The Development of Scientific Thinking and Conceptions of Science in College Science Students

FINAL REPORT OF ACTIVITIES & FINDINGS THROUGH AUG 2004

Funded by the Research in Educational Policy & Practice Program
National Science Foundation (NSF Grant No. REC-9980519)

Neil Stillings, PI, Hampshire College
Mary Anne Ramirez, co-PI, Hampshire College
Carol Smith, co-PI, University of Massachusetts, Boston

Major Research & Education Activities

Overview

Our research investigated general cognitive outcomes of college science instruction. College faculty pursuing inquiry-oriented instructional reforms hope that their students will master both concepts and general, or higher-order, scientific reasoning skills that will transfer across content domains. They also hope that students will develop sophisticated views of the nature of science, including an understanding of the role of theories and models in explanation and of the interplay between theory and evidence in the development of research programs and controversies. In this project we have pursued several lines of investigation:

- Epistemological development project. In this project we developed a new version of Carey & Smith's Nature-of-Science Interview (NOS) to study epistemological development in college students. We looked at patterns of responding and change over one semester in a number of courses and, in a project that is still ongoing, followed a longitudinal sample during their college careers.

- Scientific reasoning project. In this project we developed an essay-style assessment of scientific reasoning ability that was administered to students in a number of reform-oriented college and university courses. Detailed classroom observations were also made in these courses, using a classroom observation protocol that we developed for coding inquiry-oriented instructional and learning activities.

- Intervention projects: We developed collaborations with three faculty groups, in which we used results from the scientific reasoning project to design new teaching strategies and learning activities:

  - *Reading primary literature in first-year science courses*: The Hampshire College Natural Science faculty pioneered the use of primary literature in first-year college courses in a range of fields. In this project we worked with ten members of the faculty to improve this innovative practice and to assess student outcomes more carefully.

- *Model-based reasoning in large-lecture university biology*: In this project we worked with the introductory biology faculty at the University of Massachusetts, Amherst to redesign the lecture portion of the course around reasoning with core biological models.

- *Software-based model construction in introductory chemistry*: In this project we worked with William Vining and two other chemistry faculty at the University of Massachusetts, Amherst to refine the instruction and assessment in small introductory honors sections that are taught in an electronic classroom using Vining's ChemLand software.

The following five sections give further detail on the activities pursued in these projects.

Epistemological development project: Activities

This is the first project to look systematically at the performance of college students on Carey and Smith's Nature-of-Science (NOS) interview, which had previously been administered to pre-college students. During the first year and a half of the project the scoring system for the NOS was extended and refined to meet the needs of the project. Two sections were also added to the NOS, which deals mainly with students' understanding of the role of hypotheses and theories in scientific explanation. The first added section deals with students' conceptions of scientific controversy and uncertainty, and the second surveys their opinions concerning how their college science courses have affected their ideas about the nature of science. A reliable scoring system was developed for the controversy section of the interview during the first year and a half of the project.

Each fall, beginning in 2000, we interviewed a cohort of 45-60 incoming students in Hampshire College Natural Science courses and in introductory courses in the University of Massachusetts Biology and Chemistry Departments. The students were drawn from courses that had a range of reform characteristics. Students were then interviewed again in the spring of their first years and once in each of the following years, creating a series of cascading longitudinal cohorts. The interviews were conducted by two of the project post-doctoral fellows, Laura Wenk and Loel Tronsky, and by research associate Faith Conant. The transcripts were scored by Wenk, Co-PI Carol Smith, and Carolyn Houghton, a long-time research associate of Smith's. Transcripts were checked for accuracy, and the inter-scorer reliability of the interview coding was checked extensively. Interviews continued through the no-cost extension year of 2003-04 and the post-grant year of 2004-05. Data analysis will extend into 2005-06.

Scientific reasoning project: Activities

During the first year of the project we developed an essay-style assessment of scientific thinking skills that could be administered in an hour. The questions on the instrument make use of very simple scientific scenarios that do not depend on any field-specific knowledge. The four questions assess the following areas:

- Generating ideas and stating testable hypotheses
- Critiquing empirical results from a flawed experimental design

- Making inferences from graphically displayed data that contains error variance
- Responding to scientific controversy

In pilot work a reliable scoring protocol was developed and applied to the student essays by co-PI Mary Anne Ramirez and graduate students Samia Khan and Alisa Izumi. In 2000 and 2001 the assessment was administered pre and post semester to samples of entering students at Hampshire College and in introductory biology and chemistry courses at the University of Massachusetts.

For the courses from which the students were drawn, a sample of class meetings was observed by Izumi, Khan, Ramirez, or PI Stillings. Observers recorded their observations on a Method-by-Goal grid that was designed around inquiry-oriented instruction. Coding was tested for reliability by having multiple, independent observers in a sample of class meetings.

Intervention 1: Reading primary literature in first-year undergraduate science: Activities
A hallmark of first-year science courses at Hampshire College is assignments that ask students to read and critically respond to primary scientific articles. First-year science courses at Hampshire are not survey courses. Rather, they delve into more restricted topics in a way that allows students to appreciate the conceptual issues and begin to do their own research projects. The faculty have been successful in designing courses that involve beginning students in the conduct of science. With the exception of physics, in which secondary sources are often used, all courses involve reading primary literature that is accessible, with appropriate scaffolding, to beginning students. For example, students in chemistry courses read papers with significant chemistry content in environmental chemistry, forensic chemistry, or biomedical research.

In this study we researched the student outcomes from working with primary literature, and we worked with faculty members to refine their instructional practice. The assessment assignments asked students to respond pre and post-semester to brief, novel primary articles by answering a seven-question essay assignment. The assessment was administered in the fall semesters of 2001, 2002, and 2003. In 2002-03 we completed the development of a scoring rubric for this assessment that is reliable and practical for use by faculty members. Post-doctoral fellows Laura Wenk and Loel Tronsky led this project with the assistance of our research associate Faith Conant.

The intervention with faculty members centered on three issues: (1) Clarifying the learning goals associated with the practice; (2) Refining and insuring implementation of appropriate scaffolding for the first-year college student attempting to read a journal article; and (3) Assessing students for progress on the learning goals.

Intervention 2: Model-based reasoning in biology: Activities
Our results from the scientific reasoning project led to the conclusion that the reforms in the UMass biology course needed more focus and consistency. In 2002-03 & 2003-04 Stillings and Ramirez worked with the biology faculty to refine their learning goals for students and to strongly align their instruction and assessment with the goals. Learning to reason with the core theoretical models of biology (e.g. gene expression, cellular metabolism) has become the integrating theme of the lecture part of the course. The instructors have become full partners in

this research effort. The instructional practices for the first semester (molecular and cell biology) of the course have now gone through a design, test, redesign, test cycle. Assessment practices are still being refined.

In the redesign of the course, research-oriented pre-post assessments were co-developed by the research team and the instructors to follow the students' development of model-based reasoning skills. The assessments combined essay-style items, multiple-choice items, and hybrid multiple-choice items in which students must defend their selection of an answer with a short essay. Ramirez and Izumi developed reliable scoring rubrics for the essay material and scored the assessments for the quality of students' causal reasoning.

The redesigned course makes extensive use of a peer-learning model, in which, several times per class meeting, small groups of students discuss problems posed by the instructor, followed by whole-class discussion and summarizing comments by the instructor. To investigate the dynamics of this process we videotaped the small group and whole-class discussions for a semester in a small section that met in a room set up with multiple cameras and microphones. Ramirez and Izumi reported an analysis of the whole class discussions and did an initial analysis of the small group discussions.

Intervention 3: Software-based model construction in introductory chemistry: Activities
In this project we investigated the instructional techniques and learning outcomes in chemistry courses in which the class meetings were organized around student exploration of the ChemLand software modules, developed by Prof. Bill Vining of the University of Massachusetts Chemistry Department. Each module is a small simulation that allows students to interactively explore, graphically and/or quantitatively, the relationships among two or more variables, which are governed by physical laws. In the classroom, pairs or triples of students work with the simulations, under structured prompting from the instructor, to try to understand the relationships and generate candidate underlying causal models. This model-generation activity serves as a bridge between students' initial conceptions and the later, more formal treatment of the relevant laws and processes by the instructor or textbook. An example is a simulation of boiling points, which allows students to vary temperature, molecular weight, and molecular composition. Students are able to discover that boiling point is not fully determined by molecular weight, a common initial conception, and they can begin to hypothesize the existence of other factors that are related to chemical structure.

The methodologies for this study included the structured classroom observation and scientific thinking survey from our scientific reasoning project, student attitude surveys, and a laboratory study in which pairs of students and the instructor were taped as the students explored two of the simulations under structured guidance from Prof. Vining. The intervention in this project involved making Vining's instructional scaffolding of student use of the software more explicit so that other instructors in the department could learn the techniques. The chemistry project was the doctoral research of graduate student Samia Khan.

Major Findings

The findings, thus far, from each of the five projects are described below. For each project, the associated theses and major conference presentations are listed.

Epistemological development project: Findings

This project has so far led to four conference presentations and a doctoral dissertation. Other publications are submitted or forthcoming.

Wenk, L. (1999). Developmental measures as evaluation tools for inquiry science programs. Presented at the annual meeting of the National Association for Research in Science Teaching (NARST), Boston, MA March 28-31, 1999. (Reported work done during the proposal phase of the project, prior to the receipt of funds.)

Wenk, L. (2000). Improving science learning: Inquiry-based and traditional first-year college science curricula. Doctoral dissertation, School of Education, University of Massachusetts, Amherst.

Smith, C. & Wenk, L. (2003). The Relations Among Three Aspects of College Freshmen's Epistemologies of Science. Presented at the annual meeting of the National Association for Research in Science Teaching (NARST), Philadelphia, March 23-26, 2003.

Wenk, L. (2004). Inquiry and Epistemology in Science: Implications for Instruction in the Sciences and Beyond. Invited presentation, American Association of Higher Education (AAHE) Focus on Learning: What Faculty and Administrators Need to Know about Learning February 20-22.

Smith, C. & Wenk, L. (2004). The Impact of 1st-Year College Science Courses on Epistemological Thinking: A Comparative Study. Presented at the annual meeting of the National Association for Research in Science Teaching (NARST), Vancouver, BC, March 31-April 3, 2004.

Smith, C. & Wenk, L. (Submitted to *Journal of Research in Science Teaching*). Relations Among Three Aspects of First-Year College Students' Epistemologies of Science.

<u>New instrument for assessing epistemological development</u>: We believe that our new Nature-of-Science interview with its scoring protocols is a major advance over the epistemological interviews used in previous research on undergraduates, which do not probe crucial features of science, such as the role of theories in scientific research and explanation. The painstaking work of establishing the reliability of the scoring system and coding enough interviews to look at an initial pattern of results was completed in Fall 2002. The new NOS interview and the initial results were described in major conference presentations (Smith & Wenk, 2003, 2004). A revised version of the 2003 paper has been submitted to the Journal of Research in Science Teaching, and the 2004 paper is also being revised for publication.

<u>Cross-sectional findings</u>: One initial hypothesis in this project was that inquiry-oriented science instruction would produce observable epistemological change over one semester of instruction. In Wenk's (2000) dissertation this hypothesis was supported for Hampshire College students in inquiry science courses when compared to students at a comparison college. Our data since then (Smith & Wenk, 2004) have supported this hypothesis at Hampshire but not in the inquiry-

oriented courses that we studied at the University of Massachusetts. The one-semester gains at Hampshire are also fairly modest. It appears that the nature of science must be explicitly and consistently addressed in introductory inquiry-oriented courses to produce reliable change in students' explicit conceptions of science over one semester. And these changes should be thought of as setting the stage for further epistemological development, which may require continued scaffolding in more advanced courses. This issue will be addressed more thoroughly as we continue to analyze our longitudinal data.

Smith & Wenk (2003) also found significant patterns within the new NOS. Student responses across the three parts of the interview were significantly related. Consistently differentiating theory and evidence was strongly related to appreciating the inherent uncertainty of scientific knowledge and with having a deeper understanding of the reasons for scientific controversies and how to resolve them. Our initial interpretation of this finding is that a metacognitive appreciation of the theory-evidence distinction is a prerequisite to reasoning about scientific uncertainty and controversy. Our results indicate the many beginning college students are still making the transition to possessing a cognitively-stable, reflective ability to distinguish theory from evidence. For many students, with greater or lesser probability, science is a collection of "findings," or "results," or "observations," accompanied by a limited ability to reflect on how the findings relate to an underlying hypothesis or to a broader theoretical framework.

<u>Developmental findings</u>: Longitudinal data collection for the project continued through the Spring of 2004 and data analysis is continuing. Our hypotheses are that epistemological development will be related to persistence as a science student through college and to significant experiences with scientific inquiry.

Scientific reasoning project: Findings

This project led to five conference presentations and a dissertation:

Stillings, N. A., Ramirez, M. A., & Wenk, L. (1999). Assessing critical thinking in a student-active science curriculum. Presented at NARST meeting, Boston, MA March 28-31, 1999. (Reported work done during the proposal phase of the project, prior to the award of funds.)

Stillings, N. A., Ramirez, M. A., & Wenk, L. (2000).Teaching and learning the nature of science in inquiry-oriented college science courses. Presented at the 2000 AERA meeting, New Orleans, LA, April 24-28, 2000.

Rea-Ramirez , M. A. (2001). Reform in Higher Education Introductory Science Classes: Integration of Strategies to Promote Scientific Literacy. Symposium, M. A. Rea-Ramirez, Organizer. Annual meeting of NARST, St. Louis, MO, March 25-28, 2001

Gibson, H. L., Bernhard, J., Kropf, A, Rea-Ramirez, M. A., Van Strat, G. A. (2001). Enhancing the science literacy of preservice teachers through the use of reflective journals. Presented at NARST, St. Louis, MO, March 25-28, 2001

Stillings, N. A. (2002). An approach to software for college-level inquiry learning. Invited paper presented at Instructional Design for New Technology-Enabled Approaches to Learning, a national workshop sponsored by the Learning Federation, Dec. 5-6, 2002, Orlando, FL.

Izumi, A. (2003). Measuring measurements of scientific thinking: Do multiple forms of assessment exhibit similar evidence of student understanding on complex scientific reasoning problems. Doctoral dissertation, School of Education, University of Massachusetts, Amherst.

Classroom observation: Stillings et al. (2000) reported wide variation in the implementation of inquiry-oriented instruction across the courses observed. Although all courses observed showed large differences from conventional instruction, the inquiry courses varied widely both in the amount of time devoted to particular inquiry goals and in the instructional methods employed. An important background finding was that most of the reform-oriented science courses at the University of Massachusetts that were considered for inclusion in the research could not be included because inquiry-oriented activities were not a significant part of the instruction. This finding increased the importance of the intervention phase of the project considerably.

Scientific reasoning: The hypothesis underlying the development of the essay-style assessment of scientific reasoning was that inquiry-oriented instruction is related to increases in general scientific reasoning skill. To some degree our results supported this hypothesis. In Stillings et al. (1999), which reported work that began at the time of the proposal for the project, we found gains in Hampshire students enrolled in inquiry-oriented science courses but not in Hampshire students who were not enrolled in science courses or in students at a comparison college who were enrolled in more conventionally-taught science courses. We continued to find gains in Hampshire students in a second study, but the gains also occurred in the sample of students who were not enrolled in science courses. The best interpretation of this finding is that courses in the social and cognitive sciences at Hampshire also have an inquiry orientation and that because of Hampshire's first-year distribution requirement, the sample of comparison students who were not enrolled in natural science courses had significant experiences with scientific inquiry, broadly defined, in other courses. Gains in inquiry-oriented courses at the University of Massachusetts have sometimes either not occurred or have been comparable to changes in comparison groups. We found that the variability of results was higher at the University than in college settings because of variability in student motivation, as measured by the number of questions completed and the lengths of answers at the beginning vs. the end of the assessment. These results, again, heightened the importance of our intervention projects. It became clear that in the state university setting reasoning goals had to be defined more narrowly and pursued more aggressively in instruction and student assessment than the faculty members had realized. In small college settings, such as Hampshire, where the classes are smaller and where the typical student may have a more intellectual orientation, the general scientific reasoning goals are realistic, but the gains that we observed were small enough that it is clear that more care in the design of instruction and assessment could be productive.

In terms of the existing literature in science education we consider these results to be of great significance. In most published studies the researchers exert extensive control over instruction and carefully design their assessments to detect specific effects of the instruction. We began our project studying naturally-occurring educational reforms that were designed and implemented by committed faculty members and by administering assessments that were not specifically tied to their instruction. Our results suggest that there is less generality and transfer in the student outcomes of inquiry-oriented instruction than is often suggested.

The results of our scientific reasoning research served mainly to promote the three intervention projects that followed. Using independent assessments, we were able to show faculty members that they were not consistently achieving some of the things that they most wanted to achieve with students. This led to three strong partnerships.

Intervention 1: Reading primary literature in first-year undergraduate science: Findings
This project has produced three conference presentations and a manuscript to be submitted. Laura Wenk's (2000) doctoral thesis also contributed to the development of the project.

Wenk, L. (2001). Improving students' understanding of the nature of science: Primary literature in the introductory college science course. Presented at the 2001 NARST Annual Meeting, St. Louis, MO, March 25-28, 2001.

Wenk, L. (2002). Improving students' understanding of the nature of science: inquiry-based instruction in introductory college science courses. Paper presented at Pathways to Change: An International Conference on Transforming Math & Science Education in the K16 Continuum, Washington, DC, April 18-22, 2002

Wenk, L., Tronsky, L., McNeal, A., and Conant, F. (2005, forthcoming). Improving College Students' Primary Literature Skills: Using Primary Literature in First Year Science Courses. To be presented at the AERA annual meeting, Spring, 2005.

Wenk, L., Tronsky, L., McNeal, A., and Conant, F. (To be submitted to the *International Journal of Science Education*). First-year college students benefit from reading primary scientific articles.

From Wenk (2000, 2001), from classroom observations (Stillings et al., 2000), and from discussions with the faculty, it seemed clear that a significant practice in Hampshire College's inquiry-oriented first-year science courses is assignments that involve students in critically reading primary journal articles. Students reported fairly frequently in our interviews that their attitudes about science were deeply affected by reading primary articles. On the other hand our classroom observations confirmed a faculty complaint that the practice was implemented in widely different ways across courses and was not always successful.

To examine student outcomes associated with reading primary literature, the research team developed a structured assessment that can be administered to students as a response assignment for their reading of a primary article. In the assessment the student writes a critical response to a novel primary journal article, addressing the following areas:

- Questions addressed by the research. Theoretical context of the research.
- Importance of the research.
- The theory or hypothesis that was investigated in the study.
- The research design of the study. The data that were collected.
- Analysis of how well the results support the study's theory or hypothesis.
- Possible alternative explanations for the findings and changes in the way the research was done that might help decide among the alternatives.
- Other further research suggested by this study.

The analysis of student papers from the first two terms in which this assessment was given pre and post-semester produced a reliable scoring rubric. In part the rubric involves categorizing each statement in a student's paper according to the seven headings above and rating the maturity or sophistication of the statement on a 3-point scale according to criteria for each category. For example, if the research design of a study includes control conditions, the most sophisticated answers mention the control conditions and explain why they are present in the design. Less sophisticated answers fail to explain the importance of the conditions or fail to mention them at all.

We found a strong relationship between student progress in understanding primary research papers and classroom pedagogy. Students showed progress in their ability to interpret primary articles in a course that required students to read a number of research articles and in which the instructor explicitly taught the requisite interpretive skills. Students showed unreliable or no progress in courses that devoted less time to reading research papers and to reflecting on how to understand them. These results will be reported in the forthcoming conference presentation and journal submission listed above.

This research project essentially became an example of a design experiment in inquiry-oriented science instruction, with the assessment results guiding our intervention with the participating faculty members. The assessment rubric itself became the basis of several improved classroom practices.

Intervention 2: Model-based reasoning in biology: Findings
This project has resulted in three conference papers, an unpublished paper, grant-supported dissemination within the University of Massachusetts, numerous workshop presentations, a forthcoming manual for teachers, a dissertation, and further grant proposals:

Stillings, N. A. (2002). Advancing our understanding and assessment of student learning through classroom-oriented research partnerships. Paper presented at Pathways to Change: An International Conference on Transforming Math & Science Education in the K16 Continuum, Washington, DC, April 18-22, 2002.

Rea-Ramirez, M. & A. Izumi (2003a). Empirical Evidence On Student Understanding Of Biological Models: Fat Metabolism. Presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia, March 23-26, 2003.

Khan, S., Rea-Ramirez, M. and Izumi, A. (2004). Instructor Scaffolded Model Construction and Model Based Reasoning in College Introductory Biology. Presented at the annual meeting of the National Association for Research in Science Teaching (NARST), Vancouver, BC, March 31-April 3, 2004.

Rea-Ramirez, M. & A. Izumi (to be submitted). Instructor Scaffolded Model Construction and Model Based Reasoning in a College Introductory Biology Class.

Izumi, A. (2003). Measuring measurements of scientific thinking: Do multiple forms of assessment exhibit similar evidence of student understanding on complex scientific reasoning problems. Doctoral dissertation, School of Education, University of Massachusetts, Amherst.

Phillis, R. (forthcoming). *A Guide to Promoting Student Reasoning and Problem-Solving Skills in the Introductory Biology Classroom*. Prentice-Hall.

The completion of the learning materials for a biology course organized around reasoning with theoretical models is a significant accomplishment. An initial set of several hundred questions has been created and tested in the design cycle. The questions are organized along several dimensions: (1) Core biological model(s) involved; (2) Relation to higher-level themes that motivate sequences of questions and reinforce the concept map of the field (e.g. cell diversity, or mechanisms of cancer); (3) Type and complexity of reasoning typically required to generate an answer; (4) Suitability for different learning and assessment activities, such as peer learning, quizzes, or exams. Faculty collaborator Randy Phillis is completing a teacher's manual for active learning and the promotion of reasoning skills in introductory biology, to be published by Prentice-Hall. The University also received a $295,000 grant from the Davis Educational Foundation, titled Creating Active Learning Through Technology, to disseminate the model-based reasoning reforms adopted in this project, as well as the use of interactive web-based quizzes and the use of student Personal Response Systems (PRS) in the lecture hall.

In the fall of 2001, pre-post testing with hybrid multiple-choice essay questions showed significant student gains in model-based reasoning skills from the beginning to the end of the semester (Stillings, 2002). This finding contrasts with our earlier failure to find gains in general scientific reasoning skills. We attribute the improvement in results to focusing the learning goals and instruction in the course more carefully and restricting the assessment to reasoning with concepts presented in the course.

Rea-Ramirez & Izumi (2003a) reported an additional significant methodological finding from the hybrid multiple-choice essay assessments. In these questions we presented students with a model-based scenario followed by several increasingly complex M-C questions concerning the scenario. Students were required to defend each of their M-C answers with a short essay, which was scored via a reliable rubric. We found low correlations between students' M-C answers and their essay scores. The lack of correlation was partly attributable to some students' inability to defend correct M-C answers, but an additional contribution to the effect came from students who wrote good essays defending incorrect M-C selections. In some of the latter cases students clearly had misread their choice, i.e. their reasoning was correct, but they did not select the answer that accorded with their reasoning. In other cases they gave well-reasoned defenses of incorrect choices that reflected gaps in their factual knowledge, i.e. the student reasoned well with a factually incorrect model.

Rea-Ramirez & Izumi (2003b) is an initial report on the whole-class discussion aspect of our study of classroom dynamics. The results show very strong trends over the course of the semester of the instructor speaking less and removing scaffolding and of the students making increasing use of reasoning strategies. We also found that students (re)construct models while they are solving problems. That is, as they attempt to apply a model to a problem they realize gaps in their understanding and work to fill them in, in order to complete their reasoning. This finding provides strong evidence that the approach we have developed is not just a matter of knowing or guessing the answers to problems or questions. Many of our problems are written in a way that students can only attack them using their understanding of underlying causal mechanisms.

Our faculty participants Prof.s Randy Phillis and Steve Goodwin have disseminated the work widely around the country in presentations to biology teaching workshops (at e.g. Arizona State University, Johns Hopkins University, University of Wisconsin, Cal. State Fullerton, and W.H. Freeman).

We are currently seeking funds for research to strongly validate the student outcomes for the instructional design that we have arrived at.

Intervention 3: Software-based model construction in introductory chemistry: Findings
This is a mature intervention project that has produced a journal publication, a doctoral thesis, and eleven conference presentations:

Fermann, J. T., et al. (2000). Discovery Learning Using Chemland Simulation Software. *Chemistry Educator, 5*, pp 32-38.

Khan, S. (2001). Developing inquiry skills while learning about unobservable processes in chemistry. Presented at the annual meeting of the American Educational Research Association, Seattle,WA. April, 2001.

Khan, S. (2001). A guided discovery approach to developing inquiry skills using multiple, compact simulations. Presented at the annual meeting of the National Association for Research in Science Teaching, St. Louis, MO, March 25-28, 2001.

Stillings, N. (2001). From the visual to the mental model and back: A cognitive perspective on molecular visualization. Invited address at Science Education and Visualization: International. A Gordon Research Conference. August 5-10, 2001, Mt. Holyoke College, South Hadley, MA

Stillings, N., Khan, S., & Clement, J. (2002, April). Guiding hypothesis construction & evaluation in college chemistry. Presented at the Annual Meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Khan, S., Stillings, N., Tronsky, L., Wenk, L, & Izumi, A. (2002, April). The integration of multiple,compact simulations to achieve process and content goals in an introductory chemistry course. Presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Khan, S., Stillings, N., Tronsky, L., Wenk, L, & Izumi, A. (2002, April). Reasoning processes of introductory chemistry students using multiple compact simulations. Presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Khan, S., Stillings, N., Clement, J. & Tronsky, L. (2002, April). The impact of an instructional strategy using computer simulations on inquiry skills in chemistry. Presented at the Annual Meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Khan, S. A. (2002). Teaching chemistry using guided discovery and an interactive computer tool. Doctoral dissertation, School of Education, University of Massachusetts, Amherst

Khan, S. (2003). Model construction processes in a chemistry class. In *Proceedings of the Annual Meeting of the National Association for Research in Science Teaching (NARST),* Philadelphia, PA.

Khan, S. (2003, May). *Chemland computer software in the chemistry classroom.* Presented at the Catalyst Conference BC Science Teachers Federation, Richmond, BC.

Khan,S. (2003, May). *Teaching chemistry with Chemland, an interactive computer tool*. Presented at Canadian Society for the Study of Education. Halifax: NS
Khan, S. (2004). Student construction of visualizable models in chemistry: Teacher guidance strategies. In *Proceedings of the 18th International Conference on Chemical Education (ICCE),* Istanbul, Turkey.

In her classroom laboratory study Samia Khan investigated both student reasoning processes and instructor scaffolding. She found that interactive software modules effectively motivated student exploration and supported a type of inquiry cycle that John Clement had earlier dubbed the GEM cycle, for *generate, evaluate, modify*. In a GEM cycle students propose a relationship among variables and then try to further evaluate it by collecting more information within the simulation. Often their hypothesis is not confirmed, and they have to modify their hypothesized relationship. As the cycle proceeds, they try to understand underlying causal relationships as well as the law-like relationship among variables.

Khan also systematized a set of instructor interventions that Prof. Bill Vining had developed spontaneously over the years. Rather than giving students answers, Vining relies on a range of open-ended prompts to move students forward in their reasoning. Examples are, suggesting that students check extreme values of a variable, asking them what their prediction is right now, asking them to think about whether their data is actually consistent with the hypothesis they are entertaining, and so on. Vining has demonstrated that these scaffolding techniques can be used in a way that encourages student exploration rather than shutting it down.

Khan presented some evidence that students' model-oriented reasoning skills improved during the term. Although overall score increases were not consistent, some items on our general scientific reasoning inventory showed gains. In addition, there was evidence from course examinations of improvements in reasoning.

Synergistic findings
The partnerships that have arisen through our intervention projects have proven to be unexpectedly rich and sustainable. The decision to begin by studying naturally-occurring reforms using independent methods was an effective way of establishing research-based relationships with faculty who have a strong commitment to student learning and a willingness to face its complexities. We have come to favor sustained partnerships between researchers and practitioners in sustainable instructional settings rather than a simple sequence of research in an artificial environment followed by "adoption" in ordinary environments.

A common thread that has emerged in all of our work is the importance of helping students understand the role of theory or model-based causal explanation in modern science. We believe that even within many reform settings beginning students are given too little active experience with models that feature reasonably complex underlying causal processes. Our NOS interview looks carefully at students' understanding of the theory-driven nature of scientific explanation, and in each of our intervention projects we have worked with faculty members to help students think actively with causal explanations.

Opportunities for Training & Development

This project has been a significant site for training and development. The number of people involved with the project has been significant, given its location at a small college. Perhaps more importantly, the nature of the training may be unique. Post-doctoral fellows and graduate students were given significant responsibility for the diverse research projects carried out under the grant; they observed and worked closely with dedicated, experienced, inquiry-oriented faculty members from Hampshire and the University of Massachusetts; and they had the opportunity, using Hampshire's matching funds, to practice inquiry-oriented teaching at Hampshire, often in collaboration with veteran Hampshire faculty members. Three graduate students completed doctoral dissertations on the project. One graduate student (Samia Khan) and two post-docs (Laura Wenk and Loel Tronsky) have made the transition to regular faculty positions (at the University of British Columbia, Hampshire College, and Albertus Magnus College respectively).

The focus on intervention led to a large number of faculty participants (20) with significant involvement in the research. The faculty participants came to the project having already initiated reforms in undergraduate teaching, making their work on the project truly collaborative. Several of them have essentially become co-researchers on the project and are planning future proposals.

Post-doctoral fellows: The grant turned out to be a more significant site for post-doctoral training than anticipated, significantly involving two post-docs:

Laura Wenk: Wenk came onto the project as a graduate student, completing her thesis during the first year of the grant. In the succeeding years, she was appointed as a post-doc. Wenk and co-PI Carol Smith have jointly conducted the epistemological development project. The NOS interview developed for the project incorporates aspects of Wenk's doctoral thesis. Wenk was also a primary researcher on the scientific reasoning project and the primary literature project. She has become a sought-after consultant for other educational reform projects, e.g. serving as an evaluator on NSF CCLI, CRUI, and REU grants and offering workshops on student-active learning at several AAHE workshops, at Mitchell College, and at Winona State Univesity. Wenk (whose appointment as a post-doc on this grant was less than full time) also taught part-time at Hampshire during the grant, developing innovative courses on science education, educational research, and inquiry-oriented instruction. In September 2003 Wenk became a regular faculty member in the School of Cognitive Science at Hampshire, having been hired in a national search during 2002-03.

Loel Tronsky: Tronsky joined the project after receiving his Ph.D. in psychology from the University of Massachusetts, Amherst. Tronksy's graduate training was in the cognitive psychology of reading and mathematics performance, with thesis work involving laboratory studies of cognitive processes in elementary mathematics. On this project Tronsky expanded his methodological expertise considerably, doing structured classroom observation, conducting NOS interviews, and co-directing the primary literature project. He was the chief statistician on the entire project. Tronsky also worked with PI Neil Stillings to develop a new course in cognitive psychology at Hampshire and developed an innovative course of his own on cognition and education. In 2003-04 Tronsky served as the evaluator for a curriculum development project in the biological sciences at Hampshire, funded by the Howard Hughes Medical Institute. In Fall 2004 Tronsky became a tenure-tract assistant professor in psychology and education at Albertus Magnus College in New Haven CT.

Graduate students: Laura Wenk completed her doctoral dissertation during the first year of the project, before becoming a post-doctoral fellow. Her dissertation research was conceived before the initiation of this grant but her work became an integral part of the project.. Two graduate students did their doctoral thesis research within the project and were fully supported by project funds:

Samia Khan: Khan was a graduate student in education at the University of Massachusetts. Her work was jointly supervised by Stillings and Prof. John Clement of UMass, who was also a consultant to this project. The chemistry project was Khan's thesis research, which was completed in Summer 2002. During the project, Khan mastered a number of methodological skills, including structured classroom observation, scoring essays for scientific reasoning, analyzing transcripts of peer and student-faculty interaction, and designing and analyzing student attitude surveys. With Stillings she co-taught an undergraduate seminar at Hampshire on cognition and education. Since September 2002 Khan has been a tenure-track assistant professor of education at the University of British Columbia.

Alisa Izumi: Izumi was a graduate student in education at the University of Massachusetts. She was a primary researcher on the model-based reasoning in biology project, observing classes, scoring student essays, and generally maintaining liaison with the UMass Biology Department. Her thesis concerned the influence of format differences on student assessment in the biology project. Her work defined a set of challenges for assessment in very large reasoning-oriented courses (200-400 students) that we hope to address in future work.

Several other graduate students at the University of Massachusetts have worked for briefer periods on the project or have interacted extensively with the project research team, including Maria Nunez, Nancy Karp, Mary Jane Else, and Cynthia Stein.

Undergraduate students: A number of undergraduates participated significantly in the project, by attending lab meetings, training as classroom observers, and working on data analysis. They included Erica Hayes, Sarah Urban, Jun Fukukura (Smith College), Janelle Butler (UMass), Jonah Charney-Sirott, and Lionel Claris. Many other students have worked on transcribing the audiotapes of the NOS interviews (identities of interviewees were blind to transcribers). Although transcription has limited educational value, these students did become familiar with one of our projects and with procedures for accurate transcription, managing transcript data, and complying with IRB regulations. Undergraduate students were supported largely from Hampshire College funds.

Faculty participants: The faculty participants in the project all had initiated reforms in higher education prior to our contact with them. Project work followed a rhythm of researching their teaching methods and student outcomes and then working collaboratively with them to develop refinements in their instruction and assessment practices. This approach produced exceptionally vibrant collaborations that will be sustained after the end of this project. Several of the faculty members have made major contributions to the research designs and have essentially become co-researchers. The main faculty participants were as follows:

Hampshire College: Eleven faculty members in Hampshire's School of Natural Science participated in scientific reasoning or primary literature projects: Dula Amarasiriwardena (chemistry), Merle Buno (biology), Elizabeth Conlisk (health science/epidemiology), Charlene D'Avanzo (ecology), Christopher Jarvis (molecular biology), Nancy Lowry (organic chemistry), Ann McNeal(physiology), Lynn Miller (genetics), Steven Roof (geology), Brian Schultz (entomology/ecology), and Fred Wirth (physics). Bruno, D'Avanzo, and McNeal worked with the research team to design the assessments and instructional interventions for the primary literature project.

Department of Biology, University of Massachusetts: Professors Randall Phillis and Steven Goodwin were the instructors for the model-based biology course. They originated the reform of the course under a Pew Center for Academic Transformation grant, emphasizing deeper concept mastery and peer learning. They collaborated with the research team on developing the model-based reasoning concept and the associated measures of student outcome. They developed the instructional materials for the model-based approach to instruction in a large lecture hall. Professor Steven Brewer, the technology director for the biology department, developed the web component of the course and worked with the research team to use the web materials to support the new emphasis on models. Professor Lawrence Schwartz taught a conventional large-lecture section of introductory biology in the Fall of 2000, which was used as a comparison group for an early version of the model-based course. Professor Elizabeth Connor contributed to research planning and to data interpretation.

Department of Chemistry, University of Massachusetts: Professor William Vining developed the ChemLand software and the techniques for employing the software in moderate-sized electronic classrooms (30-40 students). He worked with the research team to design the lab studies of the instructional technique, on student assessments, and on training other faculty members in the ChemLand approach. Professors Justin Ferman and David Adams taught ChemLand-oriented sections of introductory chemistry, in which we made classroom observations and administered student assessments. Prof. Bea Botch taught a large-lecture section of introductory chemistry, which we observed and in which we did some student assessment.

<u>Co-PI Mary Anne Ramirez</u>: Although co-PI's are not normally considered trainees, it is important to note that this was Ramirez's first grant as a co-PI. The grant marked several significant advances in her career: (1) She managed the model-based reasoning in biology project, described above; (2) She secured another grant as a co-PI, based on her thesis research (ESI-9911401); (3) She worked as a visiting assistant professor at Hampshire, teaching several courses in the cognitive science and education studies programs.

Outreach Activities

In a sense the entire project is an outreach effort, since a major goal was to bring cognitive/educational research questions and methods to the undergraduate science education reform community. The project led to significant outreach well beyond the faculty groups immediately involved in the research, however. These outreach activities fall into the following categories:

<u>Organization of conferences on undergraduate education</u>: Neil Stillings has worked to disseminate the partnership approach developed on this project more widely in the undergraduate education community. The core of our partnerships is the collaboration between faculty members who are committed to improving teaching in their disciplines and researchers in education and cognition. Stillings worked with leaders in two fields to initiate national-level partnerships between the disciplinary and cognition/education research communities.

In the first case he worked with Loretta Jones of the University of Northern Colorado and Ken Jordan of the University of Pittsburgh to organize the 2001 national workshop on molecular visualization in science education, sponsored by the NSF (REC-0090523), and attended by roughly 30 national leaders in chemistry, chemistry education, and the cognitive and learning sciences. The workshop was coordinated with the 2001 and 2003 Gordon International Research Conferences on visualization, with a mini-grant program open to Gordon participants, and with a follow-up workshop with NSF program officers in December 2002. The final report of the workshop is available at http://pro3.chem.pitt.edu/workshop/index.html.

In the second case he worked with Cathy Manduca of Carleton College and David Mogk of Montana State University to organize the 2002 national workshop titled Bringing Research on Learning to the Geosciences, co-sponsored by the NSF (REC-0213065) and the Johnson Foundation, and attended by roughly 25 leaders in the geosciences, geoscience education, and the cognitive and learning sciences. The website for this ongoing effort is at http://dlesecommunity.carleton.edu/research_on_learning/workshop02/index.html.

<u>Planning and assessment work on other projects</u>: The approach, methodologies, and expertise of the project research group has led to a number of invitations to help other groups design and assess educational reform projects. During the period of the project Stillings was co-PI on a CCLI project to develop simulation software for ecology (CCLI-9972486). On this project he designed an adaptation of the scientific reasoning assessments for middle school students. Two other mature examples are the Mind Project (CCLI-EMD-9981217) on cognitive science learning, headed by David Anderson, and based at Illinois State and Indiana Universities, and the geoscience curriculum reform effort at Mount Holyoke College (CCLI-A&I-9952822). Stillings will also be overseeing evaluation on Al Werner and Steven Roof's REU project on climate change in the high arctic (NSF #0244097). Laura Wenk of the project staff is handling evaluation for Alan Goodman's CRUI grant in physical anthropology (NSF #9978793), and Debra Martin's SES-REU grant (#0354037). Co-PI Mary Anne Ramirez worked with the teacher education program at Springfield (MA) Technical Community College on the assessment of its students' scientific reasoning abilities.

Consulting relationships and other contacts with other projects and groups: Members of the project research staff have consulted formally or informally with many other research and development projects in science education reform. Stillings, for example, was on the national advisory boards for Physiology Education Research Consortium's project on active learning (REC-9909411) and Spelke, Kanwisher, Hauser, and Carey's project on the sources of mathematical thinking (REC-0087721). He has also consulted with the FAS's Learning Federation on instructional design for technology-enabled approaches to learning and with the Concord Consortium on several proposals. Stillings and Wenk have given frequent presentations to the Five-College STEMTEC teacher education collaborative (NSF #9653966), based in Amherst, MA and headed by Prof. Morton Sternheim of UMass, and have advised many of the STEMTEC faculty members. Stillings and Wenk worked with faculty participant and STEMTEC co-PI Charlene D'Avanzo to organize a symposium at STEMTEC's 2002 national conference Pathways to Change, held in Washington, DC.

Faculty participants in the project are also very active on the national undergraduate education scene, appearing frequently at science-teaching conferences and consulting widely with colleagues around the country. They have incorporated their work on this project into their participation in broader educational communities in a variety of ways.

Pre-college school outreach: Stillings and Wenk initiated a project in the Springfield, MA school system to encourage inquiry-oriented science teaching at the 9th-grade level and to adapt the epistemological and reasoning assessments of the current project to this inner-city setting. That project received funding from the Department of Energy, which has allowed the hiring of a post-doctoral researcher and a science education coordinator/trainer.