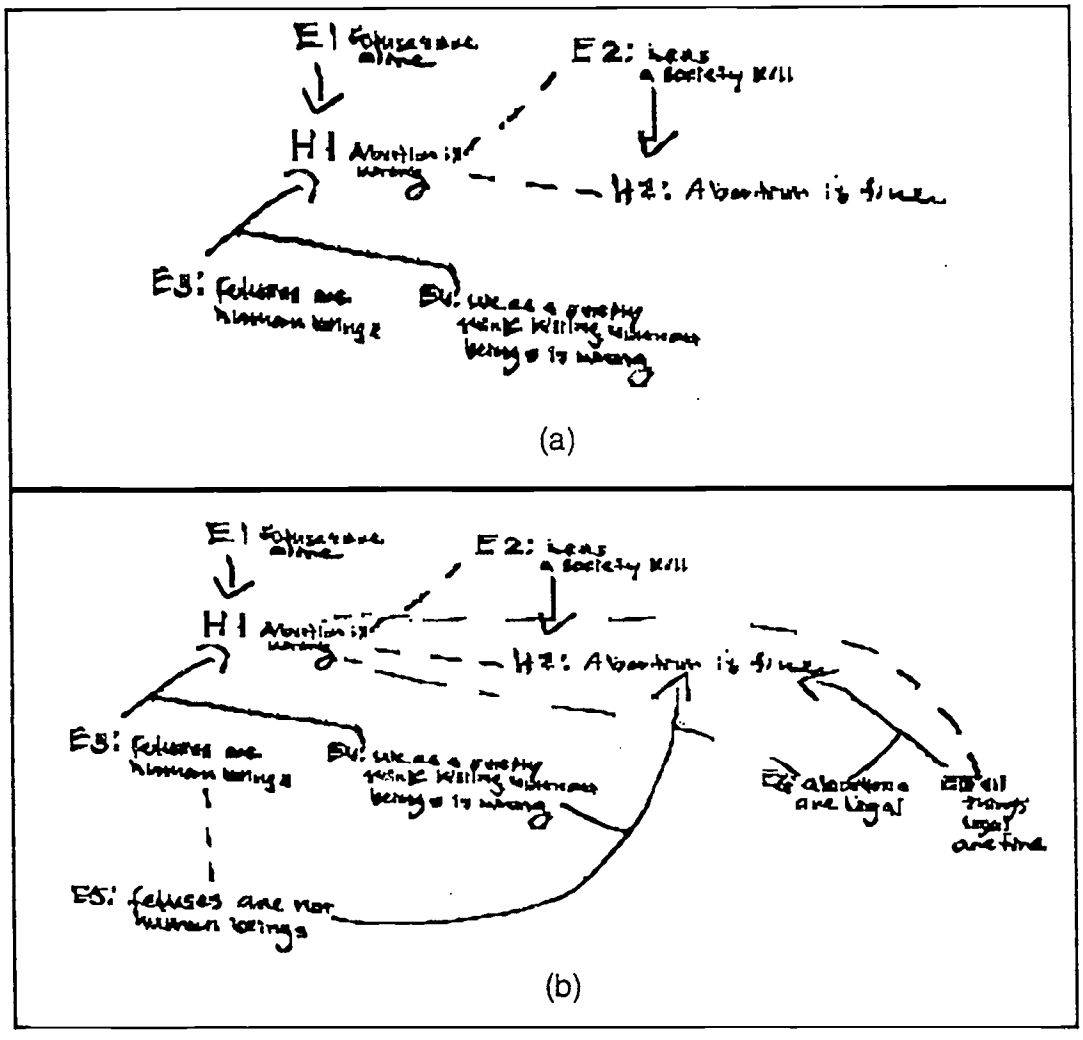


Figure 7. Original "abortion" argument as constructed by subject 17 with a diagram (a) and argument revisions after re-representing with *Convince Me* (b).

Multiple Representations

The above examples demonstrate that multiple representations can be useful for enhancing reasoning skills, allowing each student the opportunity to express himself or herself with a representational form that is comfortable, while providing access to linked representations that offer a different evaluative perspective. As subject 8 so aptly phrased this conclusion, "Although, I think that all three exercise [representations] were extremely useful in their own right, I feel that all three were essential to seeing all aspects of an argument." In affirmation of Confrey's Epistemology of

Multiple Representations, the instructional value of these argument representations lies in their ability to: (a) *Highlight different aspects of the concept:* The diagram ("visualization") tool highlights complete/connected argument structures; the propositional listing ("bookkeeping" and "brainstorming") tool highlights the details of argument composition; and, the *Convince Me* system ("analysis") tool highlights reflective reasoning strategies. (b) *Lead to a convergence across representations that may improve or strengthen our depth of understanding:* The spatial organization of a diagrammatic representation makes often-implicit or hidden (*Convince Me*) knowledge explicit and immediately viewable; the detailed listing representation prompts ordered reflection of a *Convince Me* argument; and the conceptual "scaffolding" and immediate feedback from the *Convince Me* system challenges the reasoning strategies used in constructing a diagrammatic representation of an argument. (c) *Promote examination of the potential conflict among forms of representation:* All three representations give rise to reflection and argument revision driven by new, and possibly conflicting, perspectives. (d) *Allow for assessing how changes in one representation affect another:* The addition of an on-line diagrammatic representation to the *Convince Me* interface will enable student's to see revisions to their *Convince Me* argument automatically represented in the diagram (and, potentially, vice versa). It is debatable whether the benefits gained from a *written* propositional listing would carry-over to an on-line representation—the "bookkeeping/brainstorming" potential appears to derive from the detailed written accounting of the *Convince Me* argument, and any benefits gained from seeing the entire argument "at a glance" might be derived from a diagram, as well.

This experiment substantiates the potential for multiple representational forms to enhance reasoning skills—as long as the use is carefully integrated. However, questions still remain regarding the best way to implement these representations. Suggestions from subjects (reported on the exit questionnaire) favor the current proposed modification of adding an on-line diagrammatic representation to the *Convince Me* interface: "'Convince Me' was most useful because it allowed for all the listing. If 'Convince Me' could have some diagramming feature (for visual people) then it would be great on its own. (Subject 10)"

Educational Implications

The *Convince Me* software and its associated reasoning curriculum aid students in developing a consistent argument, that takes into account as much of the relevant data and hypotheses as possible, and that considers diverse opinions as objectively as possible. *Convince Me* is essentially the only working system that both assists the elucidation of students' thinking while providing them with simulation-based feedback about the coherence of their articulated beliefs and mental representations. In attempting to "convince" *Convince Me,* students are encouraged to reflect on their reasoning strategies. Current assessments indicate that *Convince Me*-like systems may be to coherent reasoning what word processors can be to writing—a useful tool that may even yield transfer to unsupported practice.

Use of the *Convince Me* system to improve scientific reasoning need not be limited to individual interactions. *Convince Me* may prove useful as a tool for collaboration, in similar ways that the CSILE environment supports group dialectical processes (Scardamalia & Bereiter, 1991). The software has been incorporated into the BioQUEST library of biology simulations to aid students in hypothesis generation and testing and theory construction. The BioQUEST Consortium endorses hypotheses as solutions to scientific problem-solving, an activity which encompasses problem posing, problem solving and persuasion of peers (Peterson & Jungck, 1988, p. 49): "Students need to learn very early in their careers that they haven't done science (no matter how many experiments have been done, how much data collected, how many puzzles solved) until they have both reported their results and convinced their peer group as to the reasonableness of their hypothesis." *Convince Me* could support students in their scientific problem-solving activities and serve as a forum for collaborative argumentation.

Students obviously enjoy exploring scientific and everyday controversies using *Convince Me*. In the future, students working with this "reasoner's workbench" will benefit from a greater selection of reasoning "tools." It is hoped that continuing improvements in the software's design will produce a product that is useful for every student, across the curriculum.

ACKNOWLEDGEMENTS

Much of this research was graciously supported by the Spencer Foundation and the National Academy of Education, as well as NSF Traineeship Grant # GER-9355034. We would also like to thank the members of the Berkeley Reasoning Research Group, Marcia Linn, Emily Van Zee, Michelle Million, Lawrence Hamel, and Brian Tierney for their helpful insights, comments, assistance, and support.

REFERENCES

Bruning, J. L., & Kintz, B. L. (1987). *Computational handbook of statistics.* (3rd Ed.) Glenview, IL: Harper Collins.

Cavalli-Sforza, V., Moore, J. D., & Suthers, D. D. (1993). Helping students articulate, support, and criticize scientific explanations. In *World Conference on Artificial Intelligence in Education (AI-ED '93)* (pp. 113-120). Edinburgh, Scotland.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*(1), 1-49.

Collins, A., Adams, M. J., & Pew, R. W. (1978). Effectiveness of an interactive map display in tutoring geography. *Journal of Educational Psychology, 70*(1), 1-7.

Collins, A., & Brown, J. S. (1988). The computer as a tool for learning through reflection. In H. Mandl and A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems.* New York: Springer-Verlag.

Confrey, J., Smith, E., Piliero, S., & Rizzuti, J. (1991). The use of contextual problems and multi-representational software to teach the concept of functions. Final Project Report, Cornell University, Ithaca, N.Y.

Fisher, K. M. (1990). Semantic networking: The new kid on the block. *Journal of Research in Science Teaching, 27*(10), 1001-1018.

Giere, R. (1991). *Understanding scientific reasoning.* (3rd Ed.) Fort Worth, TX: Holt, Rinehart, & Winston.

Greeno, J. G. (1989). Situations, mental models, and generative knowledge. In D. Klahr and K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 285-318). Hillsdale, NJ: Erlbaum.

Jonassen, D. H. (1992). Semantic networking as cognitive tools. In P. A. M. Kommers, D. H. Jonassen, and J. T. Mayes (Eds.), *Cognitive tools for learning* (pp. 19-21). Berlin: Springer-Verlag.

Kindfield, A. (1994). Biology diagrams: Tools to think with. *The Journal of the Learning Sciences, 3*(1), 1-36.

Kuhn, D. (1993). Connecting scientific and informal reasoning. The Development of Rationality and Critical Thinking [Special Issue]. *Merrill-Palmer Quarterly, 39*(1), 74-103.

Larkin, J. H. (1983). The role of problem representation in physics. In D. Gentner and A. L. Stevens (Eds.), *Mental models.* Hillsdale, NJ: Erlbaum.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, 208*, 1335-1342.

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

Linn, M. C., & Siegel, H. (1984). Postformal reasoning: A philosophical model. In M. L. Commons, F. A. Richards, and C. Armon (Eds.), *Beyond formal operations* (pp. 239-257). New York, NY: Praeger.

Linn, M., & Songer, N. (1993). How do students make sense of science? The Development of Rationality and Critical Thinking [Special Issue]. *Merrill-Palmer Quarterly, 39*(1), 47-73.

Merrill, D. C., & Reiser, B. J. (1994). Scaffolding effective problem solving strategies in interactive learning environments. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 629-634). Hillsdale, NJ: Erlbaum.

Moschkovich, J., Schoenfeld, A. H., & Arcavi, A. (1993). Aspects of understanding: On multiple perspectives and representations of linear relations and connections among them. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating research on the graphical representations of functions* (pp. 69-100). Hillsdale, NJ: Erlbaum.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Nickerson, R., Perkins, D., & Smith, E. (1985). *The teaching of thinking*. Hillsdale, NJ: Erlbaum.

Noble, T., Flerlage, D., & Confrey, J. (1993). Experiment, simulation, and analysis: Teaching and learning in an integrated physics and mathematics class. Paper presented at the *American Educational Research Association Annual Meeting*, Atlanta, GA.

Paolucci, M., Suthers, D., & Weiner, A. (1994). Belvedere: Supporting critical discussion of science and public policy issues. Unpublished manuscript, Learning Research and Development Center, University of Pittsburgh.

Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77*, 562-571.

Perkins, D. N. (1986). *Knowledge as design*. Hillsdale, NJ: Erlbaum.

Peterson, N. S., & Jungck, J. R. (1988, March/April). Problem-posing, problem-solving and persuasion in biology education. *Academic Computing*, pp. 14-17, 48-50.

Ranney, M. (in press). Explorations in explanatory coherence. To appear in E. Bar-On, B. Eylon, and Z. Schertz (Eds.), *Designing intelligent learning environments: From cognitive analysis to computer implementation.* Norwood, NJ: Ablex.

Ranney, M., & Reiser, B. J. (1989). Reasoning and explanation in an intelligent tutor for program-ming. In G. Salvendy and M. J. Smith (Eds.), *Designing and using human-computer interfaces and knowledge based systems*. Amsterdam: Elsevier Science Publishers B.V.

Ranney, M., Schank, P., & Diehl, C. (in press). Competence versus performance in critical reasoning: Reducing the gap by using *Convince Me*. *Psychology Teaching Review*.

Ranney, M., Schank, P., Hoadley, C., & Neff, J. (1994). "I know one when I see one": How (much) do hypotheses differ from evidence? In the *Proceedings of the Fifth Annual American Society for Information Science Workshop on Classification Research* (139-156).

Ranney, M., & Thagard, P. (1988). Explanatory coherence and belief revision in naive physics. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*, 426-432. Hillsdale, NJ: Erlbaum.

Read, S. J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology, 65*, 429-447.

Ruggiero, V. R. (1988). *Teaching thinking across the curriculum*. New York, NY: Harpers.

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.

Schank, P. K., Linn, M. C., & Clancy, M. J. (1993). Supporting Pascal programming with an on-line template library and case studies. *International Journal of Man-Machine Studies, 38*, 1031-1048.

Schank, P., & Ranney, M. (1991). An empirical investigation of the psychological fidelity of ECHO: Modeling an experimental study of explanatory coherence. *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 892-897. Hillsdale, NJ: Erlbaum.

Schank, P., & Ranney, M. (1992). Assessing explanatory coherence: A new method for integrating verbal data with models of on-line belief revision. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, 599-604. Hillsdale, NJ: Erlbaum.

Schank, P., & Ranney, M. (1993). Can reasoning be taught? *Educator, 7*(1), 16-21.

Schank, P., Ranney, M., & Hoadley, C. M. (1994). *Convince Me* [Computer program and manual]. In J. R. Jungck, N. Peterson, & J. N. Calley (Eds.), *The BioQUEST Library*. College Park, MD: Academic Software Development Group, University of Maryland.

Schank, P., Ranney, M., Hoadley, C., Diehl, C., & Neff, J. (1994). A reasoner's workbench for improving scientific thinking: Assessing *Convince Me*. In G. H. Marks (Ed.), *Proceedings of the International Symposium on Mathematics/Science Education and Technology*. Charlottesville, VA: AACE.

Snir, J., Smith, C. & Grosslight, L. (1993). Conceptually enhanced simulations: A computer tool for science teaching. *Journal of Science Education and Technology, 2*(2), 373-388.

Stanovich, K. E. (1994). Reconceptualizing intelligence: Dysrationalia as an intuition pump. *Educational Researcher, 23*(4), 11-22.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12*, 435-502.

Thagard, P. (1992). *Conceptual revolutions.* Princeton, NJ: Princeton University Press.

Wason, P. C. (1968). 'On the failure to eliminate hypotheses...'—a second look. In P. C. Wason and P. N. Johnson-Laird (Eds.), *Thinking and reasoning*. Harmondsworth, Middx.: Penguin.

White, B. (1993). Intermediate causal models: A missing link for successful science education? In R. Glaser (Ed.), *Advances in instructional psychology, Volume 4.* (pp. 177-252). Hillsdale, NJ: Erlbaum.