Student Goal Orientation in Learning Inquiry Skills With Modifiable Software Advisors

TODD A. SHIMODA
Colorado State University

BARBARA Y. WHITE, JOHN R. FREDERIKSEN
University of California, Berkeley, USA

Received 1 January 2000; revised 8 December 2000; accepted 15 March 2001

ABSTRACT: A computer support environment (SCI-WISE) for learning and doing science inquiry projects was designed. SCI-WISE incorporates software advisors that give general advice about a skill such as hypothesizing. By giving general advice (rather than step-by-step procedures), the system is intended to help students conduct experiments that are more epistemologically authentic. Also, students using SCI-WISE can select the type of advice the advisors give and when they give advice, as well as modify the advisors' knowledge bases. The system is based partly on a theoretical framework of levels of agency and goal orientation. This framework assumes that giving students higher levels of agency facilitates higher-level goal orientations (such as mastery or knowledge building as opposed to task completion) that in turn produce higher levels of competence. A study of sixth grade science students was conducted. Students took a pretest questionnaire that measured their goal orientations for science projects and their inquiry skills. The students worked in pairs on an open-ended inquiry project that requires complex reasoning about human memory. The students used one of two versions of SCI-WISE—one that was modifiable and one that was not. After finishing the project, the students took a posttest questionnaire similar to the pretest, and evaluated the version of the system they used. The main results showed that (a) there was no correlation of goal orientation with grade point average, (b) knowledge-oriented students using the modifiable version tended to rate SCI-WISE more helpful than task-oriented students, and (c) knowledge-oriented pairs using the nonmodifiable version tended to have higher posttest inquiry skills scores than other pair types. © 2002 Wiley Periodicals, Inc. *Sci Ed* **86:**244–263, 2002; DOI 10.1002/sce.10003

INTRODUCTION

Learning science includes not only understanding concepts (such as friction or natural selection) but also understanding the process of scientific inquiry. Such inquiry skills include scientific methods, as well as social and collaboration skills. As part of most science curricula, students are required to perform experiments. However, as Chinn and Malhotra

Correspondence to: Todd A. Shimoda
Contract grant sponsor: Department of Education's Office of Educational Research and Improvement.
Contract grant sponsor: McDonnell Foundation.
Contract grant sponsor: Educational Testing Service.

© 2002 Wiley Periodicals, Inc.

(in press) argue, many experimental tasks used in schools oversimplify the process and students may learn a form of reasoning that is not like authentic scientific reasoning.

In our study, middle-school science students were given a research question about human memory, which is based on the depth-of-processing theory (Craik, 1979): "How can people best remember a list of words, for example, saying each word five times, or using each word in a sentence?" Our students were given only general inquiry project steps to complete (such as developing hypotheses, designing an investigation, analyzing data, and coming up with a general model to explain the data). Students were given general advice about the inquiry steps, but were not given step-by-step procedures to follow. Investigating the research question about memory requires students to do many of the cognitive processes (largely on their own) that Chinn and Malhotra (in press) identify as those required in authentic research, such as selecting and controlling variables, planning procedures, making observations, finding flaws, and developing theory.

To help students learn more authentic and complex scientific inquiry, White and Frederiksen (e.g., Frederiksen & White, 1998a, 1998b; White & Frederiksen, 1998) have focused on teaching science concepts through modeling, inquiry, and metacognition. Their inquiry cycle and curriculum introduces students to the cyclical nature of scientific inquiry: coming up with interesting research questions, developing hypotheses, designing and conducting investigations, analyzing data, developing models or theories to explain the data, evaluating the models, and coming up with new research questions based on the evaluations. Within each of these skills, students are introduced to the subskills involved and prompted to evaluate their use of the skills and their products (e.g., if their experiment design controls for possible confounding variables). An inquiry-based curriculum, then, has students taking control of much of their learning by providing them relatively open-ended questions and metacognitive scaffolding.

Each student comes to such a learning situation with unique levels of prior knowledge, beliefs, interest, abilities, motivation, and learning goal orientations among other factors. These factors interact with the learning task and the educational setting, and influence how the student learns the target concepts and skills. Of these factors, student goal orientation has been shown to be a strong predictor of learning performance (Ames & Archer, 1988; Cain & Dweck, 1989; Ng & Bereiter, 1995).

Learner goal orientation is an individual's disposition toward a level of achievement or target for accomplishment in a given learning situation. Researchers have noted differences in students' goal orientations that are classified by various terms such as Piaget's "success" vs. "understanding" (Piaget, 1978), Ames and Archer's "performance" vs. "mastery" (Ames & Archer's, 1988), and Dweck's "ability" vs. "mastery" (e.g., Cain & Dweck, 1989). Students with performance or ability orientations tend to view learning situations as normative evaluations, that is, comparing one's performance with others' and assigning differences to ability. Students with mastery orientations tend to view learning situations as opportunities to improve and make progress. For example, Ng and Bereiter (1995) identified three levels of goal orientations toward learning during a self-directed sequence of computer programming instruction: (a) *task-completion* goals, such as writing a GO TO statement; (b) *instructional* goals, or pursuing the manifest objectives of the course such as understanding how GO TO statements work; and (c) *knowledge-building* goals that are more personally motivated, as in applying GO TO statements in new and personally relevant ways.

Different goal orientations result in different behaviors and outcomes, including goal setting, performance, cognitive resource allocation, and affective reactions. In goal setting behavior, Ng and Bereiter (1995) showed that students with instructional and knowledge-building goal orientations were more likely to set learning goals of "understanding

something" or "mastering something" than the students with task completion goal orientations. Those in the latter category tended to select goals such as "do" or "complete" something.

Research has shown that students with higher goal orientations (i.e., mastery or knowledge building) perform at higher levels. For example, studies by Cain and Dweck (1989) showed that those with a mastery-orientation maintain or improve performance in the face of failure, while those with a performance orientation show behavior deterioration in the face of failure. These studies also showed that eighth graders who had learning goals in a science unit, compared to those who had performance goals, (a) attained significantly higher scores on the transfer test; (b) produced about 50% more work on a transfer task, suggesting that they were more active in the transfer process; and (c) produced more rule-generated answers on the test (indicating a better comprehension of the rules they learned).

Ng and Bereiter (1995) speculate that helping students work at higher level goals (i.e., knowledge building) would be beneficial to their learning. Ng and Bereiter claim, however, that observed classroom practices focus predominately on task completion goals. Much instructional research in the past was devoted to understanding how one could teach the domain specific skills underlying the tasks. Currently, much instructional research focuses on social and cognitive aspects, particularly collaborative learning and metacognition, and other higher-order thinking skills.

Scardamalia and Bereiter (1991) also argue for increasing levels of agency in instruction. Levels of agency refers to the amount of control that students exercise of their own intentional learning. In Vygotskian philosophy (Vygotsky, 1978), levels of agency increase when the students assumes more control of their zone of proximal development (ZPD), which is the development that could occur given an individual's current development and appropriate learning opportunities. Scardamalia and Bereiter argue that students have access to the ZPD and can help move themselves through it. For example, Scardamalia and Bereiter's studies with text-based learning show that children can generate educationally productive questions based on their perceived knowledge needs when given the agency to do so. These knowledge-based questions facilitate deeper, more personal learning than text-based questions, which are directed at the surface understanding of a specific passage. An example of a knowledge-based question that the students identified as productive during a discussion of the environment is "Why must we choose forests and countrysides to build more and more cities. Don't we have enough cities?"

Figure 1 presents a theoretical framework that integrates dimensions of agency, competence, and goal orientation. This framework implies that higher levels of agency facilitate higher-level goal orientations that in turn produce higher levels of competence. Toward the left end of the goal orientation layer is "task completion," or just getting the task done without regard to performance. (There could be a marker further to the left—"do nothing.") Next along the continuum is "performance," which is an orientation toward setting some level of achievement for a task, for example, doing a good job, or just good enough to pass. The next marker is the "mastery or learning" orientation that includes understanding and achieving the cognitive and social goals embodied in the task. For example, mastering inquiry-style research projects involves cognitive abilities such as reasoning carefully and social skills such as collaborating. "Personal knowledge building" refers to the orientation toward developing one's own ability to learn that goes beyond the task.

The levels of agency and goal orientation framework predicts that students doing inquiry projects would have different performance outcomes based on their goal orientations. The framework indicates many features of a computer-based educational technology including a system that can adapt to students with different goal orientations, and one that can be modified by the student users (increasing agency). Such a system should be general enough

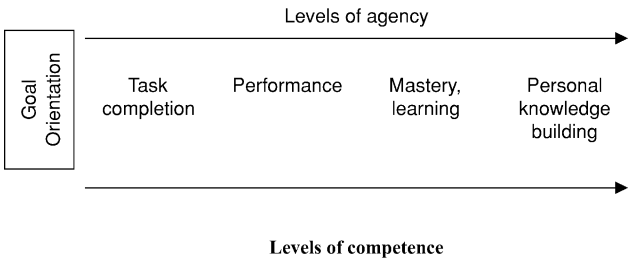

Figure 1. A model of agency, goal orientation, and competence.

to allow students to use it for personal knowledge building goals that cannot be predicted a priori.

THE SCI-WISE SYSTEM

The goals of the ThinkerTools SCI-WISE system (Social and Cognitive Improvement Within an Inquiry Support Environment, among other meanings) are to facilitate the acquisition (or improvement) of inquiry skills and the acquisition of reflective, meta-scientific ability to build general theories of the inquiry process and learning in general (Shimoda, White, & Frederiksen, 1999; White, Shimoda, & Frederiksen, 1999).

In general SCI-WISE consists of knowledge-based software advisors that provide contextual advice and other information. The advisors can also prompt students to do activities such as reflect on or assess their work. The advisors can also be modified by the students so they can meet personal knowledge building goals or learning styles. The students and advisors work in an environment consisting of a Project Journal that records the students' work, a Progress Report that keeps track of the students' general progress through tasks, a Meeting Room where advisors can be accessed by type, and a Dialogue Box that facilitates communication between students and advisors.

An in-depth discussion of the design of SCI-WISE will not be presented in this paper. For more information refer to our other papers (Shimoda, White, & Frederiksen, 1999; White, Shimoda, & Frederiksen, 1999). A brief discussion that introduces the main aspects of SCI-WISE follows.

Advisors

Advisor Types. There are three broad classes of software advisors in SCI-WISE that correspond to learning goal orientations (shown in parentheses): *task advisors*, which help students complete the steps of the inquiry process and other tasks (task completion and performance orientation), *general purpose advisors* for help in understanding and using the underlying skills embodied in the inquiry tasks (learning and mastery orientations), and *system development advisors*, which help students modify the system for their own learning goals (personal knowledge building). These are not the only classes and types of advisors that can be useful (see White, Shimoda, & Frederiksen, 1999, for additional examples), but they satisfy a model of reflective, goal-driven inquiry as proposed by White and Frederiksen (1998) and fit the levels-of-agency model shown in Figure 1. Figure 2 illustrates a taxonomy of the three classes along with specific advisors in those classes.

Advisor Knowledge. The advisors in the prototype SCI-WISE system are knowledge-based, and are intended to provide general guidance that might apply in many content areas

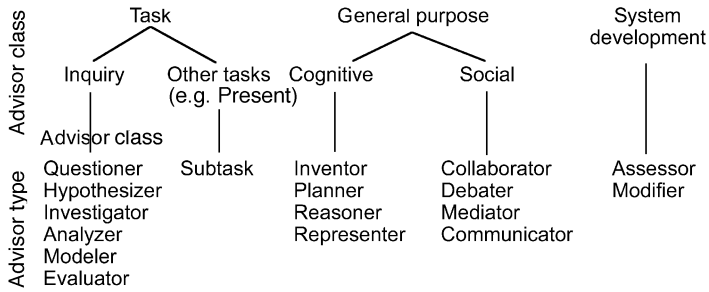

Figure 2. SCI-WISE advisor classes and types.

including those in the physical and biological sciences, as well as in the social sciences. The advisors have general domain knowledge (i.e., the Questioner's domain is research questions) such as criteria (e.g., what makes a good research question), strategies (e.g., how to come up with a good research question), examples, concepts (definitions), and referrals to other advisors. An example of a referral could be the Hypothesizer referring the student to the Investigator for advice on what "testable" hypothesis means. This type of knowledge is intended to help students grasp the complexity of scientific experimentation; that is, the process is not step-by-step, but many steps that may need to be considered to some degree at the same time (cf. Chinn & Malhotra, Table 2, in press).

Advisor Actions. The actions taken by task and general purpose advisors largely consist of giving advice appropriate for a context. An advisor determines the context through information it receives from the software interface, such as the task the student is working on. The reasoning is accomplished by advice drivers, which are rule-driven, reasoning engines for establishing beliefs about which advice or message is appropriate to display from the advisor's knowledge base. For example, a student is working on her or his research question and has just begun. This context causes the Questioner to give the student the criteria for good research questions.

Modifiability. The advisors can be modified while the students work with them in several ways. Most simply, students can change the advice driver settings by clicking a button. The settings will modify how the advisor responds to the user actions. For example, changing the advice regulator to "Ask" means that the advisor will only give advice when the student asks for advice. Students can modify the knowledge base contents by editing, deleting, or adding an entry, or by changing which advice is given for a context. The Modifier (Figure 3) assists with this task by stepping the students through the process.

Environment

The students and the advisors operate within an inquiry support environment consisting of four graphical user interfaces (GUIs): the Project Journal, the Progress Report, the Meeting Room, and the Dialogue Box. Figure 4 provides a screen view of the GUIs.

The Project Journal is where students do and store their task-related work. Each of the tasks has a page that consists of a text box where the students enter anything they feel is relevant to the task. In future versions of SCI-WISE, the Project Journal will likely allow students to store images, sound, and video as well as text. The Project Journal sends any change in context such as current task and status to the Progress Report. Also,

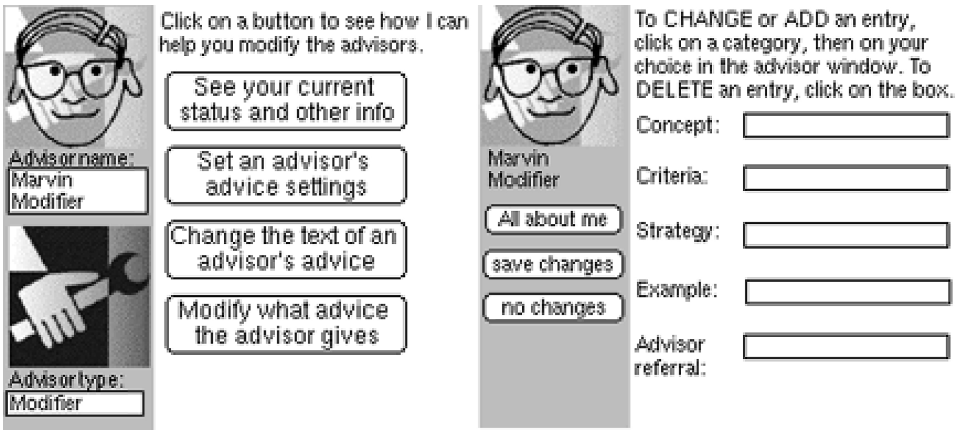

Figure 3. The Modifier (Marvin) interface.

any user actions in the Project Journal are sent to the Progress Report to be kept in the Log-file.

The Progress Report keeps track of student progress in the tasks and student actions. Three status classifications are "Haven't started" (automatically when no text has been entered on the task page in the Project Journal), "Working" (automatically when text has been entered), and "Finished" (only after students click a button when they are finished

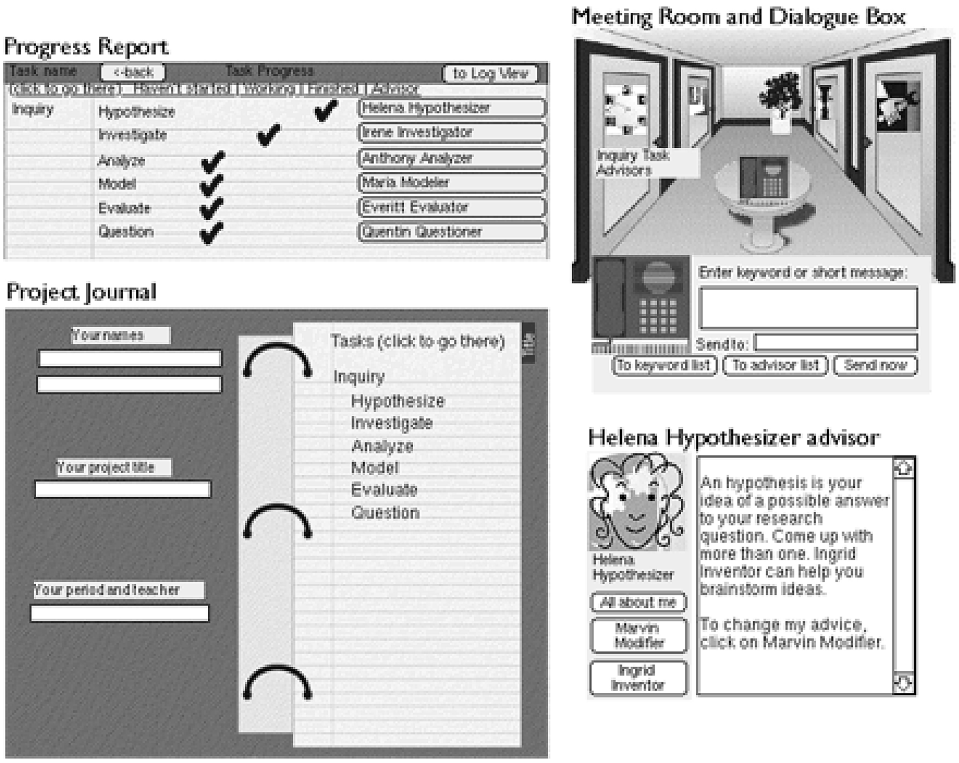

Figure 4. SCI-WISE graphical user interfaces.

with a task, this does not prevent them from working again on task). A check mark in the Progress report indicates the current status in each task. The Progress Report also provides access to the task advisor associated with a task and keeps a Log-file, or record of user actions.

The Meeting Room is where the advisors can be accessed by their classification, that is, Task, Cognitive, Social, and System Development advisors. Each category of advisor has a separate "room" of icons for the advisors in that category. To access any advisor, the students click on the advisor's icon.

The Dialogue Box in which the students can send messages to specific advisors also serves as a message builder. Any change in context is sent from the Project Journal to the message builder, which directs it to the current advisor. Message information also includes the sender (where the message originated from). Students can also send keyword messages to a specific advisor.

RESEARCH QUESTIONS AND METHODS

The levels of agency and goal orientation framework predicts that students doing inquiry projects would have different performance outcomes based on their goal orientations. In particular, we were interested in finding out if the modifiability of the advisors would affect outcomes differently. To investigate this, we developed two versions of SCI-WISE, one that was modifiable and one that was not. The modifiable version ("Modify") was broadly described in the previous section. The nonmodifiable version did not allow students access to the advice drivers and knowledge base editing features. This nonmodifiable version gave advice without student input as to what kind of advice the student might want or when they want it. The nonmodifiable version also prompted students to evaluate the advisors' advice (this version was called "Assess"). The assess activity was necessary because students would likely take more time using Modify, which introduces the potential confounding variable of time on task.

Our specific research questions were (a) How will students of differing goal orientations using different versions of SCI-WISE evaluate the system? (b) How will students of differing goal orientations learn target inquiry skills and concepts? and (c) Will students with task-completion goal orientations shift toward knowledge-building goal orientations?

Student Participants

Thirty two sixth-grade students from two urban middle schools in the East Bay area of northern California participated in the study. Thirteen of the students were female (41%) and 19 were male (59%). The ethnicities of the students were African American (21), white (7), Asian American (3), and Latino (1). The students represented a wide range of grade point averages (GPA, 4.0 maximum possible, obtained from school records), from a high of 3.86 to a low of 0.86 ($M = 2.36$, $\mathrm{SD} = 0.88$). The students were told that the study was after class and voluntary, and that they would be paid ($4/h plus a bonus for being on time.) They were told that the purpose of the study was to help us evaluate new software that helps middle-school science students do research projects and that the study would have no bearing on their grades.

Materials

The materials created for the first study were inquiry project tasks, two versions of SCI-WISE, a goal orientation pretest and posttest questionnaire, an inquiry pretest question, a GUI evaluation worksheet, a system evaluation, and an inquiry posttest question.

Inquiry Project and SCI-WISE. The students were given an inquiry project question to research through an inquiry cycle (White & Frederiksen, 1998). The question was "How can people best remember a list of words, for example, saying each word five times, or using each word in a sentence?" The inquiry steps in the cycle used in this study were Hypothesize, Investigate, Analyze, Model, Evaluate, and Question.

Each of the inquiry steps was presented on a separate page of the Project Journal. On each page, the students were given simple instructions; on the Hypothesis page, for example, the instructions were "Type your hypotheses about the best way to memorize a list of words. Type why you think so."

While the students worked, the advisors offered advice of various forms to help the students complete the project. The advice was not directly related to the inquiry project on memory; in other words, it could be generally applied to a host of inquiry projects. Both the Assess and Modify versions of SCI-WISE had the same knowledge base (see Table 1 for an example of the knowledge base for Quentin Questioner). When and which advice was displayed depended on the SCI-WISE version being used by the pairs.

In the Assess version, all of the advisors were set to display specific advice at anytime. The types of advice displayed for each of the three current status types ("Haven't started," "Working on," or "Finished") were generally according to the following patterns. For the status of "Haven't started," concept and criteria were given. For "Working on," strategies, examples, and advisor referrals were given. For "Finished," other strategies and advisor referrals were given. When an advisor's advice was displayed, AnnLi Assessor popped up (Figure 5) and students were prompted to "Please assess my advice."

In the Modify version, the first time an advisor popped up it displayed its Advice Regulation driver interface for the users to select when they want advice to be given. The users

TABLE 1
Partial Knowledge Base of Quentin Questioner

	Advice Type		
Entry/Type	Specific	General	Hints
Research question/concept	A research question is exactly what you want to find out with your project. It helps you remember what you are investigating.	A research question is something that you would like to know a possible answer.	What does "question" mean?
Good questions/ criteria	A research question is something no one knows the answer. You can make predictions about possible answers to a good research question. You can think of a way to investigate it in your project.	Research questions are interesting to you and other people and can be investigated in some way.	Can you explain what you are interested in investigating?

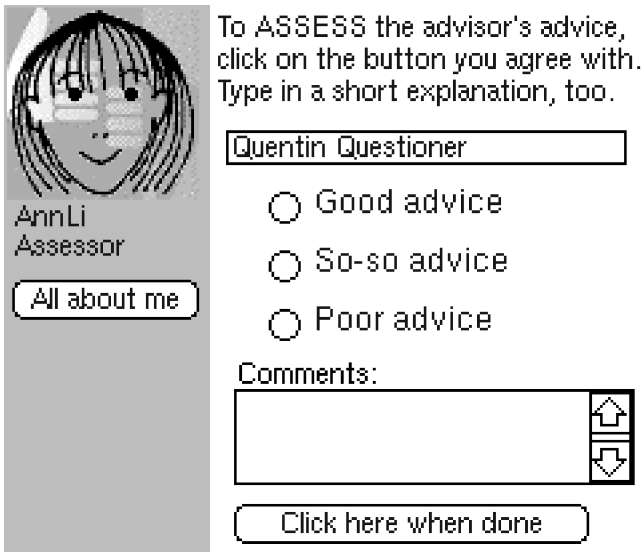

Figure 5. AnnLi Assessor, advisor assessment window.

were given three choices—Anytime, Alert, and Ask. After the students selected the advice regulation, they selected which advice type they wanted. The students were given three choices—Specific, General, and Hints. After the student selected the advice type, the advice was displayed. The initial types of advice for each of the three current status types were the same as for the Assessor. Of course, in the case of the Modify version, the students could change the advice displayed and when it was displayed. If a student did not select an option, no advice was given. Students could also access advice by clicking on the "All about me" button to access all the advisor's advice.

Goal Orientation Questionnaire. The range of questions in the goal orientation questionnaire (see the Appendix for sample questions) was limited to addressing the most likely consequences of having one goal orientation over another, for example:

(a) focusing on just getting a project finished instead of learning from it
(b) focusing on getting a good score rather than learning something
(c) putting effort into understanding the project when faced with difficulties
(d) figuring out what to do on one's own, rather than being told
(e) using the knowledge learned in new situations

Two question types were used, a biploar question and a dialogue question. The latter type has two fictional students offering different opinions; respondents select with whom they agree and to what degree (cf. Elby et al., 1997). The questionnaire was pretested with for sixth-grade students who were paid $5 for their participation.

Inquiry Pretest and Posttest. The inquiry pretest and posttest assessed students' abilities to come up with competing hypotheses for a research question and suggest an appropriate research investigation. The questions were open-ended and did not provide any further instructions on what type of answer might be appropriate. The pretest question, accompanied

by a schematic map of several neighborhoods of a fictitious city, was

> Keshawn, a middle school student, drew a map that showed all the violent acts reported to the police last year. He wants to know why there are more in the northeast neighborhood. What advice would you give him to answer his question?

A posttest inquiry question was similarly constructed with the following question:

> Shandice, a middle school student, drew a map that showed all the detentions reported to the principal last week. She wants to know why there are more in classroom 3. What advice would you give her to help her answer her question?

GUI Evaluation. The student pairs were given a worksheet to fill out together after they completed their inquiry projects. The worksheet asked them to write down the kinds of things they could do in each of the GUIs (Meeting Room, Progress Report, Dialogue Box), and what they could do with the system development advisor for the version of SCI-WISE they were using (Assessor or Modifier). The pairs were then asked to rate usefulness and ease of use of each of the GUIs and version advisor on a 1 to 3 scale (Very, So-So, Not).

System Evaluation. After the student pairs completed the GUI evaluation, they were given a posttest questionnaire that consisted of three evaluation questions (helpfulness in completing the projects, usefulness in science classes, and usefulness in other classes), and two open-ended questions (what they liked best and least about SCI-WISE).

Procedures

The study was divided into several steps. Several groups of pairs were run one at a time in a classroom with iMac computers.

Step 1: Pretest and SCI-WISE Introduction. Following a brief welcome by the researcher, the student participants were given the pencil-and-paper inquiry, beliefs, interests, and goal orientation questionnaire to complete on their own. After the questionnaire was completed, the students were given a treat. When all were finished, the students were asked to pair off. One pair at a time was then randomly selected. The pair was given a brief introduction to their version of the SCI-WISE system. The introduction only pointed out the main functions of the GUIs and the students were told that there were advisors that would help them complete the project. They were shown where to get started in the Project Journal, and then left on their own unless they had questions.

Step 2: Inquiry Project. The student pairs completed the inquiry project using the SCI-WISE system. If students had questions about completing their project, they were referred to the advisors for help. If the students had questions or problems with SCI-WISE, they were answered or shown how to fix the problem. If students needed experimental subjects for their investigations, they were told to ask other pairs for volunteers or the researcher could give them some data based on their experimental design.

Step 3: GUI Evaluation Worksheet. After the student pairs finished their project, they completed the GUI evaluation worksheet together.

Step 4: Posttest. After completing the worksheet, the students were given the individual SCI-WISE evaluation, and the inquiry posttest and goal orientation questionnaire (the same as in the pretest).

RESULTS

The results reported include time spent on the projects, scores on the pretest and posttest goal orientation and beliefs questionnaires, ratings of SCI-WISE, pretest and posttest inquiry scores, goal orientation shifts, and pairs' inquiry project scores.

Time Spent on Projects

All pairs were able to complete their projects using the software, taking on average 60 min. Students in the Assess group ($n = 14$) took an average of 62.2 min (SD = 6.9) compared with 58.1 min for those in the Modify group ($n = 18$, SD = 5.8); the difference was insignificant, $F < 1$.

Goal Orientation Questionnaire

An item analysis of the responses to the goal orientation questionnaire was performed (cf. Frederiksen et al., 1999). After deleting an outlier item and two items that measured interest in science, the coefficient alpha for the 17 remaining items in the pretest was 0.85. For the posttest, the coefficient alpha for the 17 items was 0.90. The possible range of scores was 17–85 for the 17 items.

The higher scores indicate a tendency toward knowledge-orientation, and lower scores toward task-orientation. The students' pretest goal orientation scores ranged from 40 to 79 ($M = 63.3$, SD = 10.8, Mdn = 65.0). Posttest goal orientation scores ranged from 40 to 83 ($M = 64.7$, SD = 12.0, Mdn = 64.0). Pretest and posttest scores had a Pearson correlation of 0.86, $p < 0.001$.

Goal orientation pretest and posttest scores had almost no correlation with GPA (pretest Pearson correlation: 0.036, $p = 0.85$; posttest Pearson correlation: −0.033, $p = 0.86$). To see if science interest was correlated with goal orientation, the responses to the two interest questions were combined for each participant to use in a correlational analysis. Goal orientation pretest and posttest scores had high correlation with science interest (pretest Pearson correlation: 0.64, $p < 0.001$; posttest Pearson correlation: 0.61, $p < 0.001$).

Statistical Analyses

The primary statistical procedure used in the analysis of the results was a general linear model (GLM) with SCI-WISE version as a categorical independent variable, goal orientation score as a continuous independent variable, interaction of version with goal orientation score, and GPA as a covariate. Separate analyses for each SCI-WISE version were also performed; in this case, the GLM was a multiple linear regression analysis consisting of a constant and two independent variables.

How Will Students of Differing Goal Orientations Evaluate the Different Versions of SCI-WISE System?

Two measures that evaluated SCI-WISE were collected: helpfulness in completing the study's inquiry project and usefulness in science class research projects. The evaluation scale was from strongly agree (5) to strongly disagree (1).

Helpfulness in Completing the Inquiry Project. The results showed that students generally found SCI-WISE helpful ($M = 4.23$, SD $= 1.1$), with 4 being "mostly agree" that SCI-WISE was helpful. Using the posttest goal orientation scores, there was an insignificant effect of version, adjusted least squares mean, $F < 1$. There was a significant effect of posttest goal orientation scores on helpfulness ratings, $F(1, 26) = 10.04$, $p = 0.004$. As Figure 6 shows, the helpfulness ratings generally went up with higher goal orientation scores. The interaction of version with goal orientation scores was marginally significant, $F(1, 26) = 4.00$, $p = 0.056$, indicating a trend for the version to affect the ratings differently in regard to goal orientation scores. The effect of GPA on the ratings was insignificant, $F < 1$. The model accounted for 45% of the variance.

Usefulness in Science Class Research Projects. The results showed that students generally thought SCI-WISE would be useful in science class projects ($M = 4.16$, SD $= 1.2$), with 4 being "mostly agree." Using the posttest goal orientation scores (Figure 7), the analysis found a significant effect of version on the usefulness in science class ratings, adjusted least square means: Assess $= 4.7$ (SE $= 0.27$), Modify $= 3.8$ (SE $= 0.23$), $F(1, 26) = 6.00$, $p = 0.021$, indicating that students using the Assess version generally thought SCI-WISE would be more useful in science class than did those using Modify. The posttest goal orientation scores effect was insignificant, $F(1, 26) = 1.61$, $p = 0.22$, and the interaction of version with goal orientation scores was marginally significant, $F(1, 26) = 2.67$, $p = 0.10$. The effect of GPA was insignificant, $F < 1$. The model accounted for 33% of the variance.

A multiple regression analysis for a single group, Modify, found a significant effect of posttest goal orientation scores, $t(15) = 2.50$, $p = 0.012$, indicating that students with higher scores rated SCI-WISE more useful in science class. There was a marginally significant effect of GPA on the ratings, $t(15) = 1.60$, $p = 0.065$, indicating that students with higher GPAs tended to rate SCI-WISE more useful in science class.

How Will Students of Differing Goal Orientations Learn Target Inquiry Skills and Concepts Using Different Versions of SCI-WISE?

The students' responses to the pretest and posttest open-ended inquiry questions were scored using a rubric with a range of 0–5. A second scorer scored 100% of the pretests and

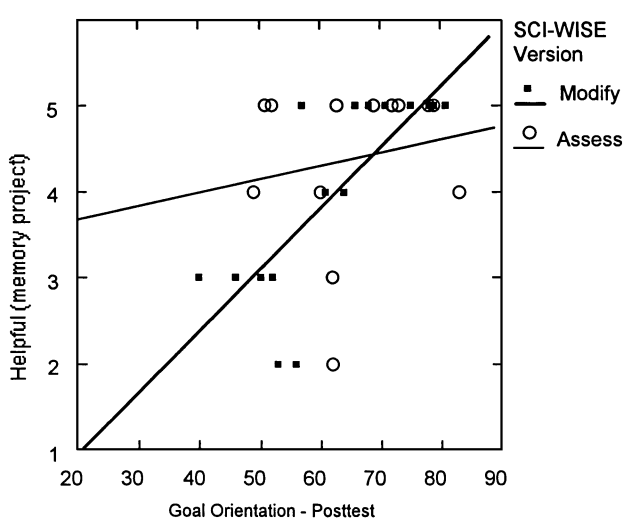

Figure 6. Students' ratings of the helpfulness of SCI-WISE in completing the inquiry task, plotted as a function of their posttest goal orientation scores.

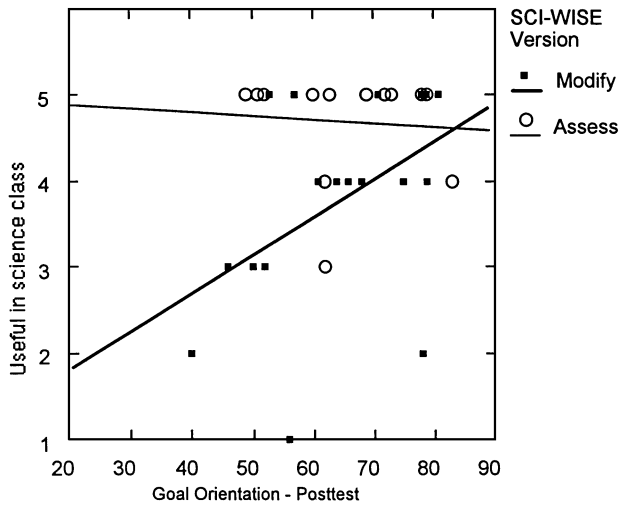

Figure 7. Students' ratings of the usefulness of SCI-WISE in science class, plotted as a function of their posttest goal orientation scores.

posttests, and the inter-rater reliability was 0.84. An example follows of a pretest response that scored 3 (the question was about higher detention rates in a classroom):

> Investigate the kids in the classes and their backgrounds. See how the teachers teach and how they deal with trouble-makers.

The above example gives two factors to investigate: (a) the kids might have something to do with the higher rate of detention and (b) the teachers might have something to do with it. The lower scoring answers tended to be single explanations or hypotheses, such as "The kids were bad," and did not mention any inquiry processes.

The overall results showed that students scored on average the same on the pretest open-ended inquiry question (scale 1–5, $M = 1.45$, $\mathrm{SD} = 0.72$) and on the posttest question ($M = 1.45$, $\mathrm{SD} = 0.85$). Table 2 shows the scores by version. A multiple linear regression analysis of the pretest inquiry scores was performed using goal orientation scores and GPA as independent variables. This regression was the same as the one used in the ratings analyses, except students in both version groups were included. Version was not included in the analysis because the pretest was taken before the students worked with SCI-WISE. In the regression analysis an insignificant effect of goal orientation scores was found, $t(27) = 0.69$, $p = 0.25$. There was a significant effect of GPA, $t(27) = 3.83$, $p < 0.001$, indicating that students with higher GPAs generally scored higher on the pretest inquiry question.

A GLM analysis was used to analyze posttest inquiry scores (Figure 8). The analysis found no significant effect of version, adjusted least square means: Assess = 1.53, SE = 0.16;

TABLE 2
Pretest and Posttest Open-Ended Inquiry Mean Scores by Version

SCI-WISE Version (n)	Pretest Mean, Scale 0–5 (SE)	Posttest Mean, Scale 0–5 (SE)
Assess (14)	1.38 (0.17)	1.53 (0.19)
Modify (18)	1.48 (0.14)	1.36 (0.16)

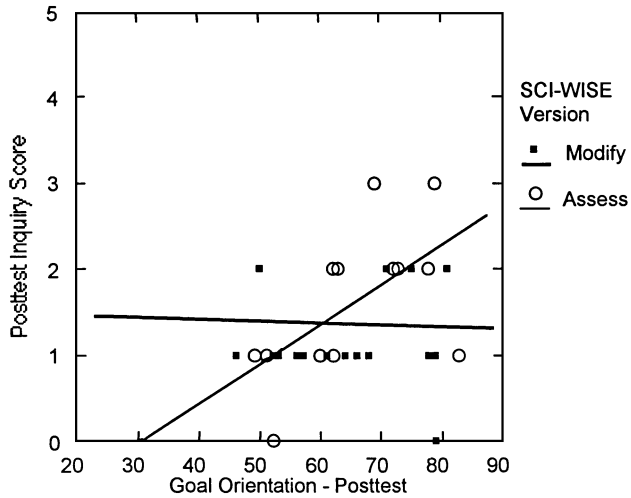

Figure 8. Students' posttest inquiry scores plotted as a function of their posttest goal orientation scores.

Modify = 1.36, SE = 0.14; $F(1, 25) = 2.00$, $p = 0.17$. The effect of posttest goal orientation score was significant, $F(1, 25) = 4.28$, $p = 0.049$, indicating posttest inquiry scores tended to increase with higher goal orientation scores. The interaction of version with goal orientation score was marginally significant, $F(1, 26) = 4.04$, $p = 0.055$, indicating that the effect of goal orientation scores tended to depend on version. The effect of GPA was significant, $F(1, 26) = 5.86$, $p = 0.023$, indicating that students with higher GPAs scored higher on the posttest inquiry question. The model accounted for 43% of the total variance.

An additional analysis was performed on the posttest inquiry scores using pair goal orientation type (Task-Task, Task-Know, and Know-Know). The grounds for this are that students' learning is influenced by how they use the software and the kinds of conversations they have with their partner while they do their project. White and Frederiksen (1998) found this to be the case: the lower-achieving students who were paired with higher students showed the greatest benefit of self-assessment in learning inquiry.

The pair goal orientation type was based on a median split of the students' goal orientation scores (Mdn = 64.0). The mean posttest scores for each group were as follows: Know = 75.3, SD = 5.1; Task = 54.9, SD = 6.9. The mean GPAs for each group were as follows: Know = 2.38, SD = 0.89; Task = 2.35, SD = 0.89. There were 4 Task-Task, 8 Task-Know, and 4 Know-Know pairs. The students' GPA pair types were also used in the analysis, based on a median split of the students' GPAs. The mean GPAs for each group were as follows: High = 3.06, SD = 0.57; Low = 1.62, SD = 0.39. There were 4 Low-Low, 8 High-Low, and 4 High-High pairs.

The analysis found a significant effect of version, adjusted least square means: Assess = 1.68, SE = 0.18; Modify = 1.17, SE = 0.14; $F(1, 23) = 5.42$, $p = 0.029$, indicating that students using the Assess version had higher posttest inquiry scores than students using Modify. The effect of goal orientation pair type was significant, $F(2, 23) = 6.59$, $p = 0.005$, indicating posttest inquiry scores were higher with Know-Know pair goal orientation types, adjusted least square means: Know-Know = 1.99, SE = 0.22; Know-Task = 1.52, SE = 0.14; Task-Task = 0.78, SE = 0.23. Using a planned comparison of Task-Task and Know-Know pairs, the difference in means was significant, $F(1, 9) = 11.38$, $p = 0.008$. The interaction of version with goal orientation pair type was also significant, $F(2, 23) = 6.46$, $p = 0.006$, indicating that the effect of goal orientation pair type tended

to depend on version (the pair types for students using Modify had little effect, while pair types for students using Assess had a large effect in the direction seen in Figure 8). The effect of GPA pair type was not significant, $F < 1$. The model accounted for 59% of the total variance.

Will Students with Lower Goal Orientations Shift Toward Higher Goal Orientations Using a Modifiable SCI-WISE Compared to a Nonmodifiable Version?

Overall, the mean posttest goal orientation score was insignificantly higher than the mean pretest score (pretest = 63.3, SD = 10.8; posttest = 64.7, SD = 12.0; $F < 1$). The shifts from pretest to posttest ranged from +13 to −13, with $M = 1.77$, SD = 5.89. The procedure used to analyze the shift was a repeated effects GLM using pretest goal orientation as a categorical variable (obtained from the median split of the scores). The categorical variable was used because the goal orientation scores were already included in the analysis as dependent variables. In the between-subjects analysis, there were no significant effects (all $F < 1$) except pretest goal orientation category, $F(1, 26) = 43.8$, $p < 0.001$, which merely indicates that pretest and posttest scores are highly correlated. In the within-subjects analysis (which analyzes each student's change in scores from pretest to posttest), the analysis found no significant effects (all $F < 1$ except for the difference from pretest to posttest, $F(1, 26) = 1.38$, $p = 0.25$).

Inquiry Project on Memory

The student projects were scored using the Inquiry Scorer (Frederiksen & White, 1998a, 1998b). The Inquiry Scorer provides a series of predetermined qualitative measures for each of White and Frederiksen's inquiry cycle steps (Frederiksen, 1998). Each of the measures is assigned a point value (e.g., a rating of "Clear" is given 1 point, "So-so" is given 0.5 point, and "Unclear" is given 0 points in response to the question: "Do the students clearly explain why someone might believe their hypotheses?"). The Inquiry Scorer also includes a holistic scoring rubric for scoring projects on several dimensions including Understanding Science, Understanding Inquiry, Reasoning Carefully, Being Systematic, Communication, and Teamwork. A previous analysis by Frederiksen of middle school teacher use of the Inquiry Scorer showed the software to be a reliable scoring instrument (Frederiksen & White, 1998a, 1998b).

An example of one of the higher scoring projects (holistic score of 3 out of 5) follows (sic):

Hypothesize:	We think the best way to memorize words is to use it in a sentence. We think this because if you have the word in a sentence it would be more easier. Our second hypotheses is that when you read the words over 5–10 times it will stick to your head.
Investigate:	We are going to test both of our hypotheses to see which ones will work. We are going to get a list of words and we are going to try to read five times and use them in sentence to try our hypothese.
Analyze:	Our results of our hypothese test is: Writing & saying 5–10 times-70% Saying in a sentence-90%. Our investigation proves our hypothese right memorizing your words in a sentence is much better than reading it 5–10 times. We did our analizing by our data and by comparing our data.

Model:	Our results were that memorizing your words a sentence will help you memorize your words better so our 1# hypothese is right. This happened because if you make up your own sentence it will better stick in your head because you created the sentence. So know it better.
Evaluate:	Our limitations are or things our modle can't explain are how the brain reactts to the memorizing or how much memorizing can the brain take thats what our model can't explain. Our theory can be used during the time of a test or inh daliy life.
New question:	How much time does it take to memorize a short story? We picked this question because it is pretty interesting.

In the present study, student pairs' point scores averaged 21.6 (SD = 7.25) out of 38 possible points; the scores ranged from a low of 2 to a high of 32. The holistic scores (from 1 to 5 possible) averaged 2.1, ranged from 1 to 3, and the point scores correlated with the holistic scores at 0.85 ($p < 0.001$). A GLM analysis was performed on the project point scores with the following variables: version, goal orientation pair-type (Know-Know, Task-Task, Know-Task), and GPA pairing based on median split (High-High, Low-Low, High-Low).

The analyses found a marginally significant effect of version, adjusted least square means: Assess = 23.5 (SE = 2.45); Modify = 17.5 (SE = 2.2), $F(1, 10) = 3.37$, $p = 0.096$, which indicated that students using the Assess version tended to have higher scores. There were insignificant effects of goal orientation pair type ($F < 1$) and the interaction of version and goal orientation ($F < 1$). The effect of GPA pair type was significant, $F(1, 9) = 5.76$, $p = 0.040$, which indicates that higher GPA pair types had higher scores. The model accounted for 58% of the variance.

DISCUSSION AND CONCLUSIONS

Goal Orientation Scores and GPA

The goal orientation questionnaire was found to be reliable; whether students who had the lowest scores can be validly classified as task-oriented and those with the highest as knowledge-oriented is not as easily determined. For instance, the students' goal orientations were self-reported, and their actual, real-time goal orientations could be different from what they report. Also, the questions used in the pretest and posttest must really address "goal orientation" to be considered valid.

The studies found that if a student has a knowledge-building goal orientation that does not necessarily mean he or she will have high academic achievement as measured by GPA. This result indicates that some knowledge-oriented students complete assigned tasks to a sufficient quality that allows them to get good grades, while others do not. This phenomenon implies that there may be more than just two kinds of students in regard to goal orientation. For example, categories of goal orientation and achievement might include those shown in Table 3, although the table does not mean to imply that the numbers in the categories would be equal, that is 25% in each cell.

Results from the study indicate that prior achievement and goal orientation had different correlations with student performance and ratings. For example, GPA was not correlated with SCI-WISE helpfulness and usefulness ratings while the effect of goal orientation scores was significant. The reverse was true for the pairs' project scores: GPA pair type was a significant factor in an analysis of project scores, while goal orientation pair type was not significant. GPA was a significant variable in the analyses of pretest and posttest inquiry

TABLE 3
Goal Orientation and Academic Achievement

	Goal Orientation	
Academic Achievement	Knowledge	Task
Higher	Knowledge-oriented/higher achieving	Task-oriented/higher achieving
Lower	Knowledge-oriented/lower achieving	Task-oriented/lower achieving

scores, while goal orientation only showed an effect in the case of the students who used the Assess version of SCI-WISE.

Ratings of SCI-WISE

The students' goal orientation scores were significantly correlated with their ratings of the SCI-WISE system. For those students using the Modify version, knowledge-oriented students tended to rate SCI-WISE higher than task-oriented students. For those using Assess, their ratings were generally flat for the range of goal orientation scores.

Task-oriented students using Modify may have found that the extra steps (setting the advisors' advice type and advice regulator drivers) required to get advice hindered task completion. For example, a task-oriented student using Modify wrote the following response in the system evaluation task: "You had to spend too much time setting up the advisor before you could actually get some advice." In contrast, a knowledge-oriented student using Modify liked "using the advice regulators." On the other hand, the Assess version gave advice immediately and did not have to interfere with task completion (students didn't have to assess the advice immediately in order to keep working on the task).

Pretest and Posttest Inquiry Scores

The inquiry scores for students using either version didn't significantly change from pretest to posttest for either task-oriented or knowledge-oriented students. The relatively short intervention of SCI-WISE, only an average of an hour, may account for this result. Overall, knowledge-oriented students (who worked with a knowledge-oriented partner) tended to score higher than task-oriented students (who worked with a task-oriented partner) on the posttest. Goal orientation pair-type actually accounted for more variance than GPA pair-type. Most of the effect came from students using the Assess version.

The posttest inquiry scores tended to be higher for students using the Assess version than students using Modify. Perhaps this occurred because students had a higher exposure to the advice that was given unsolicited with the Assess version. In other words, students in the Modify condition did not get advice unless they specified the type of advice and when the advisor was to display advice. This result fits with the study by Zellermayer et al. (1991) that showed students who received continuous, unsolicited metacognitive prompts performed better in writing tasks than students in solicited-guidance groups. In the Zellermayer et al. study, however, no differentiation was made according to student goal orientation. The results from the present study may have also been due to the metacognitive task of assessing advice. That is, by assessing the advice, it provides students with a prompted opportunity to determine how well the advice facilitates their understanding and use of the inquiry process.

Students in knowledge-oriented pairs did better on the posttest than other pair-types. Indeed, goal orientation pair-type was a greater predictor of posttest scores than GPA pair-type. The theoretical framework predicts this result—knowledge-orientation facilitates higher levels of competence.

Goal Orientation Shift

Overall, the goal orientation shifts were not significant. The theoretical framework predicts that knowledge-oriented students using the Modify version would increase their scores more than task-oriented students. This result did not occur, as students in the Modify condition on average had no gains. The individual shifts that did occur can be explained as regression to the mean. Possible explanations for the lack of an increase include (a) the goal orientation questions cannot pick up small shifts, (b) the Modify version did not provide enough of an increase in agency to affect change (or the students did not take enough advantage of the features), (c) the short time they used SCI-WISE, and (d) goal orientation as measured in the study is stable.

Inquiry Project Scores

Students with higher GPAs and those using the Assess version did better overall on their inquiry projects, particularly pairs with at least one student in the higher half of the median split. In contrast to the posttest inquiry scores, the effect of goal orientations of the student pairs was not significant. This might be explained by the fact that the inquiry projects were completed in pairs while the posttests were an individual effort. The inquiry project completion might have been dominated by students with a higher GPA. The results did show that there was little difference in project scores from pairs with two high-GPA students and pairs with one high-GPA and one low-GPA student.

Conclusions

The SCI-WISE system is based on the successes of the ThinkerTools inquiry cycle and curriculum (e.g., White & Frederiksen, 1998) and is intended to help students acquire inquiry skills through experimentation that is more epistemologically authentic (Chinn & Malhotra, in press). According to Chinn and Malhotra, there are many differences in cognitive processes between authentic experimentation and simple multivariate experimentation. For example, as in this study, the students were not given step-by-step instructions on how to investigate the research question as is often the case in simple experiments. The variables the students had to collect or control were not given to them, as is the case with authentic experiments. The students were not told what kind of theoretical model they were supposed to "discover" as a result of their research as often occurs with simple experiments. As part of their inquiry projects, the students were asked to develop their own model using the criteria of good theoretical models.

Of course, in this study, the students did not perform at the level of experienced scientists; they had much to learn. Also, they did not start from a position of being grounded in knowledge of research literature, as Chinn and Malhotra (in press) propose to be a part of authentic experimentation. The students did not necessarily grasp the concepts and theories behind the psychology of memory. Further research is needed to find out how to incorporate epistemologically authentic experimentation (embodied in educational technology or not) into science curricula.

APPENDIX: GOAL ORIENTATION QUESTIONNAIRE

BE HONEST! Circle only ONE choice (a, b, c, d, or e), please!

1. *When I do science projects I focus on getting them done and don't care much what I learn.*

   (a) I strongly agree (b) I mostly agree (c) I agree some and disagree some (d) I mostly disagree (e) I strongly disagree

2. *Tomas and Miguel are working on a science project together. Read what they both said:*

   Tomas: When I do science projects, my main goal is to understand the project.

   Miguel: When I do science projects, my main goal is to get a good score on the project.

   (a) I strongly agree with Tomas (b) I mostly agree with Tomas (c) I agree some with Tomas and some with Miguel (d) I mostly agree with Miguel (e) I strongly agree with Miguel

3. *I think working on science projects is boring.*

   (a) I strongly agree (b) I mostly agree (c) I agree some and disagree some (d) I mostly disagree (e) I strongly disagree

4. *Leticia and Miki are working on a science project together. Read what they both said:*

   Leticia: I think this project is hard, but I'm learning lots of stuff. Let's keeping working on it. What do you think?

   Miki: We're spending too much time figuring out the project. Let's just finish so we can start our other homework.

   (a) I strongly agree with Leticia (b) I mostly agree with Leticia (c) I agree some with Leticia and some with Miki (d) I mostly agree with Miki (e) I strongly agree with Miki

5. *When I work on science projects, if I don't understand something I stop and try to figure it out.*

   (a) I strongly agree (b) I mostly agree (c) I agree some and disagree some (d) I mostly disagree (e) I strongly disagree

6. *I just want to finish science projects as quickly as possible, even if I don't understand everything.*

   (a) I strongly agree (b) I mostly agree (c) I agree some and disagree some (d) I mostly disagree (e) I strongly disagree

The ThinkerTools research group at the University of California, Berkeley greatly contributed to the development of the software.

REFERENCES

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Student's learning strategies and motivation processes. Journal of Educational Psychology, 80, 260–267.

Cain, K. M., & Dweck, C. S. (1989). The development of children's conceptions of intelligence: A theoretical framework. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence, (Vol. 5, pp. 47–82). Mahwah, NJ: Erlbaum.

Chinn, C. A., & Malhotra, B. A. (in press). Toward epistemologically authentic reasoning in schools: A theoretical framework for evaluating reasoning tasks.

Craik, F. I. M. (1979). Levels of processing: Overview and closing comments. In L. S. Cermak & F. I. M. Craik (Eds.), Levels of processing in human memory (pp. 447–461). Hillsdale, NJ: Erlbaum.

Elby, A., Frederiksen, J., Schwarz, C., & White, B. (1997, March). EBAPS: Epistemological Beliefs Assessment for the Physical Sciences. Paper presented at the Annual Conference of the American Educational Research Association, Chicago.

Frederiksen, J., Elby, A., White, B., & Schwarz, C. (1999, April 14–19). Using item analyses for testing imputations about science students' epistemological beliefs. Paper presented at the Annual Conference of the American Educational Research Association, Montreal, Canada.

Frederiksen, J. R., & White, B. Y. (1998a). Teaching and learning generic modeling and reasoning skills. Journal of Interactive Learning Environments, 5(1), 33–52.

Frederiksen, J. R., & White, B. Y. (1998b, April). Assessing students' scientific inquiry. Paper presented at the Annual Conference of the American Educational Research Association, San Diego, CA.

Ng, E., & Bereiter, C. (1995). Three levels of goal orientation in learning. In A. Ram & D. B. Leake (Eds.), Goal-driven learning (pp. 355–380). Cambridge, MA: The MIT Press.

Piaget, J. (1978). Success and understanding (A. J. Powerans, Trans.). Cambridge, MA: Harvard University Press. (Original work published in 1974)

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, 37–68.

Shimoda, T. A., White, B. Y., & Frederiksen, J. R. (1999). Acquiring and transferring intellectual skills with modifiable software agents in a virtual inquiry support environment. In Proceedings of the 32nd Annual Hawai'i International Conference on System Sciences, January 5–8, 1999, Maui, Hawai'i. Los Alamitos, CA: IEEE Computer Society.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

White, B. Y., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. Cognition and Instruction, 16, 3–118.

White, B. Y., Shimoda, T. A., & Frederiksen, J. (1999). Enabling students to construct theories of collaborative inquiry and reflective learning: Computer support for metacognitive development. International Journal of Artificial Intelligence and Education, 10, 151–182.

Zellermayer, M., Salomon, G., Globerson, T., & Givon, H. (1991). Enhancing writing-related metacognitions through a computerized writing partner. American Educational Research Journal, 28(2), 373–391.