DOCUMENT RESUME

ED 440 873 | SE 063 556

AUTHOR	Diehl, Christine L.
TITLE	"Reasoner's Workbench" Program Supports Students' Individual and Collaborative Argumentation.
PUB DATE	2000-04-00
NOTE	29p.; Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (New Orleans, LA, April 28-May 1, 2000). This research was supported by an American Educational Research Association/Spencer Pre-Dissertation Fellowship, a National Science Foundation Science and Design Traineeship Grant, and a University of California Fellowship.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Abstract Reasoning; *Computer Software; *Computer Uses in Education; *Cooperative Learning; Educational Technology; Secondary Education; Thinking Skills

ABSTRACT

This research evaluates the effectiveness of computer-mediated support for students' individual and collaborative argumentation. "Convince Me" is a "reasoner's workbench" program that aids students in generating and analyzing arguments, providing feedback on argument coherence from a general computational model. Laboratory studies indicate that students working individually with Convince Me to build arguments obtain benefits that are often associated with collaborative activity. The current research investigates whether these benefits can be attributed to the feedback from Convince Me's simulation model, that is, does the program serve as a "computer partner" in place of a "student partner." Students in four urban, ninth-grade Integrated Science classes used Convince Me either with or without model feedback. Half of the students in each group worked individually with the program and half worked in pairs. Results show that in attempting to "convince" Convince Me, students who receive feedback are encouraged to reflect on their reasoning strategies. Convince Me also appears to support reflection on argument construction and evaluation for pairs of students working together in the absence of feedback from the simulation model. (Contains 38 references.) (Author/ASK)

"Reasoner's Workbench" Program Supports Students' Individual and Collaborative Argumentation

Christine L. Diehl

Education in Mathematics, Science, and Technology
Graduate School of Education
University of California, Berkeley

Paper presented at the 2000 NARST Annual Meeting in New Orleans

Contact: Christine Diehl
4533 Tolman Hall
School of Education
UC Berkeley
Berkeley, CA 94720
cdiehl@socrates.berkeley.edu

"Reasoner's Workbench" Program Supports Students' Individual and Collaborative Argumentation

Abstract

This research evaluates the effectiveness of computer-mediated support for students' individual and collaborative argumentation. *Convince Me* is a "reasoner's workbench" program that aids students in generating and analyzing arguments, providing feedback on argument coherence from a general computational model. Laboratory studies indicate that students working individually with *Convince Me* to build arguments obtain benefits that are often associated with collaborative activity. The current research investigates whether these benefits can be attributed to the feedback from *Convince Me*'s simulation model, that is, does the program serve as a "computer partner" in place of a "student partner." Students in four urban, ninth-grade Integrated Science classes used *Convince Me* either with or without model feedback. Half of the students in each group worked individually with the program and half worked in pairs. Results show that in attempting to "convince" *Convince Me,* students who receive feedback are encouraged to reflect on their reasoning strategies. *Convince Me* also appears to support reflection on argument construction and evaluation for pairs of students working together in the absence of feedback from the simulation model.

Objectives

This research evaluates the effectiveness of a computer program, *Convince Me*, in supporting students' individual and collaborative argumentation. Results from laboratory studies indicate that *Convince Me* appears to be a useful tool for structuring and revising arguments. The current research investigates the role of the program's simulation model feedback and computer representations while testing *Convince Me*'s applicability in a Science curriculum. I hypothesize that students working individually without feedback from the simulation model will show the least improvement in argument evaluation. Further, that students working individually with model feedback will show similar performance to students working in pairs without model feedback, that is, that the program will serve as a "computer partner" in place of a "student partner."

INSTRUCTIONAL AND THEORETICAL PERSPECTIVES

Researchers describe many difficulties that individuals have with formal and informal reasoning (e.g., Kuhn, 1993; Linn & Songer, 1993), and some suggest reasoning skills that can and should be taught, such as: evaluation of information sources, hypothesis formation and evaluation, and discriminating between hypotheses and evidence (Koslowski, 1996; Perkins, 1985). Educational reformers also agree that higher-order thinking should be more than just a set of skills, but should provide opportunities for students to apply their knowledge in the real world (Stanley & Whitson, 1992). Knowledge is situated, being in part a product of the activity, context, and culture in which it is developed and used (Brown, Collins & Duguid, 1989; Lave & Wenger, 1991). Many educators claim that skills and knowledge in schools have become abstracted from their uses in the world. All too often in classrooms, science has been presented as an inert body of knowledge to be assimilated rather than a process of inquiry and a way to make sense of the world (Collins, Brown & Newman, 1989; Cognition and Technology Group at Vanderbilt [CTGV], 1990). Cobern contends that the problem is "that science education too often is not about developing *scientific* understanding but about understanding scientific *concepts* with the tacit assumption that scientific understanding will follow as a matter of course" (Cobern, 1996, p. 589; emphasis in source). There has been a shift in science instruction from a focus on facts and basic skills to student use of knowledge in complex goals, such as inquiry into domain knowledge (Linn, Songer & Eylon, 1996). Engaging students in explanation activities such as argument building can support the view of science as an ongoing process in which knowledge is sought, questioned and revised (Strike & Posner, 1992). Forming a coherent explanation and evaluating the coherence of an explanation are reasoning skills that students will apply in many situations in science as well as in their own lives, such as evaluating and applying theories, decision making, policy setting and evaluation, critical debate, and evaluating persuasive arguments.

Computer Argumentation Tools: Convince Me

The ECHO Educational Program (EEP) developed the *Convince Me* "reasoner's workbench" and associated curriculum to teach students reasoning skills for coherent argumentation. This is representative of the current trend in using computers to support "authentic" activities, that is, activities that are similar to the practice of experts, that involve complex tasks, and that are relevant to students' interest and goals (CTGV, 1996; Newman, 1997). Other researchers have developed courseware and curricula intended to help students' improve their reasoning skills through scaffolded argument construction. The CSILE environment is a discussion-based tool that supports the dialogical argumentation of a group of students as they engage in research on a topic (Scardamalia & Bereiter, 1993). SenseMaker and Belvedere are two knowledge representation tools like *Convince Me* that support the construction of rhetorical arguments by individuals (Bell, 1997; Cavalli-Sforza, Moore & Suthers, 1993). Each provides scaffolding and feedback in using evidence to support theories; however, *Convince Me* is still, to our knowledge, 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.

The *Convince Me* program is a user-friendly interface for using the ECHO model of the Theory of Explanatory Coherence (TEC) as an instructional tool. TEC attempts to account for how people decide the plausibility of beliefs asserted in an explanation or argument and is based on a few "hall of fame" principles of reasoning (see Appendix A; Thagard, 1992). ECHO is a connectionist computer model based on TEC (e.g., Ranney & Thagard, 1988). 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, 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 stabilizes, the nodes representing the most mutually coherent hypotheses and evidence exhibit high activation and may be regarded as accepted, and the 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.

The *Convince Me* 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" option provides feedback in the form of a correlation between (a) an individual's believability ratings for an argument's propositions and (b) ECHO's activations. *Convince Me* asks students to:

1) input their own situational beliefs,
2) categorize them as hypotheses or evidence,
3) indicate which beliefs explain or contradict which others,
4) rate their beliefs' plausibilities,
5) run the ECHO simulation, which predicts which of their beliefs "should" be accepted or rejected, based on the structure of their argument, and
6) contrast their ratings with ECHO's predictions.

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. 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.

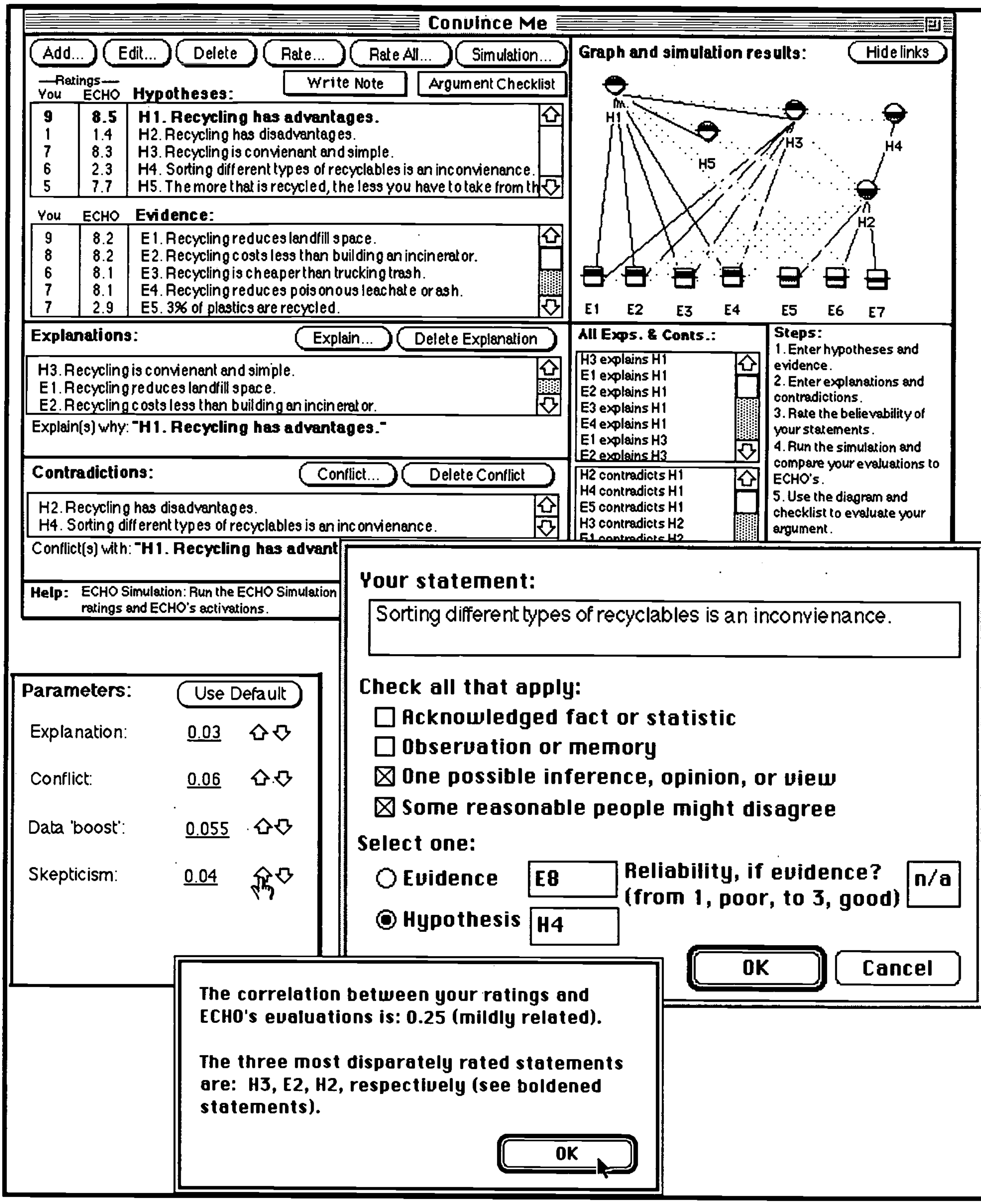


Figure 1. *Convince Me* interface with several interactive features pictured (parameter settings, statement editing dialog box, simulation model's (ECHO) fit feedback).

BEST COPY AVAILABLE

Assessing Convince Me

Studies investigating the prescriptive utility of *Convince Me* have assessed the impact that this system has on undergraduate students' ability to generate coherent arguments (Schank, Ranney, Hoadley, Diehl & Neff, 1994). An early study shows that the *Convince Me* system seems to make novice reasoners more like experts in their discrimination of hypothesis and evidence, 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, 1996). Use of the *Convince Me* system to aid students in generating and analyzing arguments has not been limited to scientific controvercies. A study of medical students reasoning about a neurophysiology problem with *Convince Me* versus think-aloud protocol showed that *Convince Me* arguments were more sophisticated (Weidner, Ranney, & Steinbach, in press). Students using *Convince Me* proposed more alternative hypotheses for the solution, provided more new, auxiliary, evidence to support hypotheses, produced more contradictory evidence, and indicated more contradictory relationships. *Convince Me* assessment has also not been limited to laboratory studies of undergraduate students. Siegel (1997) incorporated *Convince Me* as a component of a high-school science curriculum unit to determine if *Convince Me* activities would help students become better decision makers by implicitly teaching rules for decision making. The *Convince Me* program has been useful as a research method of investigating and modeling individual students' reasoning, and it appears to be a useful tool for structuring and revising arguments—and in helping to improve students' argument coherence. The interface and model feedback encourage students to explore more alternatives when reasoning about a problem and to include contradictory information in their arguments.

Collaborative Argumentation

Convince Me may serve as a forum for both individual and collaborative argumentation. The pedagogical implications of social learning theories suggest that students learn through a process of constructive engagement with the world—the world of people, of materials, of events, of ideas (Lave, 1991; Slavin, 1990). The cognitive apprenticeship model is an instructional application of this theory where teachers model cognitive and metacognitive skills and coach students as they appropriate the knowledge required to participate in authentic activities (Brown, Collins & Duguid, 1989). Another instructional application suggests creating "communities of practice" for learning, collaborative groups where the knowledge and skills to be learned are used in authentic tasks that support the learning and knowledge-use of individual students (Brown & Campione, 1994). The social construction of knowledge facilitates student learning by making their existing ideas explicit in order to evaluate these ideas with respect to the ideas of others (Scardamalia & Bereiter, 1993). These research applications are reflected in computer-mediated instructional environments like *Convince Me* which provide students with the means to construct and manipulate their own knowledge while being guided by the program and interacting with other students (Koschmann, 1994). The *Convince Me* program with it's model of coherent reasoning is designed to "make thinking visible" by modeling expert thinking, supporting individual reflections, and promoting the collaborative exchange of ideas (c.f., Collins, Brown & Holum, 1991). A study of students working on a scientific discovery task with a computer microworld found that students working in pairs participated more actively in explanatory activities, such as, entertaining hypotheses and considering alternative ideas and justifications (Okada & Simon, 1997). Research has shown that providing an explanation to oneself or to a partner can help a student generate elaborations between new and existing information and increase knowledge acquisition (Chi, de Leeuw, Chiu, & LaVancher, 1994). These benefits of collaborative activity are representative of the research results for students using the *Convince Me* program individually. Even when students are working individually with *Convince Me*, the program appears to provide a "social" context of reflection and explanation from which students benefit. Comments from students using *Convince Me* indicate that they reflect on their arguments as if they were explaining them to another student (Diehl, 1995).

> [Convince Me was the most useful exercise because] you are trying to convince a computer about your logical reasoning. It is fun because you'll get feedback from the computer and check to see how much you have convinced it.
>
> Convince Me was probably 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.
>
> [Convince Me] made me justify my thought processes and many times I had to question them and find other ways to approach a problem.

The research reported herein examines whether these benefits of collaborative activity can be attributed to feedback from *Convince Me*'s simulation model. It is possible that students working alone will not be as reflective in their argument construction in the absence of feedback on the coherence of their argument. It is likely, though, that students working collaboratively on an argument construction using *Convince Me* will use the program's representations (even without the simulation model's feedback) to negotiate a public interpretation of their individual understandings (c.f., Enyedy, Vahey & Gifford, 1997).

METHOD AND DATA

Participants

Students in four urban, ninth-grade, Integrated Science classes were assigned to four comparison groups: (1) students working individually without feedback from the simulation model (Indiv-No-Feedback), (2) students working individually with feedback from the simulation model (Indiv-Feedback), (3) students working in pairs without feedback from the simulation model (Pair-No-Feedback), (4) students working in pairs with feedback from the simulation model (Pair-Feedback). Two teachers participated in the study, and to make the teaching as similar as possible across groups, each teacher taught half the students in each of the four comparison groups. Each teacher taught one class with *Convince Me* using the simulation model's feedback and one class with *Convince Me* without the model's feedback. During the computer exercises, half of the students in each of the four classes worked individually and half worked in pairs. Students in a fifth Integrated Science class in the same school served as a control group who completed the same curriculum but did not use the *Convince Me* program at all.

Pre-tests

A questionnaire was administered to all students before the curriculum to score students' familiarity with scientific reasoning skills and their attitudes towards science. This test was adapted from measures on a variety of science epistemology surveys used in prior studies including: (a) CLP/KIE research group Beliefs about Science Test (Davis, 1998), (b) ThinkerTools Epistemology Assessment (White & Frederiksen, 1998), and (c) SEPUP Perceived Relevance of Science Questionnaire (Roberts & Henke, 1997). Students indicated their agreement with 20 statements from four categories: defining evidence, defining hypothesis, defining scientific reasoning, and relating science to real-world applications (see Appendix B for sample statements).

Convince Me Software

The *Convince Me* program was developed and run using HyperCard 2.0 on a Macintosh computer (Schank, Ranney & Hoadley, 1994). I developed two new features for the *Convince Me* program for this classroom use. The first is a new card that provides a checklist of strategies that students can use to evaluate their argument (e.g., "Have you supported each of your hypotheses with evidence?" "Are there any hypotheses that conflict with your statements?"). The student could request to see the checklist at any time by selecting the "Argument Checklist" button on the main page. The strategies on the checklist are organized under categories, such as "Hypotheses" and "Belief Ratings," and the student must select the category before the strategies are displayed (see Figure 2). The second new feature is a "notebook" window that appears at the bottom of the screen every time the student runs the simulation. The student is prompted to reflect on the believability of their argument. Students working with the simulation model are prompted to comment on the difference between their believability ratings and the simulation's ratings. This notebook is also available to the student upon request by selecting the "Write Note" button on the main page.

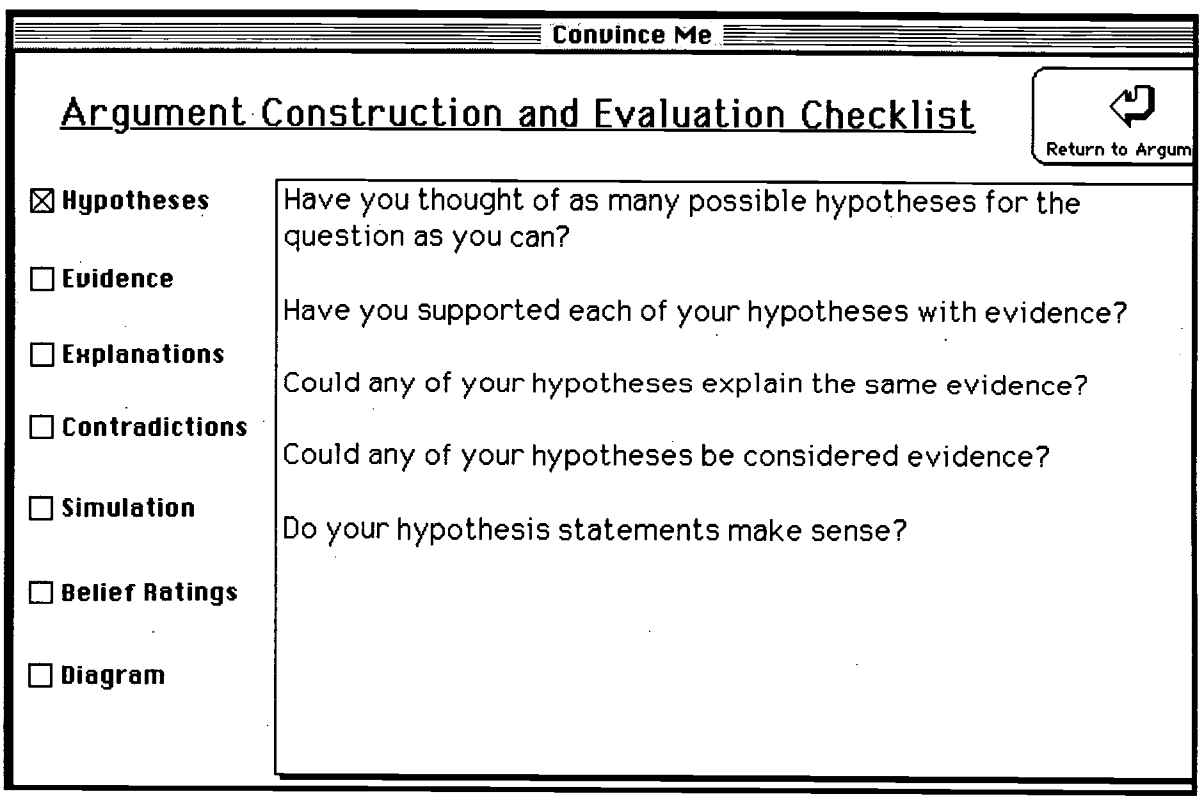


Figure 2. The new "Argument Construction and Evaluation Checklist" for *Convince Me*. Specific strategies are displayed when a student selects the category.

Curriculum

The *Convince Me* program was integrated into an existing four-week curriculum unit on waste management. Students were introduced to four methods of waste management and/or disposal, and they worked on activities to evaluate the cost and benefits of each method in relation to each other. Students worked for an hour each week to construct four arguments using the *Convince Me* program. They created one argument for each of three of the methods (landfills, incineration and recycling) and a final argument comparing all methods (including source management). Students engaged in classroom discussion, laboratory activities, and homework that provided the information necessary to construct each argument. Computer logs recorded the students' interactions with *Convince Me* while building and revising arguments. The log provides data on argument construction, argument revision, use of *Convince Me*'s program representations, final argument structure and coherence, and student reflections on their arguments and the model's feedback.

Post-tests

The scientific reasoning questionnaire was again administered to all students at the end of the curriculum. In addition, students in the four experimental groups completed a questionnaire where they indicated how useful they perceived each of the representations in the program's interface to be (e.g., diagram, dialog boxes; see Appendix C). A scientific argumentation knowledge test was also administered to these students to determine if students could evaluate the coherence of an argument and make changes to an argument to increase/decrease its believability (see Appendix D). All free-response answers were coded by two trained researchers and agreement for conflicting categorization (less than 5% of responses) was negotiated.

Data Analysis

Pre-tests were used to score students' knowledge of scientific reasoning, ability to evaluate arguments, and knowledge of the coherence principles underlying the simulation model. A regression analysis was used to determine differences among students working *individually with feedback* (Indiv-Feedback) from the simulation model, students working in *pairs with feedback* (Pair-Feedback), students working *individually without feedback* (Indiv-No-Feedback), students working in *pairs without feedback* (Pair-No-Feedback), and students with no exposure to the computer program (Control). The computer logs detail which features of the program students used in the construction and evaluation of arguments. Once again, a regression analysis was used for comparing the performance of the four groups in using the *Convince Me* program representations (e.g., differences in model's fit, number and type of changes to argument). The post-tests allow a comparison of changes in the reasoning skills of all the groups. Data from the control group of students is used to determine if gains on the post-tests are due to the use of *Convince Me* or are attributable to the curriculum unit on recycling. In addition, the post-tests assess students understanding of the *Convince Me* program and their attitudes towards the program's argument representations, which enables the correlation of students' understanding of the argument representations with their actual use in the program.

Due to high incidence of absenteeism, 25 out of 127 students were dropped from the data analysis. To be included in the study, a student must have attended at least three out of four days when the program was in use and at least four out of six days when the program was not in use (or at least seven out of ten days for students in the control group). The mean attendance rate for students included in the study was 90%, and for students dropped from the study the mean attendance rate was 39%.

RESULTS AND CONCLUSIONS

Scientific Reasoning

There were no significant differences between any of the groups for the Scientific Reasoning Questionnaire scores prior to the curriculum. However, there was a significant difference among groups on the post-curriculum scores for this test (F =4.31, $p < .01$; see Figure 3). From an analysis of pair-wise comparisons we can conclude that the significant differences lie between the scores of the Control and No-Feedback groups.

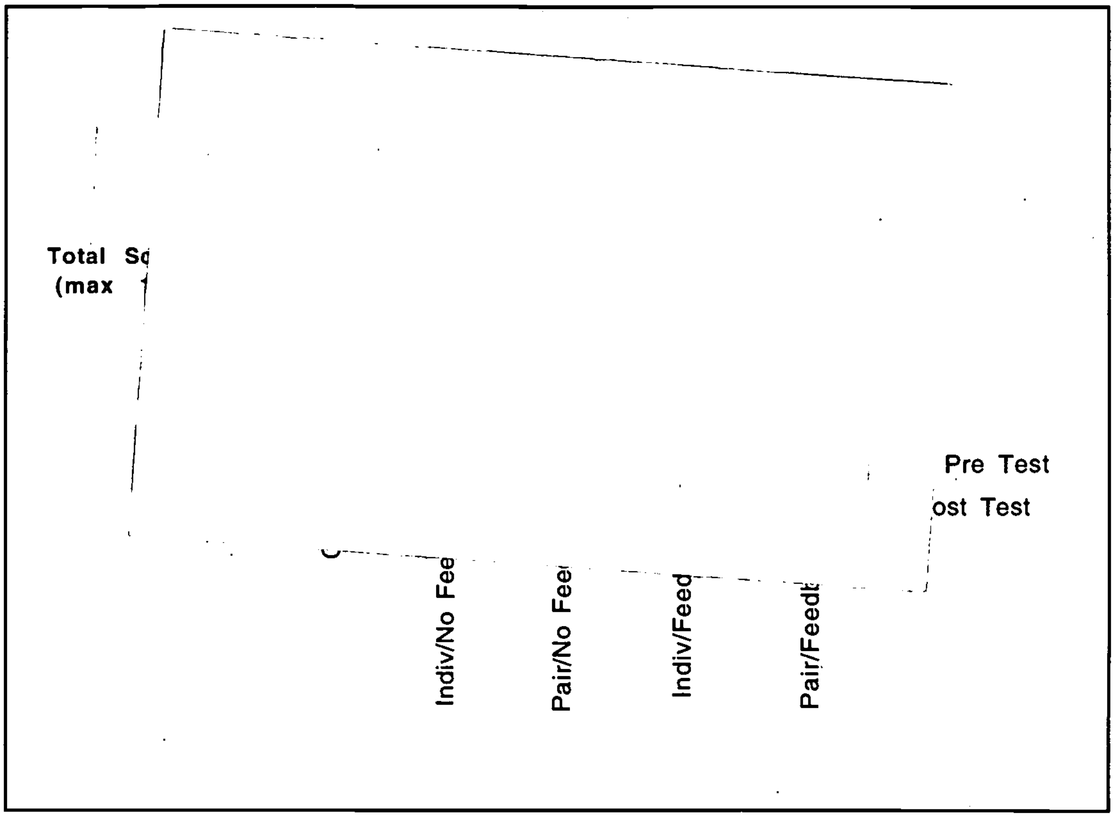


Figure 3. Scientific Reasoning Questionnaire Total Scores. This chart shows the pre- and post-test scores for each experimental group.

The mean score on the Scientific Reasoning Questionnaire for students in the Control and Pair-Feedback groups actually decreased slightly while the mean score for the other three groups increased. A detailed analysis shows that the students in the Pair-Feedback group did not develop less normative ideas of scientific reasoning, rather they often decreased the certainty of their agreement with normative responses (e.g., they responded that they "somewhat agreed" with normative responses vs. "strongly agreed"). The four experimental groups all exhibited more normative scientific reasoning skills on the post-test questionnaire; however the two No-Feedback groups scored significantly higher than the two Feedback groups ($F = .63$, $p < .01$, mean 76.5 vs. 72.1). On further analysis, the only significant difference that appears in the response of the No-Feedback versus Feedback groups was actually for the questions in the category of defining scientific reasoning ($F = 8.6$, $p < .01$, mean 19.25 vs. 17.7). The students who received feedback from the simulation model on the coherence of their arguments, and who attempted to revise their arguments to convince the computer of their reasoning (see the discussion on argument revisions below), appear to have developed a more critical attitude towards evaluating scientific arguments.

Argument Revisions

For each computer-based argument building exercise, students were given a chance to revise their argument after reflecting on (a) which side of their argument was more believable, and (for Feedback groups) (b) which side of their argument the simulation model thought was more believable. There was a significant difference among the groups in the type of argument revisions made, in particular, whether students made changes to the structure of their existing argument (intrinsic changes) versus adding new information to the argument (extrinsic changes) (see Figure 4). Intrinsic argument changes include changing the categorization of a hypothetical or evidential proposition, changing the reliability rating of evidence, changing the believability rating of a proposition, or adding contradictions or explanations between existing propositions. Extrinsic argument changes include adding new hypothetical or evidential propositions, or adding contradictions or explanations linked to new propositions. Students in the No-Feedback groups were more likely to add new information to their arguments—in particular, explanatory evidence—than were students in the Feedback groups (F = 3.72, p<.05). Students working together were also more likely to make extrinsic changes to their argument than students working individually (F = 8.45, p < .01). Many more students in the Feedback groups revised their existing argument structure (72%) than did students in the No-Feedback groups (23%) (F = 27.74, p < .01)—the most common revisions were changing the category of a proposition from hypothesis to evidence (or vice versa) and changing the explanatory or contradictory relations among the propositions in an argument. Feedback from both the computer model and a student partner appears to encourage students to more critically evaluate their arguments and make changes to them.

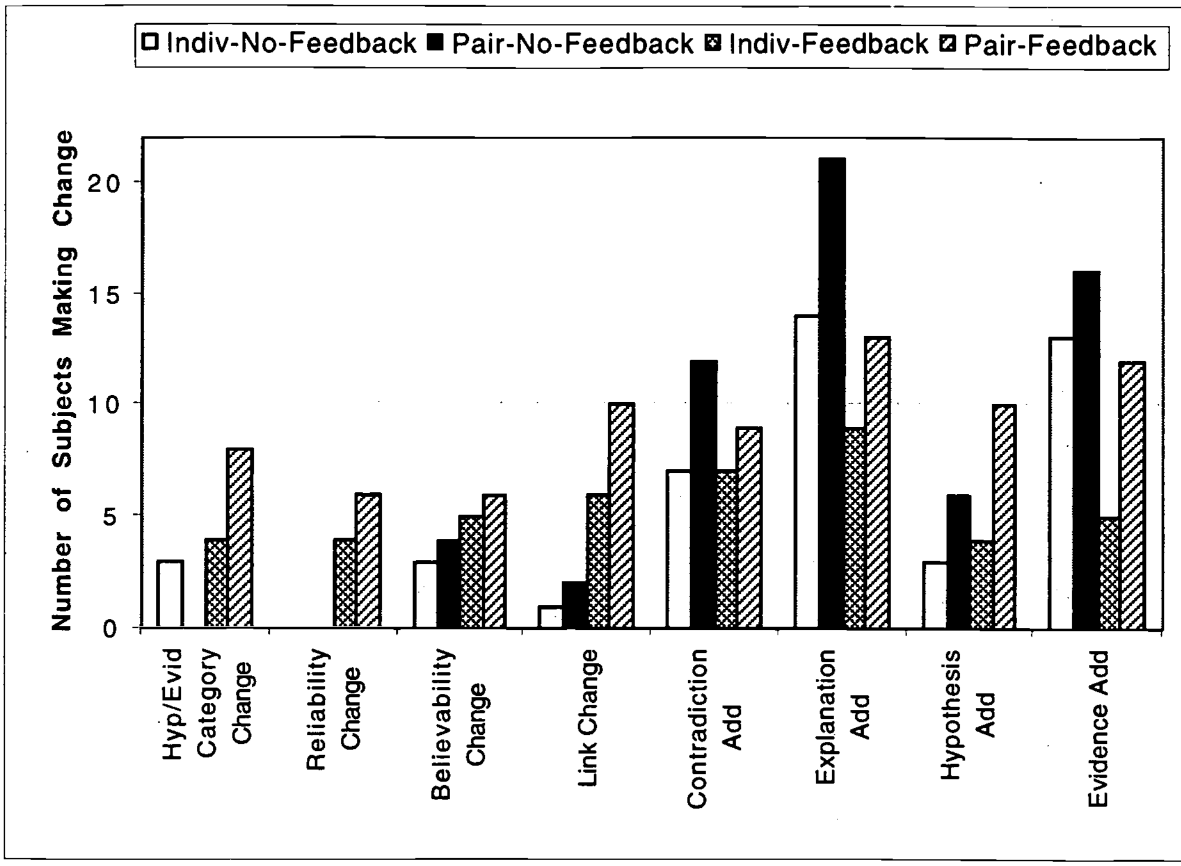


Figure 4. Argument Revisions. This chart shows the types of changes students in each experimental group made to their third argument on the topic of recycling.

Argument Coherence—"Model's Fit"

A "model's fit" measure was obtained for each students' initial and revised arguments which reflects how well the students' argument structure seems to match their beliefs. This measure of argument coherence is a correlation between a student's believability ratings and the simulation model's predictions of the propositions' plausibilities (as derived from the model's final activations). The higher the overall correlation, the more the simulation agrees with the students' belief ratings—based on their argument. Students in the Feedback groups received the model's fit value as feedback from the simulation model. For students in the No-Feedback group, the model's fit values were obtained by the researcher posthoc. There was no significance difference among the four groups' model's fit values on the initial arguments. However, after given the opportunity to reflect on and revise their arguments, the model's fit for the Feedback groups was significantly higher than that for the No-Feedback groups ($F = 4.94$, $p < .05$). As can be seen in Figure 5, there was very little change in the correlation between students' belief ratings and the simulation's ratings for the No-Feedback groups after their argument revision. An increase in model's fit in a student's revised argument was significantly correlated with Intrinsic changes to argument structure ($F = 5.93$, $p < .05$). Students working in pairs with feedback from the simulation model had the greatest improvement in argument coherence as measured by the belief correlation value represented by the model's fit.

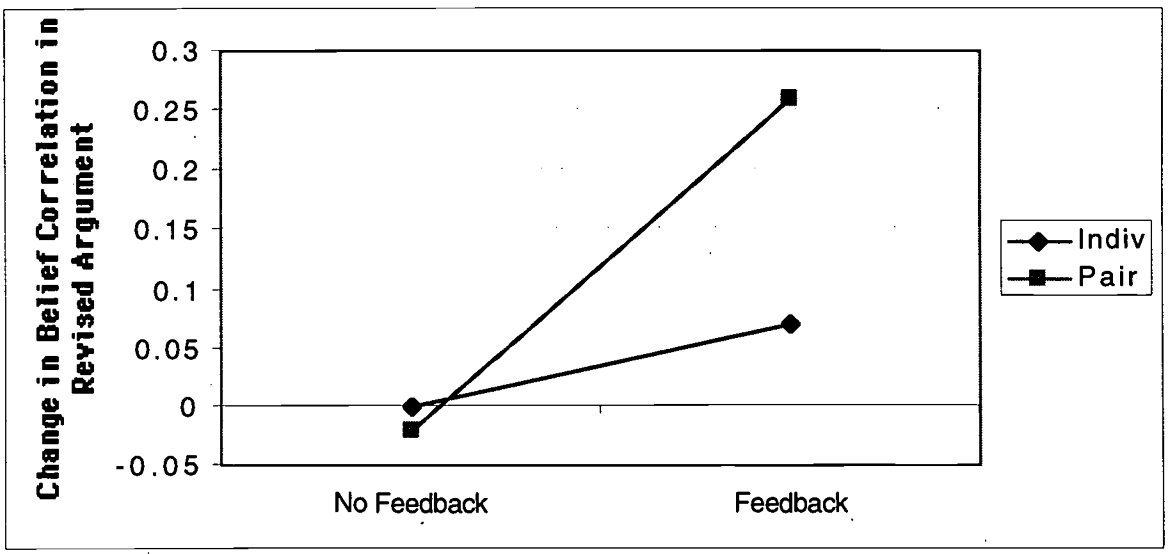


Figure 5. Change in Model's Fit. This chart represents the difference between the model's fit value for the initial and revised arguments for each experimental group.

Computer Representations

Prior research has shown that student performance and reasoning strategies are influenced by both pragmatic (i.e., helping to use the program) and conceptual (i.e., helping to build an argument) attributes of *Convince Me*'s program representations. In this study, students' descriptions of the usefulness of the various representations were coded as either "not-useful," "pragmatic," or "conceptual" (see Appendix E for sample coded responses). The categorization of each representation was strongly correlated with the usefulness rating that students assigned the representation—e.g., students who attributed conceptual "coaching" to a particular representation rated it as more useful than students who attributed pragmatic "guidance" to the representation ($F = 157.2, p < .01$). The conceptual "coaching" attributed to the program representations varied among the four groups ($F = 3.27, p < .05$). Both feedback from the simulation model and working in pairs (even without model feedback) resulted in students reporting more conceptual benefit from the program.

Scientific Argumentation

Questions on the scientific argumentation knowledge test were scored for correctness and categorized by the source of knowledge used to answer the question. Responses for each question were categorized as (a) "structure" if they referred to the structure of the argument, (b) "content" if they referred to specific information given in the argument, and (c) "real life" if they referred to personal knowledge or experience beyond the information given in the argument (see Appendix F for sample coded responses). Mean scores (for correctness) were higher for both students working with feedback from the simulation model and students working in pairs (even without model feedback) ($F = 3.37$, $p < .05$). The source of knowledge used to answer the questions varied among the four groups ($F = 4.17$, $p < .05$) with feedback from the simulation model resulting in students reflecting more on the structure of the argument in their responses. Students working alone without feedback were more likely to respond with real-life explanations than any other group of students ($F = 4.7, p < .01$).

EDUCATIONAL IMPORTANCE

"Learning to think" has been called the central purpose of education. The emphasis on thinking skills in the federal Goals 2000 framework and state curriculum frameworks is representative of the current advocacy of teachers, administrators, and policy makers for the adoption of a *four Rs* curriculum supporting the foundational skills of reading, writing, arithmetic, and reasoning. There is little agreement on what skills a *thinking curriculum* should highlight, but likely candidates include metacognition, argumentation skills, and the evaluation of information sources. The project described in this paper supports scientific reasoning in the classroom with computer-mediated instruction. Computer-mediated instruction has the potential to change the social context in which it is used; for example, to create the opportunity and the content for discussion among students or between teachers and students (Rubin, 1996). Computer environments can support collaboration by providing opportunities for explanation and reflection and by helping students keep track of and refer to ideas under discussion. *Convince Me* aids students in generating and analyzing arguments, providing feedback from a general computational model. Results show that in attempting to "convince" *Convince Me,* students who receive feedback from the simulation model are encouraged to reflect on the structure of their arguments and their reasoning strategies. *Convince Me* also appears to support reflection on argument construction and evaluation for pairs of students working together, even in the absence of feedback from the simulation model. Analysis is continuing on the arguments students constructed, the logs of student interactions with the program, and videotapes of students working with the program to determine the pragmatic, explanatory and reflective role of student versus computer partner.

ACKNOWLEDGEMENTS

This research was supported by an AERA/Spencer Pre-Dissertation Fellowship, an NSF Science and Design Traineeship Grant, and a University of California Fellowship. I would like to thank all of the members of the Reasoning Group and the SAND Research Group for their helpful insights and support during the design, implementation, and analysis of this research. I would especially like to thank my dissertation committee members Michael Ranney, Bernard Gifford, Barbara White, and Alison Gopnik for their invaluable assistance.

REFERENCES

Bell, P. (1997). Using argument representations to make thinking visible for individuals and groups. In R. Hall, N. Miyake, & N. Enyedy (Eds.), *Proceedings of the Computer Support for Collaborative Learning Conference '97* (pp. 10-19). Toronto, Canada: University of Toronto.

Brown, A. L. & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed), *Classroom Lessons: Integrating Cognitive Theory and Practice* (pp. 229-270). Cambridge: MIT Press.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

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.

Chi, M.T.H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.

Cobern, W.W. (1996). Worldview theory and conceptual change in science education. *Science Education, 80*(5), 579-610.

Cognition and Technology Group at Vanderbilt [CTGV] (1996). Looking at technology in context: A framework for understanding technology and education research. In D.C. Berliner, & R.C. Calfee (Eds.) *Handbook of educational psychology* (pp. 438-490). New York, NY: Simon & Schuster Macmillan.

Collins, A., Brown, J.S. & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator, 6*(11), 38-46.

Collins, A., Brown, J.S. & Newman, S.E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. Resnick (Ed.), *Knowing, learning, and instruction* (p.453-495). Hillsdale, NJ: Erlbaum.

Davis, E. (1998). *Scaffolding Students' Reflection for Science Learning.* Unpublished doctoral dissertation. University of California, Berkeley, CA.

Diehl, C.L. (1995). Assessing Graphical and Propositional Tools for Enhancing Reasoning Skills. Unpublished manuscript. University of California, Berkeley, CA.

Enyedy, N., Vahey, P. & Gifford, B. (1997). Active and supportive computer-mediated resources for student-to-student conversations. In R. Hall, N. Miyake, & N. Enyedy (Eds.), *Proceedings of the Computer Support for Collaborative Learning Conference '97* (pp. 27-36). Toronto, Canada: University of Toronto.

Koschmann, T. D. (1994). Toward a theory of computer support for collaborative learning. *The Journal of the Learning Sciences*, *3*(3), 219-225.

Koslowski, B. (1996). *Theory and Evidence: The Development of Scientific Reasoning.* Cambridge, MA: MIT Press.

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

Lave, J. (1991). Situated learning in communities of practice. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 63-82). Washington, D.C.: American Psychological Association.

Lave, J. & Wenger, E. (1991): *Situated learning: Legitimate peripheral participation* New York: Cambridge University Press.

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.

Linn, M.C., Songer, N.B., & Eylon, B. (1996). Shifts and convergences in science learning and instruction. In D.C. Berliner, & R.C. Calfee (Eds.) *Handbook of educational psychology* (pp. 438-490). New York, NY: Simon & Schuster Macmillan.

Newman, D. (1997). Functional environments for microcomputers in education. In M. Cole, Y. Engestrom, & O. Vasquez (Eds.), *Mind, culture, and activity* (pp. 279-291). Cambridge, UK: Cambridge University Press.

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

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

Ranney, M., Schank, P., & Diehl, C. (1996). Competence versus performance in critical reasoning: Reducing the gap by using Convince Me. *Psychology Teaching Review, 4*(2), 153-166.

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.

Roberts, L., & Henke, R.R. (1997). Mapping middle school students' perceptions of the relevance of science. In M. Wilson (Ed.), *Objective Measurement, Vol. 4*. Norwood, NJ: Ablex Publishing Corp.

Rubin, A. (1996). *Educational Technology: Support for Inquiry-based Learning.* Unpublished manuscript. MA: TERC.

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.

Scardamalia, M., & Bereiter, C. (1993). Computer support for knowledge-building communities. *The Journal of the Learning Sciences, 3*(3), 265-283.

Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice.* Englewood Cliffs, NJ: Prentice-Hall.

Stanley, W.B., & Whitson, J.A. (1992). Citizenship as practical competence: A response to the new reform movement in social education. *International Journal of Social Education, 7*(2), 57-66.

Strike, K.A. & Posner, G.J. (1992). A revisionist theory of conceptual change. In R.A. Duschl & R.J. Hamilton (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (pp. 147-176). New York: State University of New York Press.

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

Weidner, J., Ranney, M., and Steinbach, A. (in press). Using *Convince Me* to assess medical reasoning skills (and vice versa). To appear in the Proceedings of the International Conference for the Learning Sciences.

White, B.Y. & Frederiksen, J.R. (1998). Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. *Cognition and Instruction, 16*(1), 3-118.

APPENDIX A

Some principles that underlie the Theory of Explanatory Coherence.

Symmetry: Coherence and incoherence are symmetric relations. This means that if one belief explains (or conflicts with) another, the beliefs "send activation" back and forth to each other.

Explanation: A belief that explains a proposition coheres with it. Also, beliefs that jointly explain a proposition cohere with it, and cohere with each other.

Data Priority: Results of observations, such as evidence, have an extra measure of acceptability.

Simplicity: The plausibility of a proposition is inversely related to the number of explaining statements needed to explain it. That is, lots of assumptions (or co-beliefs) are often counterproductive, compared to fewer assumptions.

Contradiction: Contradictory hypotheses incohere. This means that beliefs that conflict with each other send "negative activation" to each other.

Acceptability: The acceptability of a proposition depends on its coherence within the system of propositions in which it is embedded. The acceptability of a proposition increases as it coheres more with other acceptable propositions, and vice versa.

Overall Coherence: The overall coherence of a network of propositions depends on the local pairwise cohering of its propositions.

APPENDIX B

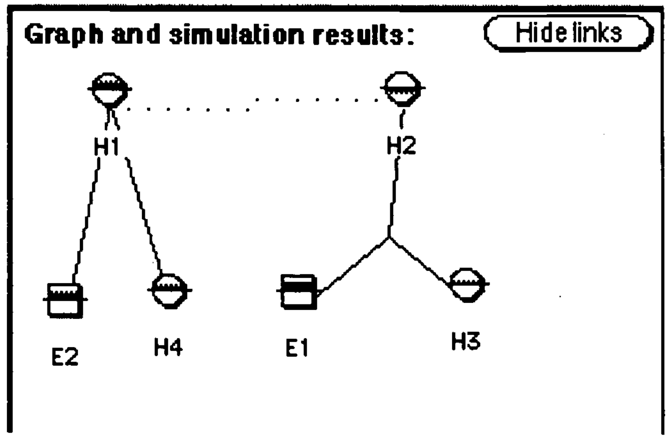
Scientific Reasoning Questionnaire (sample statements)

DIRECTIONS: *For each of the following items, please read the statement, and check the answer that best describes how strongly you agree or disagree.*

	strongly disagree	somewhat disagree	in the middle	somewhat agree	strongly agree
Evidence is used to support a hypothesis.	❐	❐	❐	❐	❐
Explaining why or how things happen is important in science class but not in my life outside school.	❐	❐	❐	❐	❐
Some hypotheses do not accurately represent reality.	❐	❐	❐	❐	❐
A bad argument can still have a lot of evidence supporting a hypothesis.	❐	❐	❐	❐	❐

APPENDIX C

Representations Questionnaire (sample exercise)



1. I liked using this feature of the program.

strongly disagree	somewhat disagree	in the middle	somewhat agree	strongly agree
☐	☐	☐	☐	☐

2. I understand how to use this feature of the program.

strongly disagree	somewhat disagree	in the middle	somewhat agree	strongly agree
☐	☐	☐	☐	☐

3. This feature of the program was useful.

strongly disagree	somewhat disagree	in the middle	somewhat agree	strongly agree
☐	☐	☐	☐	☐

How was this feature of the program useful?

____

____

4. I learned a lot from this feature of the program.

strongly disagree	somewhat disagree	in the middle	somewhat agree	strongly agree
☐	☐	☐	☐	☐

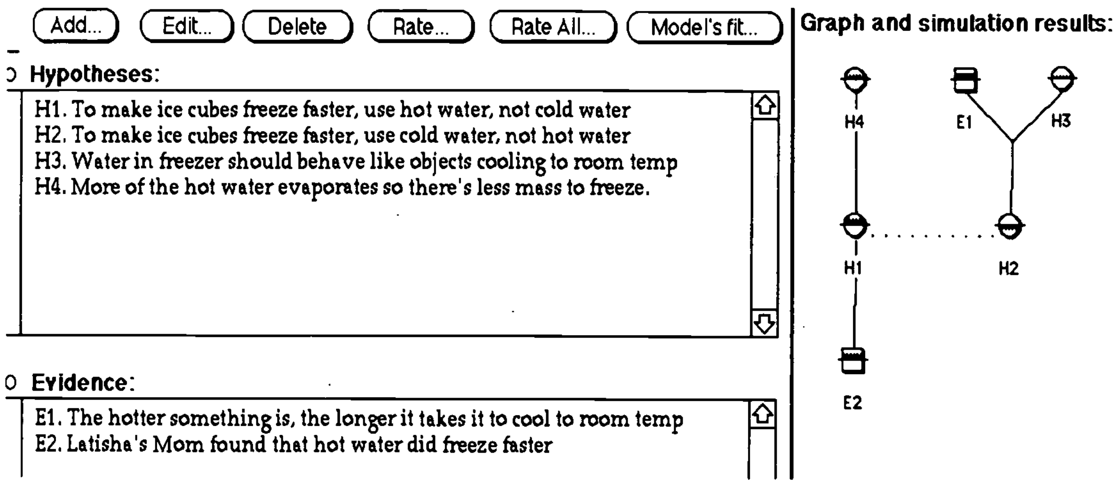
What, if anything, did you learn from this feature of the program?

____

____

Appendix D

Scientific Argumentation Questionnaire



Sample questions for students who received model feedback

1. Which side of the argument, H1 or H2, would the Convince Me simulation (ECHO) think is stronger? Why?
2. If I add a piece of evidence that conflicts with H1, would the Convince Me simulation (ECHO) think H1 is more or less believable? Why?
3. Describe two specific changes you could make to the argument to make the Convince Me simulation (ECHO) believe H2 more.

Sample questions for students who *did not* receive model feedback

1. Which side of the argument, H1 or H2, do **you** think is stronger? Why?
2. If I add a piece of evidence that conflicts with H1, would you think H1 is more or less believable? Why?
3. Describe two specific changes you could make to the argument to make H2 more believable.

APPENDIX E

Sample Coded Responses from Representations Questionnaire

PRAGMATIC CATEGORY

- The [diagram] is useful because it connects to all evidence or hypothesis that goes together. Is easier to tell which statement is connected to the other.
- [The add a proposition dialog box] is kind of useful because I don't need to write out the statement, just type it in.
- [The relational listing] is useful because it shows you how the information that is here gets to the diagram.
- [The argument checklist] was useful because it's a checklist to see what you have accomplished.

CONCEPTUAL CATEGORY

- The diagram helped me organize my ideas, unlike the another feature, which just listed the explanations and contradictions. I learned that several evidences can support a hypothesis and that a hypothesis can support another hypothesis, too.
- [With the add a proposition dialog box] I learned that the quality and reliability of an evidence is important.
- [The ECHO simulation model feedback] helped me figure out what was wrong with my evidence and hypothesis.
- [The relational listing] help me put an evidence to my hypothesis. I learn how to contradict hypothesis and evidence.
- [The argument checklist] helps me to make sure if we're sure that are statements are in the correct categories.

APPENDIX F

Sample Coded Responses from Scientific Argumentation Questionnaire

STRUCTURE CATEGORY

- You could either add more evidence supporting H2 or delete evidence that supports H1.
- Because H2 is more because right now there's like a tie in H1 & H2 but if you add a conflict to H1, then of course the ECHO would agree on H2 if the answer's are right.
- Less believable because H2 would have more evidence and the computer only knows what you tell it.

CONTENT CATEGORY

- I think H1 because somebody had tried it and found out that is evidence.
- [Add] better evidence--more than one sentences about water cycle.
- [I believe] H1 because of E2.

REAL LIFE CATEGORY

- I think they would pick #2 because it is faster if you use cold water. I try using hot water to make ice cube before, and it takes longer to freeze than the cold water.
- Let five people try it at home and see how much or which is more to freeze faster by checking it every 15-20 mins.