Developing the Changes in Attitude about the Relevance of Science (CARS) Questionnaire and Assessing Two High School Science Classes

Marcelle A. Siegel,[1] Michael A. Ranney[2]

[1]*Lawrence Hall of Science, #5200, University of California, Berkeley, California 94720-5200*

[2]*Graduate School of Education, University of California, Berkeley, California 94720-1670*

Received 28 January 2002; Accepted 12 November 2002

Abstract: This study has two purposes: (a) methodological—to design and test a new instrument able to reflect changes in attitudes toward science over time, and (b) investigative—to find out the effect of two similar curricular treatments on the attitudes of two classes. Items about the relevance of science to students' lives were developed, pilot-tested, and analyzed using Rasch modeling. We then divided reliable items into three equivalent questionnaire forms. The final three forms of the questionnaire were used to assess high school students' attitudes. Over 18 weeks, one class used a core curriculum (Science and Sustainability) to learn science in the context of making decisions about societal issues. A second class used the same core curriculum, but with parts replaced by computer-based activities (Convince Me) designed to enhance the coherence of students' arguments. Using traditional and Rasch modeling techniques, we assessed the degrees to which such instructional activities promoted students' beliefs that science is relevant to them. Both classes tended to agree more, over time, that science is relevant to their lives, and the increases were statistically equivalent between classes. This study suggests that, by using innovative, issue-based activities, it is possible to enhance students' attitudes about the relevance of science. © 2003 Wiley Periodicals, Inc. J Res Sci Teach 40: 757–775, 2003

"Rose-colored glasses." "Half full or half empty?" Such sayings remind us of the effect that one's attitude can have on one's experience (Ranney, 1996). In the educational and psychological research literature, many ways of categorizing attitudes have been developed, from disposition to opinion, and from affect to belief. Generally, the research on dispositions tends to examine human sensibilities and approaches; the work on opinions tends to include individuals' current attitudes or views of a particular issue; research on affect deals with emotional expression and feelings; and the belief literature concerns deeply held doctrines, such as epistemological ideas about the nature of knowledge.

Contract grant sponsor: NSF Research Training Grant.
Correspondence to: M.A. Siegel; E-mail: mcgull@uclink.berkeley.edu
DOI 10.1002/tea.10110
Published online in Wiley InterScience (www.interscience.wiley.com).

Attitudes toward Science

Work in the realm of students' attitudes toward science has been motivated by the desire to increase interest, performance, and student retention in science (Third International Mathematics and Science Study, 2001). Educational studies have produced mixed results but tend to show that attitudes affect students' persistence and performance (for a review, see Schommer, 1994). Modest positive correlations between science attitude and science achievement have been reported in many studies (Schibeci & Riley, 1986; Keeves & Morganstern, 1992). Models propose that science instruction influences attitudes in ways that predict achievement (Hegarty-Hazel, 1990; Schibeci & Riley, 1986; Simpson & Oliver, 1990). In particular, science instruction that is activity-based (Dickinson, 1976; Fraser, 1980; Freedman, 1997) and issue-oriented (Iskandar, 1991; McComas, 1993) has been shown to enhance positive attitudes toward science.

Another reason to extend research on scientific attitudes is that attitudes and beliefs are part of "cognition" itself (Schoenfeld, 1985). Research has suggested that people who view science in a sophisticated way are better able than others to use their knowledge in more contexts and to make sense out of complex information (Davis, 1998; King & Kitchener, 1994; Linn & Songer, 1993). Thus, science attitude research may be viewed as fundamental to understanding scientific cognition.

Measuring Attitudes toward Science

To measure attitudes toward science, researchers have used Likert scales, Thurstone scales, semantic differential scales, direct interviews, and, much less frequently, indirect or projective methods and interviewing techniques (Ramsden, 1998). The most common method, Likert scales, has several advantages: They are not difficult to create, can include a large number of items that can be answered quickly, can provide precise information about a respondent's degree of agreement or disagreement, and can provide high reliability (Oppenheim, 1992). Disadvantages of Likert scales can sometimes be overcome using special techniques or multiple methods. For example, one criticism of measuring attitudes with Likert scales is that the neutral midpoint might not represent neutrality in the subject's mind, but rather a state of confusion or misunderstanding. Researchers can mitigate this problem by adding another response choice, such as "do not understand." Another criticism of Likert scales is that they might decontextualize the issue studied. For example, students' diverse views of science might hinge on what is meant by *science*, and a scale might not be able to account for these different meanings (Bell & Linn, 2002). This type of problem is lessened by crafting Likert items to address the specific subtleties of issues (Oppenheim, 1992) and by using multiple methods. Interviewing and using new open-ended instruments such as the Views of Nature of Science questionnaire (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002) offers practical solutions that can add qualitative depth to a study—although they also increase the classroom time used. Whereas attitude scales entail problematic assumptions owing to the ways participants might interpret a given item (Lederman et al., 2002), open-ended instruments entail similarly problematic assumptions regarding how to interpret participants' statements. Projective or indirect techniques allow the researcher to uncover ideas without participants' direct awareness; however, such techniques have disadvantages as well, such as lower reliabilities and more interpretive difficulties (Oppenheim, 1992).

To our knowledge, reliable questionnaires to measure changes over time in scientific attitudes have not been developed. Previous research on attitudes toward science has concentrated on students' beliefs about the nature of science (Lederman et al., 2002; Moore & Foy, 1997; Moore &

Sutman, 1970; Ryan & Aikenhead, 1992), students' values (Koballa, 1988), and students' images of scientists (Driver, Leach, Millar, & Scott, 1996; Mead & Metraux, 1957; Song & Kim, 1999). Others have studied affective orientation toward science, especially as it relates to career choice, the interest and intention to engage in science activities, enjoyment of classes, study practices, and perceived ability (Andre, Whigham, Hendrickson, & Chambers, 1999; Butler, 1999; Conwell & Prichard, 1992; Crawley & Koballa, 1994; Marsh & Yeung, 1997; Shamai, 1996; Stake & Mares, 2001; Tamir & Amir, 1987). Several researchers have also demonstrated gender differences regarding these factors (Kahle & Lakes, 1983; Mattern & Schau, 2002; Neathery, 1997; Weinburgh, 2000). Some studies have focused on whether instructional interventions can alter students' views of the nature of inquiry in science (Bell & Linn, 2000; Davis, 1998).

Research Issues: New Questionnaire Development, Attitude Change Assessment

This study has two main purposes: (a) methodological—to design and test a new science attitude instrument able to reflect attitude change over time and (b) investigative—to find out the effect of two similar curricular treatments on the attitudes of two classes. In this section, we describe the need for and the objectives of the study.

Methodologically, evaluating students repeatedly with the same questionnaire is not adequate. Students taking a test more than once could remember past responses or learn from the questionnaire itself (i.e., the retest effect). New concerns arise if the tests are not exactly the same, as they should have equivalent difficulties. For testing attitudes, questions should also be equally balanced in terms of being hard to agree with and easy to agree with (we will call this item *extremeness* instead of *difficulty*). If the questionnaires diverge, they also must be consistent so that they measure the same variables (we call this *issue consistency*). With these criteria in mind, three similar versions of the attitude questionnaire were developed through pilot-testing and quantitative modeling using Rasch techniques.

In addition, we used the questionnaire with two innovative science classes to determine whether (a) collectively, classes' attitudes toward science would change over time, and (b) there would be a difference between groups. Based on previous work cited above, one would expect an attitude change if a curriculum were sufficiently activity-based and engaging; yet it is rare to find dramatic changes in beliefs or attitudes in a short time. For example, epistemological beliefs—entrenched ideas about the nature, origin, justification, and limits of knowledge, often characterized as implicit or below conscious awareness and expression—are usually further characterized as developmental in nature (King & Kitchener, 2002; Kuhn, 1991) and not easily altered with the presentation of new information [summarized in Aronson (1995)]. In addition, research has shown a decline in positive attitudes toward science as students reach the high school level [summarized in McComas (1996) and Piburn and Baker (1993)], suggesting that one might even expect a decline with the two 10th-grade classes engaged in the present study. The design and history of the curriculum, though, led us to expect attitude enhancement to be possible albeit difficult to achieve.

In this study, we used two activity-based science curricula. The first materials were developed by the Science Education for Public Understanding Program (SEPUP) to promote scientific literacy and decision making. The second set of instructional materials was created by the first author to be used with the Convince Me (CM) software (developed by Schank, Ranney, & Hoadley, 1996) to enhance the coherence of students' arguments. The SEPUP-only (S) group was compared with the SEPUP/CM (CM) group. When both groups used SEPUP curricular materials (described in the next section), they learned about scientific evidence that they then used to make decisions involving social consequences.

Various possibilities could explain why students' attitudes might change. Both groups might be expected to have a more positive view of science at the end of the study because they both used more of the realistic-issues curriculum developed by SEPUP. By exclusively using SEPUP's hands-on science activities that relate to the real world, perhaps the SEPUP-only, S, group would gain a more positive view of the applicability of science than would the mixed group. The CM group used Convince Me software (described in the next section) during part of the semester. The CM group might be expected to develop positive views toward science because of their additional experience with a scientific method of reasoning. CM is structured to help people reason coherently about two or more sides of an issue, and these features might also generate more positive science attitudes. Furthermore, students used CM to make choices involving social and personal topics. Perhaps by using CM more intimately and possibly by learning to be a better decision maker, students would strengthen their attitudes that science is relevant to them.

These are many reasons why one might expect our participants' science attitudes to shift, but finding out exactly why is beyond the scope of this study. Our purpose was to develop an instrument and use it to measure changes over 18 weeks in the attitudes of the two classes of students using the innovative science curricula. The S and CM groups responded to the science attitude questionnaire three times: before, immediately after, and 3 weeks after CM was taught. Would students perceive science as being more relevant to their lives over the 18-week semester? Note that the prediction for the attitude study does not include a comparison of the two groups. In the overall study, however, the CM group was expected to have higher gains on other constructs regarding reasoning and decision making than the S group) (Siegel, 1999a,b).

The Common Part of Both Curricula: SEPUP

The central S curriculum used to different degrees by both groups in this study is issue-oriented in that students learn science in the context of making decisions about realistic societal issues. For example, providing food for the world is the motivating issue for students to learn about genetics, energy transfer, and plant biology; selective breeding and genetically engineered crops are explored as tools that agriculturists use to increase crop yield (SEPUP, 1999a). The goals of issue-oriented programs are to (a) develop an understanding of science and problem-solving processes related to social issues without taking an advocacy position, and (b) let students make their own choices based on what they learn (Thier & Nagle, 1994). The instructional materials are intended to promote scientific literacy and enhance the role of students as independent thinkers and active participants in science and society (SEPUP, 1999a).

Research efforts have shown that, in contrast to control materials, SEPUP materials increase student achievement along certain variables measured (Roberts & Wilson, 1998; Wilson, Sloane, Roberts, & Henke, 1995). SEPUP middle school students made statistically and educationally significant gains in their ability to use evidence and weigh tradeoffs (Wilson, Sloane, Roberts, & Henke, 1995), and high school students also improved in this regard (Siegel, 1999a,b). SEPUP instructional materials, assessment rubrics, and moderation activities are also powerful professional development tools; they improve teachers' abilities to assess learning (Roberts, Sloane, & Wilson, 1996; Roberts & Wilson, 1998) and they improve teaching practices, such as clarifying learning goals and establishing fair standards (Roberts & Wilson, 1998).

Participants in this study used a new SEPUP course, Science and Sustainability, to learn about science concepts, especially in biology, as they relate to sustainability.[1] The course package included written materials for students and teachers, as well as laboratory equipment. The SEPUP approach to science teaching involved investigations, discussions, debates, and other activities

(SEPUP, 1999a). The course was designed to encourage students to (a) explore the scientific principles and tradeoffs involved in obtaining their desired quality of life, and (b) develop a global perspective by investigating issues from the point of view of people living in different regions of the world (SEPUP, 1999a). Classes in this study used the field test version of this course that is now published (SEPUP, 1999b).

CM Activities

The CM software (Figure 1) and curricular system have been used in many domains from physics to abortion issues, and primarily for evaluating hypotheses and evidence—although this task readily converts to decision making (Ranney & Schank, 1998; Thagard & Millgram, 1995). We used the CM program with one of two classes of students to see whether it engendered more, less, or no attitude change compared with the S curriculum. The primary instructional goal in this study was to provide more guidance during the decision-making process.

While using the program, a student enters hypotheses and evidence about a controversy. The student connects each statement to at least one other statement with (a) explanation links if the statements support each other or (b) contradiction links if the statements conflict with each other. Also, the student rates each proposition according to how much she believes it and, if it is evidence, how reliable it is. The steps are summarized in Table 1.

Convince Me can think using its connectionist network (ECHO) which simulates aspects of human reasoning (Ranney & Thagard, 1988). Using ECHO, Convince Me offers feedback (model's fit scores) on the degree to which a student's evaluation of each proposition matches the value predicted by the computer. ECHO's rules for prediction are based on the Theory of Explanatory Coherence articulated by the philosopher of science, Paul Thagard (1989). This theory is based on foundational ideas about sound arguments—for example, that the more support

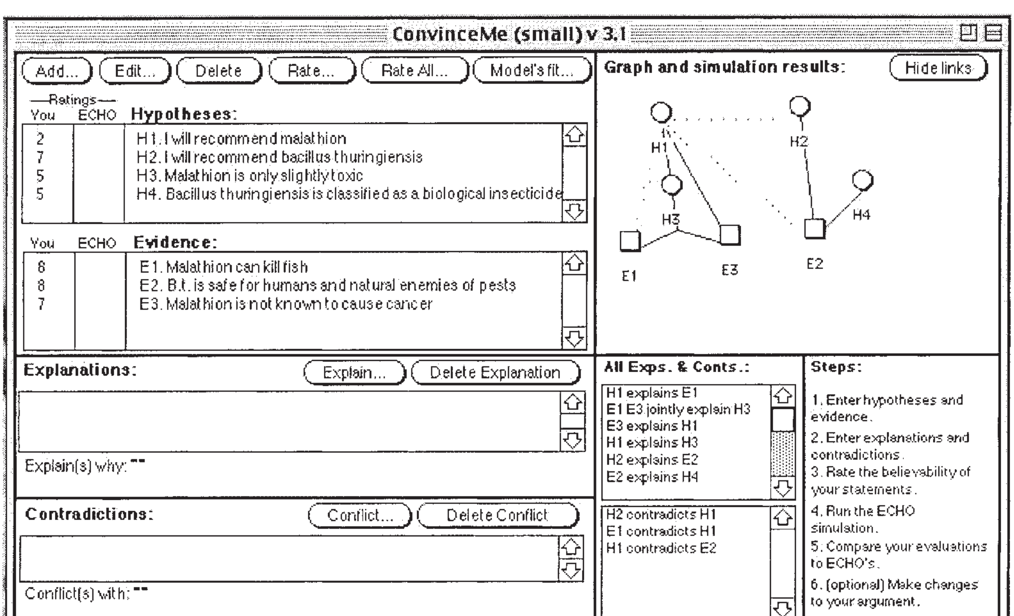

Figure 1. Convince Me main screen.

Table 1
Steps for using Convince Me in this study

Step No.	Action
1.	Articulate beliefs about an issue presented.
2.	Categorize each belief as ‘‘evidence’’ or ‘‘hypothesis.’’ Check off descriptive statements about the belief (e.g., ‘‘Some reasonable people might disagree.’’) If evidence, rate its reliability.
3.	Connect propositions with ‘‘explanation’’ or ‘‘contradiction’’ links.
4.	Rate the plausibility of the belief on a 1–9 scale.
5.	Run the ECHO simulation.
6.	Obtain feedback (‘‘model’s fit’’), revise argument, and cycle through the steps again.

Note. The particular sequence shown here is not mandatory.

a statement has from accepted sources, the more believable it is. The theory and ECHO were tested empirically and were repeatedly found to be effective predictors of subjects' beliefs in various domains (Read & Marcus-Newhall, 1993; Schank & Ranney, 1991, 1992).

Prior studies with Convince Me indicated that, like SEPUP, it is a useful tool for learning about reasoning (Diehl, Ranney, & Schank, 2001; Ranney, Schank, & Diehl, 1995). Students using the program performed better than students doing similar pen and paper exercises (Schank, 1995). Also, undergraduates and adolescents working with Convince Me improved at distinguishing between hypotheses and evidence (Ranney, Schank, Hoadley, & Neff, 1994) in ways particularly facilitated by Convince Me (Diehl, Castro, & Ranney, 1997; Schank, 1995; Siegel, 1999a). For example, using Convince Me, they tended to most associate a piece of evidence with an ‘‘acknowledged fact or statistic,’’ whereas a hypothesis correlated most with ‘‘one possible inference, opinion, or view’’ (Schank, 1995). High school students using Convince Me were able to support their claims with objective evidence and generated more than one alternative while making complex decisions (Siegel, 1999a,b). Furthermore, Convince Me was used to study the process of reasoning; for example, high school students' and experts' reasoning about global warming was compared (Adams, 1998).

Method for the Two Groups

This study was conducted within a larger study that also examined students' reasoning, including how students supported decisions with scientific evidence, weighed tradeoffs, and understood how Convince Me evaluates arguments and propositions (Siegel, 1999b).

Participants

Two 10th-grade classes participated in the study, with 28 pupils in Group S and 19 in Group CM. The two classes' scores on the standardized Terra Nova test (CTB, 1997) were not significantly different ($p = .35$) before instruction. The overall mean score was 2.45 on a 5-point scale, which was below the test's nearing proficient score of 3.00.

The participating school faces socioeconomic challenges typical of the inner city: 50% of its students qualify for Aid for Families with Dependent Children (AFDC); 42% of students are identified as Limited English Proficient by the school district. SAT scores reported for the 41% of the school's seniors who took the test were dramatically below the national average: 321 on the verbal portion and 437 on mathematics, compared with the national averages of 423 and 479, respectively.

Design, Procedure, and Materials

Figure 2A shows the timing of activities, evaluations, and topics covered. For the first 7 weeks, both classes participated in the same Science and Sustainability activities at the same time. Then, for 8 weeks (indicated by the middle double-headed line), the CM class used both the SEPUP and the CM activities. Because of a scarcity of the school's computer facilities, half the class used CM one day while the other half completed SEPUP activities, and the next day they switched. The time that the CM class spent on decision-making activities with the software was short (effectively 4 weeks, or 22% of the time), compared with the time spent on SEPUP and other materials. The S class continued to use SEPUP issue-oriented activities during the time the CM class used the software in this manner. During the last 3 weeks of the study, both classes engaged in activities and review (Figure 2A). Thus, the study proceeded for half of a school year, and the majority of the time, both classes were using SEPUP materials. Figure 2A also shows the times when attitude questionnaires were administered. Students spent less than one period on each questionnaire. The first author interacted with both classes daily and equally based on the procedure just described. From September through January, before the study began, both classes were using the Science and Sustainability course.

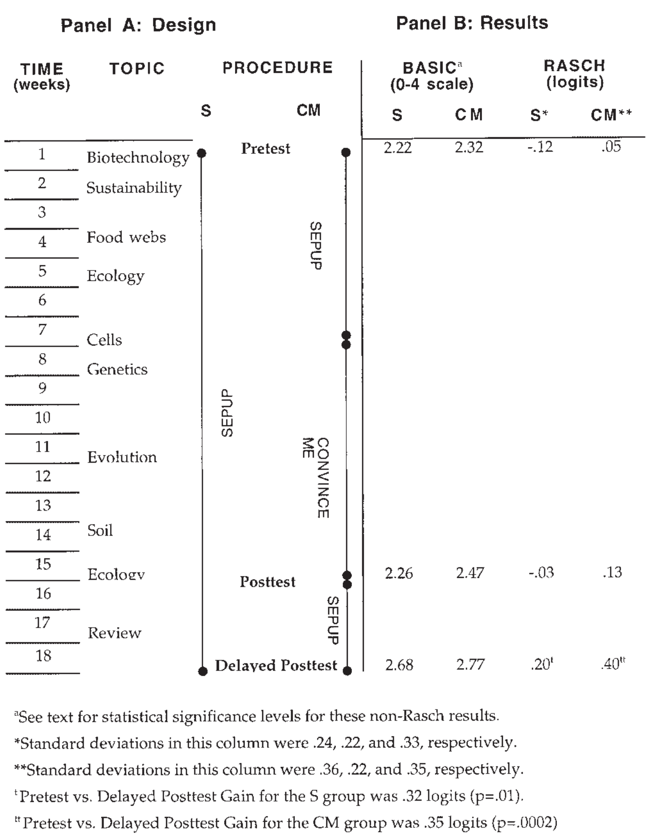

Figure 2. (A) Timing of instruction and assessment. Labels on double-headed lines indicate weeks when groups were working on SEPUP or Convince Me activities. (B) Science attitude results on the pretest, posttest, and delayed posttest which, as indicated in (A), occurred during the 1st, 15th, and 18th weeks.

The regular teacher taught both classes, which were at the Advanced level of First-Year Biology. The teacher was experienced in using SEPUP materials, and she taught all the other sections of Advanced Biology for the school, as well as an Environmental Science Institute. During the CM computer activities, the group was split into alternating halves. The first author acted as the CM instructor in the computer lab; the regular teacher used SEPUP activities in the classroom. A computer teacher, who was primarily an English teacher, was also present in the computer center.

The SEPUP science topics covered during the study were sustainability, biotechnology, food webs, genetics, and ecology. Figure 2A shows the sequence. Students in the CM group learned and practiced a model of decision making they could use in other situations involving complex choices. Once these students were adept at using the program [see Siegel (1999a) for details of instruction], each made decisions while working alone with CM. Students entered their alternatives, beliefs, and/or goals into the program and supported these with evidence and other hypotheses. They also had an opportunity to use CM to make a personal decision of their own choice. The goal was for students to learn ways to make an evidence-based, coherent decision.

Development of Three Versions of the Changes in Attitudes about the Relevance of Science (CARS) Questionnaire

We developed three versions of a questionnaire concerning students' attitudes toward science to assess students' views at different times. The CARS questionnaire, shown in Appendix A, consisted of 59 items that were eventually separated into three equivalent versions (pretest, posttest, and delayed posttest) by pilot-testing the questions with 188 9th- and 10th-grade students (none of whom were in the S and CM groups). The final three versions of the questionnaire contained 8 repeated and 17 unique items, for a total of 25 items on each version. The items include and expand upon SEPUP's Perceived Relevance of Science items [with permission from Roberts and Henke (1997)].

The purpose of the questions was to investigate students' views on the nature of science—specifically the relevance of science to their lives as they make decisions. For example, do students believe that science may be helpful in making better decisions about their health or about the environment? Does using a scientific process help the student make good decisions? For this study, science is broadly defined and includes students' commonsense notions of science. Accordingly, the questionnaire was designed to conceptualize science in multiple ways. The terminology included *scientific*, *science class*, *science experiments*, *learning science*, and *scientific methods*. Topics in the questionnaire included several arenas in which science might be relevant or irrelevant to students, including health, the environment, sports, computers, college, thinking skills, social skills, future success, and course selection.

Reliability and Experimental Concerns

As alluded to earlier, a potential problem with evaluating students over time on the same variable arises if they take the same test multiple times. This retest effect was diminished because some of the items were different on each test (Appendix A).

Other concerns involve item extremeness, issue consistency, item fit, and internal consistency. Item extremeness, issue consistency, and item fit were computed using Quest software (Adams & Khoo, 1993; a newer version is now available) in the aforementioned pilot study with 188 other students. Again, item extremeness is the relative ease of agreeing or disagreeing with an item. It is important for items to have the same level of extremeness on equivalent test forms. Therefore, we assessed item extremeness on each questionnaire to ensure that any observed increase or decrease in agreement was not because it was easier or harder to agree with the items themselves. Also, issue consistency was assessed to ensure that items consistently represented the variable being measured. Item fit (a measure of both item consistency and quality, computed by Quest software) and item consistency (Cronbach alpha) were also calculated and addressed in the pilot studies. After these four factors were examined in such studies, items were then adjusted—removed or switched to a different test form—and recalculated so that the three questionnaires used in this study would be valid and reliable. Statistics illustrating the validity and reliability of attitude questionnaires are reported in the next section.

Experimental Concerns Addressed by Rasch Techniques

Data from the pilot and final tests were analyzed using advanced measurement techniques from Rasch modeling (Masters, 1982). In Rasch analyses, the raw information is transformed using techniques from item response theory. We used a slightly generalized form of Master's partial credit model (Masters, 1982). With Rasch analyses, objectivity is obtained by estimating person ability and item difficulty (attitude and extremeness in the present study) with a probability function for a particular answer such that the probability of getting an item correct is a function of the difference between person ability and item difficulty (Rasch, 1960).

Rasch models offer a few benefits compared with classical measurement techniques. One advantage of using Rasch models is that participant characteristics and test characteristics can be separated rather than interpreted only in the context of each other. Using classical measurement models, a participant's attitude is defined only in terms of a test such that if the test is extreme, she will appear to have a moderate attitude. In addition, classical measurement procedures have not provided complete solutions to important testing issues, such as designing tests (Lord, 1980), exposing biased items (Lord, 1980), and equating test scores (Cook & Eignor, 1991). Using Rasch techniques, one can also compare the extremeness of particular items in detail, rather than using items without knowing how they differ. Classical theory does not account for items, but rather for a composite test score, thus only providing for comparison of people (norm referencing). Rasch modeling is both norm-referenced and criterion-referenced, which offers more avenues for interpreting the results. Criterion referencing means that performance is measured according to specified standards or criteria, rather than compared with the performances of other test takers. The model assumes a progression such that if a student scores well on a hard item (or one that is difficult to agree with), he will also score well on an easier item (or one with which he is more comfortable to agree). When items fit the model, they adhere to this rule. Because we are investigating changes in attitudes about science, a model that expresses change in a criterion-referenced way is preferable.

As noted above, the CARS questionnaire consisted of 59 items separated into three partially redundant final scales (pretest, posttest, and delayed posttest) with 25 items each. The questionnaire was a reasonably reliable instrument, as indicated by the internal consistency of each scale being above .80, and for all 59 items it was .91. The first part of Appendix B shows fit statistics for each item. The items did not completely represent attitudes toward science consistently, as shown by the item fit statistics in Appendix B.

Qualitatively, the Rasch model's participant measures for this scale represent individual differences in attitudes toward science—that is, from strongly believing that science is relevant to everyday life to strongly believing that science is irrelevant. Appendix C summarizes opinions characteristic of four levels of opinion out of the 8 categorized. Rasch units are called logits, and Appendix C provides a qualitative sense of what a logit score means in terms of how strong a representative student's opinion was.

Results for the Two Groups

Basic (Nonscaled) Results

Basic non–Rasch-scaled results are included here for readers not familiar with Rasch modeling. Findings from the three tests showed increasingly more positive views toward science. On a 0–4 continuum, from "strongly disagree" to "strongly agree" regarding the relevance of science, both groups averaged between 2 and 3 (neutral and agree) at every time point. The average scores for the S and CM groups are shown in Figure 2B. Both groups tended to agree more over time that science is relevant to their lives. The 3×2 (Time $\times$ Participant group) repeated-measures analysis of variance (ANOVA) showed no significant difference between groups and revealed a significant difference over time, $F(1, 18) = 17.98$, $p < .000$. Simple pairwise comparisons between time periods for the full sample revealed that differences between each test were significant: pretest and posttest, $F(1, 37) = 4.94$, $p < .05$; posttest and delayed posttest, $F(1, 18) = 16.92$, $p < .005$; and pretest and delayed posttest, $F(1, 25) = 23.71$, $p < .001$. The overall gain was equivalent for both classes, increasing by about half a point (.45 and .46).

Rasch Analyses

Rasch analyses should provide a more reliable temporal picture of students' attitude changes toward the relevance of science, by modeling the measurement error. Rasch analyses are also better able to handle missing values, resulting in different sample sizes for the Rasch results than the basic results. Figure 2B shows average Rasch ability estimates (in logits) and reveals more positive attitudes toward science over time. As noted, standard deviations ranged from .22 to .36. Figure 2 also notes the average gain for each class. On average, student attitudes increased more at the end of the semester, between the posttest and delayed posttest, but not significantly. The 3×2 (Time $\times$ Participant group) repeated-measures ANOVA of ability estimates for each student again showed no significant difference between groups and again revealed a significant difference over time, $F(1, 20) = 23.50$, $p < .001$. Simple pairwise comparisons between time periods for the full sample revealed that differences between each test were significant: pretest and posttest, $F(1, 39) = 4.23$, $p < .05$; posttest and delayed posttest, $F(1, 20) = 29.12$, $p < .001$; and pretest and delayed posttest, $F(1, 27) = 33.73$, $p < .001$. The Rasch results show the exact same pattern of results as the basic analyses with some slightly greater F values.[2]

Students' opinions about science ranged from a score of 3.13 logits down to −0.81 logits. The logit scores can be interpreted qualitatively by comparing them with similar entries in Appendix C. For example, the S class felt that science was not very relevant to everyday life, during the pretest (−0.12 logits on average) (Figure 2). On the posttest, their opinion was more neutral (−0.03 logits), and on the delayed posttest (0.20 logits) they had a mild inclination toward the relevance of science. The CM class started from a neutral position on the pretest (0.05 logits), became more positive on the posttest (0.13 logits) and felt that science was somewhat relevant to them on the delayed posttest (0.40 logits).

Rasch analysis also revealed which items were more or less extreme. For example, the item regarding whether learning science would affect one's election vote was easy to agree with. Also, more students agreed with this item on the posttest and delayed posttest than on the pretest. Students from both classes tended to be neutral about science helping them evaluate their own work. Furthermore, on the delayed posttest, students strongly agreed with the item, "Science helps me to make sensible decisions." Although students became more positive over time on most items, a few items were difficult to agree with, such as one about science class being useful in one's everyday life.

Discussion

Much has been accomplished toward developing a reliable and valid instrument to measure students' attitudes toward science (Moore & Foy, 1997; Roberts & Henke, 1997). However, we are not aware of any instruments designed to measure attitude about the relevance of science over multiple intervals. Such a measure requires careful selection of items and incorporation into versions of questionnaires that are equivalent yet different enough so that not all of the items are repeated. In this study, we developed three versions of a science attitude questionnaire, using a Rasch model, and applied these to compare students' attitudes over time on two variants of an innovative biology course.

Our results regarding students' attitudes about science are encouraging in that our analyses show significant improvements over time. The basic results demonstrated that on average, students had a slightly positive view of science and that it became more positive over the semester. In addition, the qualitative analysis using the Rasch scale demonstrated that the gains in both classes represented modest but discernible changes. For example, on average the CM group tended to be neutral on the pretest and somewhat in agreement on the delayed posttest that science helps in making sensible decisions. (A slight difference between the basic and Rasch-scaled results such as this is expected.)

These findings are interesting not just because they point to the success of the curricula—they would need to be further tested with other contrasting curricula—but also because they show that scientific attitudes may change. That students' beliefs about the relevance of science can be altered over one semester, simply by using realistic, issue-oriented science activities, seems worthwhile for educators to know. The result is especially critical given that students' attitudes about science are important determinants regarding their future involvement and performance in scientific classes and careers.

Both groups' attitudes changed as measured by reliable questionnaires. No significant differences were found between groups, although the CM group had numerically more positive attitudes and gains (Figure 2) about the relevance of science than the S group. Thus, as expected, no claim can be made that the use of CM and related activities designed to teach students about forming a coherent argument contributed to their attitude changes. A possible personal attention confound is that the CM group had a much smaller class size during the computer activities because they were split into two rooms. Even so, as alluded to above, additional data collected for the overall study revealed that CM students improved in terms of reasoning as well as attitude. Specifically, the use of evidence and the ability to weigh tradeoffs in decisions appeared to be enhanced, with a significant nonscaled gain between the pretest and delayed posttest of 1.48 (1.05 for the S group) (Siegel, 1999b). Although it seems reasonable that the addition of CM-type activities to the SEPUP curriculum would be beneficial for students in terms of both reasoning and attitudes, further studies would be required to draw more general conclusions about the added benefits of incorporating Convince Me.

We have reported (a) the development of a reliable set of questionnaire items for measuring scientific attitudes over time and (b) results suggesting it is possible to enhance students' attitudes that science is relevant by using innovative, issue-based activities. As educators develop new instruments and assess students, it is important to consider that change may not be immediate but may take time. We are curious about what changes occur over a longer period, such as students in an 8th-, 9th-, and 10th-grade sequence of SEPUP courses. Research involving how and why attitudes change is also promising. For instance, why exactly do we see a positive effect on students' beliefs toward science? Furthermore, we mentioned that both attitudes and performance on reasoning tasks (Siegel, 1999b) seemed to be enhanced during the entire study. Are the attitude and performance benefits for these disadvantaged students also correlated with persistence in science? A critical future step for education researchers will clearly include devising more developed, and more classroom-based, models to help explain the links among attitude, performance, and persistence.

This research was supported in part by a National Science Foundation Research Training Grant. However, the ideas expressed herein may not be representative of positions endorsed by the sponsoring agency. The authors especially thank Katharine Noonan, Herbert D. Thier, Marian C. Diamond, Mark Wilson, the Reasoning research group (particularly Janek Nelson), the SEPUP development team, and the students from Periods 3 and 4. Many thanks to our *JRST* editors and reviewers, and Edward Wolfe, Stephen Adams, and Florian Kaiser, for helpful suggestions and reviews.

Notes

[1]For example: "to be sustainable is to provide for food, fiber, and other natural and social resources needed for the survival of a group—such as a national or international society, an economic sector, or residential category—and to provide in a manner that maintains the essential resources for present and future generations" (Wimberley, 1993).

[2]Multiple analyses—for instance, repeated-measures ANOVA using averages to replace missing values—revealed the same pattern of results.

Appendix A

CARS Items

Students checked a box labeled with one of the following choices: "strongly agree," "somewhat agree," "neutral," "somewhat disagree," "strongly disagree," or "don't understand." Students were encouraged to inquire about items before checking "don't understand"; the few resulting responses were coded differently from nonresponses and missing data. Each questionnaire ended with: "Please comment on any of these issues in your own words" and "Thank you very much!"

Numbers below represent the same items shown in Appendix B and C. Eight repeated items, numbered 18–25, appeared on each version of the questionnaire.

	Repeated Items on Versions A, B, and C
18	Much of what I learn in science classes is useful in my everyday life today.
19	Learning science can help me when I pick food to buy.
20	Caring about people is part of making a scientific choice, such as whether to use pesticides on plants.
21	Science helps me to make sensible decisions.
22	The things I do in science have nothing to do with the real world.
23	Science helps me to make decisions that could affect my body.
24	Learning science will have an effect on the way I vote in elections
25	Making decisions can be difficult without reliable evidence.

	Items Only on Version A
1	My parents encourage me to continue with science.
2	I plan to take more science classes in high school.
3	Science helps me to work with others to find answers.
4	Science class helps me to evaluate my own work.
5	Learning science helps me understand about the environment.
6	Emotion has no place in science.
7	Science class helps me to judge other people's points of view.
8	Science will help me understand more about world-wide problems.
9	Science has nothing to do with my life outside of school.
10	Experiments in science help me to learn with a group.
11	Science teaches me to help others make decisions.
12	Knowing science will not help me in sports.
13	Science has nothing to do with buying things, such as food and cars.
14	Knowledge of science could make it easier to fix a bicycle.
15	Science teaches me to think less clearly than I already do.
16	Making a good decision is a scientific process.
17	Science class will help prepare me for college.
	Items Only on Version B
26	Science class helps me to work with others to make decisions.
27	I am interested in learning more about computer technology and designing video games.
28	Science has nothing to do with local issues, such as waste from nearby factories.
29	Science can help me make better decisions about what I buy.
30	Science experiments can help me to better understand the world.
31	I would like to learn more about strategies for thinking in my science class.
32	Knowledge of science helps me to prevent the spread of colds/diseases.
33	Using scientific methods helps me make environmental decisions.
34	Learning science is not important for my future success.
35	I only take science because it is a required course.
36	In most cases, personal feelings are important for making choices in science.
37	Knowing science can help me to make better choices about medical issues.
38	Collecting evidence is an important part of making a decision.
39	Science class will help prepare me for major decisions in my future.
40	I will only take math classes for as long as I have to.
41	Learning science enables me to explain my thoughts better to others.
42	Knowledge of science will help me protect the environment.
	Items Only on Version C
43	Science will help me to understand the effect I have on the environment.
44	Science helps me to ask others for help with my work.
45	Using scientific methods helps me think things through.
46	Science can help me decide how to treat my cold or illness.
47	Usually, it is bad to have any feelings about the scientific issues I am considering.
48	Science should be required in school.
49	Science could help me figure out how to spin/shoot/throw/hit a ball.
50	Science class helps me to evaluate my own work.
51	I do not expect to use science much when I get out of school.
52	I am interested in a career as a scientist or engineer.
53	Making decisions can be difficult when I don't understand the choices.
54	My intuition helps me make decisions in science.
55	I have support from others to excel at science.
56	Using scientific methods helps me decide what to buy in the store.
57	Science will help me understand the importance of recycling.
58	Learning science can help me understand about things that affect people's health.
59	Science can help me to make better choices about various things in my life (e.g., food to eat, car to buy).

Appendix B

Science Attitude Item Fit Statistics

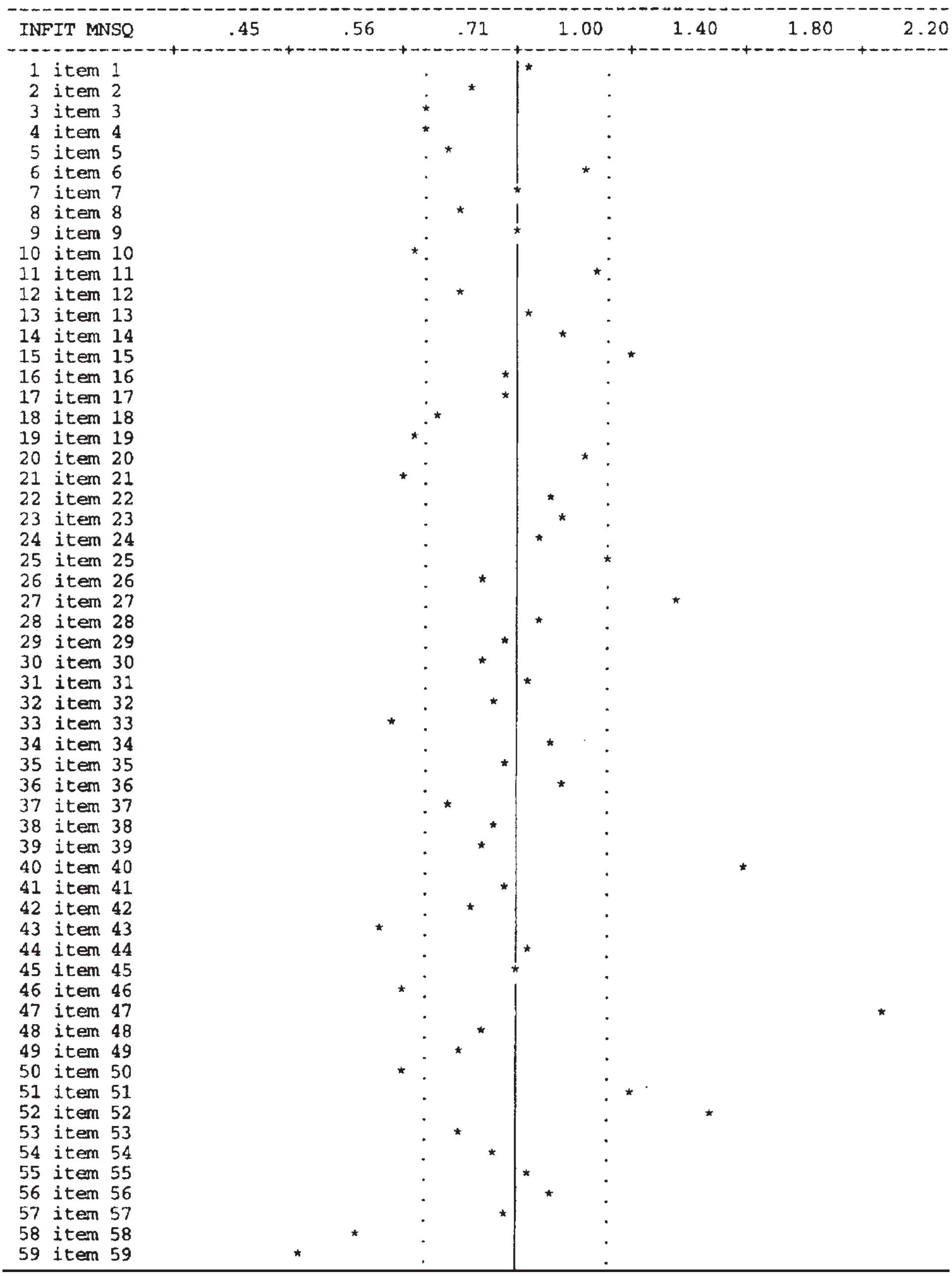

Figure B1. Item fit.

The proportion of items outside the dotted lines (Figure B1) is about 25%, which is large. Item 58 concerned the usefulness of science for making decisions about health and did not fit the model, possibly because the Science and Sustainability course did not include health directly. Item 47, which dealt with having feelings about scientific issues, also did not fit the model. Items 40, 52, and 59 also did not fit as well as predicted from pilot testing. It appears that students reacted strongly to these items that were about ‘‘taking math as long as I have to,’’ wanting a ‘‘career as a scientist,’’ and life choices. These items are perhaps not helpful indicators of students’ views of the relevance and usefulness of science for making decisions.

Infit Mean Square		Outfit Mean Square		Infit t		Outfit t	
Mean	1.01	Mean	1.06	Mean	0.02	Mean	0.15
SD	0.30	*SD*	0.40	*SD*	1.38	*SD*	1.41

As mentioned earlier, the questionnaire was a reasonably reliable instrument as indicated by the internal consistency of each scale being above .80, and for all 59 items being .91.

Appendix C

Interpretation of Science Attitude Logit Scale

This appendix offers a sense of what the logit score numbers mean. The qualitative descriptions in this appendix are representative of responses that were close to the characteristic logit score. The first column provides various representative students’ logit levels described in the other columns. The second column describes the attitude of the representative students. The third column gives examples of student responses to particular items.

Logit	Attitude Description	Examples of Student Response to Particular Statements
2.4	Very strongly believes that science is relevant to life	Strongly agree with: Science helps me to make sensible decisions (Item 21). Learning science can help me when I pick food to buy (Item 19). Strongly disagree with: The things I do in science have nothing to do with the real world (Item 22).
1.3	Believes that science is relevant to most areas of life.	Somewhat agree with: Science helps me to make sensible decisions (Item 21). Learning science can help me when I pick food to buy (Item 19). Neutral to: Learning science will have an effect on the way I vote in elections (Item 24).
0.5	Believes that science is sometimes relevant to life.	Somewhat agree with: Science helps me to make sensible decisions (Item 21). Learning science can help me when I pick food to buy (Item 19). The things I do in science have nothing to do with the real world (Item 22).
−0.8	Strongly believes that science is irrelevant to life	Somewhat disagree with: Science helps me to make sensible decisions (Item 21). Neutral to: The things I do in science have nothing to do with the real world (Item 22). Strongly disagree with: Learning science will have an effect on the way I vote in elections (Item 24).

References

Adams, S. (1998). What is “good reasoning” about global warming? A comparison of high school students and specialists. Doctoral dissertation, University of California at Berkeley.

Adams, R.J. & Khoo, S.-T. (1993). Quest: The interactive test analysis system. Hawthorn, Victoria, Australia: Australian Council for Educational Research.

Andre, T., Whigham, M., Hendrickson, A., & Chambers, S. (1999). Competency beliefs, positive affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. Journal of Research in Science Teaching, 36, 719–747.

Aronson, E. (1995). The social animal. New York: W.H. Freeman.

Bell, P. & Linn, M.C. (2000). Scientific arguments as learning artifacts: Designing for learning on the Web in KIE. International Journal of Science Education, 22, 797–817.

Bell, P. & Linn, M.C. (2002). Beliefs about science: How does science instruction contribute? In Hofer, B.K. & Pintrich, P.R. (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (pp. 321–346). Mahwah, NJ: Erlbaum.

Butler, M.B. (1999). Factors associated with students’ intentions to engage in science learning activities. Journal of Research in Science Teaching, 36, 455–473.

Conwell, C.R. & Prichard, M.K. (1992). Expanding students’ horizons in mathematics and science. School Science and Mathematics, 92, 267–271.

Cook, L.L. & Eignor, D.R. (1991). NCME Instructional module: IRT equating methods. Educational Measurement: Issues and Practice, 10, 37–45.

Crawley, F.E. & Koballa, T.R. (1994). Attitude research in science education: Contemporary models and methods. Science Education, 78, 35–55.

CTB. (1997). TerraNova. CTB/McGraw-Hill.

Davis, E.A. (1998). Scaffolding students’ reflection for science learning. Unpublished doctoral dissertation, University of California at Berkeley.

Dickinson, D.H. (1976). Community college students’ achievement and attitude change in a lecture/laboratory approach to general education biological science courses (Doctoral dissertation, Utah State University, 1975). Dissertation Abstracts International, 36, 5968A.

Diehl, C., Castro, S., & Ranney, M. (1997, March). Student models of hypothesis and evidence. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Diehl, C., Ranney, M., & Schank, P. (2001). Model-based feedback supports reflective activity in collaborative argumentation. In Dillenbourg, P., Eurelings, A., & Hakkarainen, K. (Eds.), European perspectives on computer supported collaborative learning: Proceedings of the First European Conference on Computer Supported Collaborative Learning (pp. 189–196). Maastricht, The Netherlands: Universiteit Maastricht.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). Young people’s images of science. Buckingham: Open University Press.

Fraser, B.J. (1980). Science teacher characteristics and student attitudinal outcomes. School Science and Mathematics, 80, 300–308.

Freedman, M.P. (1997). Relationship among laboratory instruction, attitude toward science, and achievement in science knowledge. Journal of Research in Science Teaching, 34, 343–357.

Hegarty-Hazel, E. (1990). The student laboratory and the science curriculum: A model. In Hegarty-Hazel, E. (Ed.), The student laboratory and the science curriculum (pp. 27–30). New York: Routledge.

Iskandar, S.M. (1991). An evaluation of the science–technology–society approach to science teaching. Unpublished doctoral dissertation, University of Iowa, Iowa City.

Kahle, J.B. & Lakes, M.K. (1983). The myth of equality in science classrooms. Journal of Research in Science Teaching, 20, 131–140.

Keeves, J.P. & Morgenstern, C. (1992). Attitudes towards science: Measures and effects. In Keeves, J.P. (Ed.), The IEA Study of Science III: Changes in science education and achievement: 1970 to 1984 (pp. 122–142). New York: Pergamon.

King, P.M. & Kitchener, K.S. (1994). Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults. San Francisco: Jossey-Bass.

King, P.M. & Kitchener, K.S. (2002). The reflective judgment model: Twenty years of research on epistemic cognition. In Hofer, B.K. & Pintrich, P.R. (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (pp. 37–61). Mahwah, NJ: Erlbaum.

Koballa, T.R. (1988). Attitude and related concepts in science education. Science Education, 72, 115–126.

Kuhn, D. (1991). The skills of argument. New York: Cambridge University Press.

Lederman, N., Abd-El-Khalick, F., Bell, R.L., & Schwartz, R.S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. Journal of Research in Science Teaching, 39, 497–521.

Linn, M.C. & Songer, N.B. (1993). How do students make sense of science? Merrill-Palmer Quarterly, 39, 47–73.

Lord, F.M. (1980). Applications of item response theory to practical testing problems. Hillsdale, NJ: Erlbaum.

Marsh, H.W. & Yeung, A.S. (1997). Coursework selection: Relations to academic self-concept and achievement. American Educational Research Journal, 34, 691–720.

Masters, G.N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47, 149–174.

Mattern, N. & Schau, C. (2002). Gender differences in science attitude–achievement relationships over time among White middle-class students. Journal of Research in Science Teaching, 39, 324–340.

McComas, W.F. (1993). STS education and the affective domain. In Yager, R.E. (Ed.), What research says to the science teacher, Volume 7: The science, technology and society movement (pp. 161–168). Washington, DC: National Science Teachers Association.

McComas, W.F. (1996). The affective domain and STS instruction. In Yager, R.E. (Ed.), Science/Technology/Society as reform in science education (pp. 70–83). Albany: SUNY Press.

Mead, M. & Metraux, R. (1957). Image of the scientist among high school students. Science, 126, 384–390.

Moore, R. & Foy, R. (1997). The scientific attitude inventory: A revision (SAI II). Journal of Research in Science Teaching, 34, 327–336.

Moore, R. & Sutman, F. (1970). The development, field test and validation of an inventory of scientific attitudes. Journal of Research in Science Teaching, 7, 85–94.

Neathery, M.F. (1997). Elementary and secondary students' perceptions toward science and the correlation with gender, ethnicity, ability, grade, and science achievement. Electronic Journal of Science Education, 2(1).

Oppenheim, A.N. (1992). Questionnaire design, interviewing and attitude measurement. London: Pinter, 200–210.

Piburn, M.D. & Baker, D.R. (1993). If I were the teacher . . . Qualitative study of attitude toward science. Science Education, 77, 393–406.

Ramsden, J.M. (1998). Mission impossible? Can anything be done about attitudes to science? International Journal of Science Education, 20, 125–137.

Ranney, M. (1996). Individual-centered vs. model-centered approaches to consistency: A dimension for considering human rationality. Vivek, A Quarterly in Artificial Intelligence, 9, 35–43.

Ranney, M. & Schank, P. (1998). Toward an integration of the social and the scientific: Observing, modeling and promoting the explanatory coherence of reasoning. In Read, S. & Miller, L. (Eds.), Connectionist models of social reasoning and social behavior (pp. 245–274). Mahwah, NJ: Erlbaum.

Ranney, M., Schank, P., & Diehl, C. (1995). Competence and performance in critical reasoning: Reducing the gap by using Convince Me. Psychology Teaching Review, 4, 153–166.

Ranney, M., Schank, P., Hoadley, C., & Neff, J. (1994, October). "I know one when I see one": How (much) do hypotheses differ from evidence? In: Proceedings of the Fifth Annual American Society for Information Science Workshop on Classification Research, Alexandria, VA, pp. 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 (pp. 426–432). Hillsdale, NJ: Erlbaum.

Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen: Danmarks Paedogogiske Institut, 1980.

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

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

Roberts, L., Sloane, K., & Wilson, M. (1996, April). Local assessment moderation in SEPUP. Presented at the annual meeting of the American Educational Research Association, New York, NY.

Roberts, L. & Wilson, M. (1998). An integrated assessment system as a medium for teacher change and the organizational factors that mediate science teachers' professional development. BEAR Report Series SA-98-2. Berkeley, CA: University of California.

Ryan, A.G. & Aikenhead, G.S. (1992). Students' preconceptions about the epistemology of science. Science Education, 76, 559–580.

Schank, P.K. (1995). Computational tools for modeling and aiding reasoning: Assessing and applying the theory of explanatory coherence. Doctoral dissertation, University of California, Berkeley.

Schank, P. & Ranney, M. (1991). An empirical investigation of the psychological fidelity of ECHO: Modeling an experimental study of explanatory coherence. In K.J. Hammond & D. Gentner (Eds.), Proceedings of the 13th Annual Conference of the Cognitive Science Society (pp. 892–897). Hillsdale, NJ: Erlbaum.

Schank, P. & Ranney, M. (1992). Assessing explanatory coherence: A new method for integrating verbal data of on-line belief revisions. In J.K. Kruschke (Ed.), Proceedings of the 14th Annual Conference of the Cognitive Science Society (pp. 599–604). Hillsdale, NJ: Erlbaum.

Schank, P., Ranney, M., & Hoadley, C. (1996). Convince Me [Revised computer program (on CD) and manual]. In Jungck, J.R., Vaughan, V., Calley, J.N., Peterson, N.S., Soderberg, P., & Stewart, J. (Eds.), The 1996–1997 BioQUEST Library (4th ed.) College Park, MD: Academic Software Development Group, University of Maryland.

Schibeci, R.A. & Riley, J.P. Jr. (1986). Influence of students' background and perceptions on science attitudes and achievement. Journal of Research in Science Teaching, 23, 177–187.

Schoenfeld, A.H. (1985). Mathematical problem solving. Orlando, FL: Academic Press.

Schommer, M. (1994). Synthesizing epistemological belief research: Tentative understandings and provocative confusions. Educational Psychology Review, 6, 293–319.

Science Education for Public Understanding Program's Science and Sustainability (SEPUP). (1999a). Science Education for Public Understanding Program's Science and Sustainability course. University of California, Berkeley. [On-line]. Available: http://sepuplhs.org

SEPUP. (1999b). Science and sustainability. Ronkonkoma, NY: Lab-Aids.

Shamai, S. (1996). Elementary school students' attitudes toward science and their course of studies in high school. Adolescence, 31, 677–690.

Siegel, M.A. (1999a). Teaching science for public understanding: Developing decision-making abilities. Doctoral dissertation, University of California, Berkeley.

Siegel, M.A. (1999b). Changes in student decisions with Convince Me: Using evidence and making tradeoffs. In M. Hahn & S.C. Stoness (Eds.), Proceedings of the 21st Annual Conference of the Cognitive Science Society (pp. 671–676). Mahwah, NJ: Erlbaum.

Simpson, R.D. & Oliver, J.S. (1990). A summary of major influences on attitude toward and achievement in science among adolescent students. Science Education, 74, 1–18.

Song, J. & Kim, K.-S. (1999). How Korean students see scientists: The images of the scientist. International Journal of Science Education, 21, 957–977.

Stake, J.E. & Mares, K.R. (2001). Science enrichment programs for gifted high school girls and boys: Predictors of program impact on science confidence and motivation. Journal of Research in Science Teaching, 38, 1065–1088.

Tamir, P. & Amir, R. (1987). The relationship between instructional strategies, study practices, and attitudes toward biology. Journal of Biological Education, 21, 291–295.

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

Thagard, P. & Millgram, E. (1995). Inference to the best plan: A coherence theory of decision. In Ram, A. & Leake D.B. (Eds.), Goal-driven learning. Cambridge: MIT Press.

Thier, H. & Nagle, B. (1994). Developing a model for issue-oriented science. In Solomon, J. & Aikenhead, G. (Eds.), STS education: International perspectives on reform. Ways of Knowing Science series (pp. 75–83). New York: Teachers College Press.

Third International Mathematics and Science Study. (2001). TIMSS and TIMSS-R [On-line]. Available: http://nces.ed.gov/timss/

Weinburgh, M.H. (2000). Gender, ethnicity, and grade level as predictors of middle school students' attitudes toward science. Georgia State University, GA. (ERIC Document Reproduction Service No. ED442662).

Wilson, M., Sloane, K., Roberts, L., & Henke, R. (1995). SEPUP Course I, Issues, Evidence and You: Achievement evidence from the pilot implementation. University of California, Berkeley. (BEAR Report Series, SA-95-2).

Wimberley, R.C. (1993). Policy perspectives on social, agricultural and rural sustainability. Rural Sociology, 58 (1), 1–29.