Inquiry-Based Learning:
Cognitive Measures & Systems Support

Funded by the Learning and Intelligent Systems Program
National Science Foundation
NSF Grant REC-9720363

Neil Stillings, PI


This section reports the main activities and findings of the entire project, including a no-cost extension that ended on August 31, 2001. The project had an educational research component and a software development component. Each component included self-contained activities and goals, as well as elements that interacted with and informed the other component.

Summary

All of the proposed research was completed and the goals of the project were met and in most respects exceeded. The results led to three further projects currently in progress, one funded by NSF-CCLI, a second by NSF-REPP, and the third by the FIPSE comprehensive program.

• Educational Research
  - We designed and carried out a study of inquiry-oriented science education at the college level, investigating faculty beliefs, classroom activities, and student outcomes.
  - We laid the organizational and methodological foundations for extending this research to other institutions. REPP has funded a larger comparative study.

• Software Research and Development
  - We identified aspects of college-level inquiry-oriented instruction that could potentially be supported by software, and we developed approaches to extending existing software technologies to realize this potential.
  - We developed three software prototypes to demonstrate the feasibility of the approaches and to begin to test their effectiveness in the classroom.
  - The prototypes are currently in classroom use locally. Further development, testing, and dissemination of one program, FOREST, is being supported by a CCLI grant. A second program, Geology Observer, is being developed further, evaluated, and disseminated under a grant from the FIPSE comprehensive program. The third program, CHAT, will be distributed by the MIT Press after further development.

Although the work on the two components of the project proceeded interactively, it is convenient to describe them in detail separately below, with comments on their interaction.

Educational Research
Our educational research treated introductory science instruction at Hampshire College as a naturally-occurring environment, or living laboratory, for the study of inquiry-oriented teaching and learning at the college level. Our main research goals were to characterize the goals and practices of science instruction at Hampshire and to use them as benchmarks in the development and testing of new educational software.

The educational research was directed by PI Neil Stillings and Mary Anne Ramirez, who joined the project in June 1998 as a post-doctoral researcher. Professor John Clement of the University of Massachusetts served as a consultant on this research. Two of Clement’s graduate students, Deborah Maclin and Samia Khan participated. Yaniris Fernandez, another education graduate student, who is also on the Hampshire staff, participated as well in fulfillment of her research internship requirement at the University.

**Study of Faculty Conceptions of Inquiry-Oriented Instruction: Description & Methods**

Since Hampshire’s founding in 1970, beginning science instruction has been guided by the core principles of introducing students to the nature of scientific inquiry and actively involving them in research projects. Individually and collectively, the college’s science faculty has received substantial funding for curriculum development from the NSF, NIH, Howard Hughes Medical Institute (HHMI), and other foundations. Methods and subsidiary goals have evolved and matured over time. In this first phase of our educational research we attempted to characterize the Hampshire science faculty’s conception of their instructional goals and practices.

Over a period of several months in the Spring of 1998 Stillings and Fernandez conducted a series of structured interviews with a sample of faculty spanning the fields of biology, ecology and environmental science, geology, chemistry, and physics. The open-ended questions covered both educational goals and classroom practice. Interviewees were asked to describe specific courses and assignments that they felt best illustrated their work as well as the challenges of inquiry-oriented and student-active instruction. The questionnaire also included items concerning the use of computers and software in instruction.

Audiotapes of the interviews were transcribed by Fernandez and subjected to an initial content analysis by Ramirez during the spring and summer of 1998. For the content analysis, all propositions concerning goals or methods of instruction were extracted and logged individually into a computerized data file. The data were used to distill a set of categories that organized the propositions into a small sets of overarching goals and methods for inquiry-oriented instruction.

**Study of Faculty Conceptions of Inquiry-Oriented Instruction: Results**

The results of the study were fairly complex and are summarized in some detail in papers presented at the 1999 American Educational Research Association (AERA) and National Association of Research in Science Teaching (NARST) conventions (the latter paper is available at http://helios.hampshire.edu/lspector/NSF-LIS/ed-research.html). Briefly, we found that inquiry-oriented faculty members emphasize both the acquisition of certain higher-order
cognitive skills and developmental changes in conceptions of the nature of science (scientific epistemology). These instructional goals rest on a complex network of often implicit hypotheses about learning and cognition. Three central hypotheses are

1. Students can acquire higher-order scientific reasoning skills and epistemologies that transcend specific disciplinary content.
2. The acquisition of higher-order reasoning skills and epistemological change support each other and must occur jointly.
3. Involvement in authentic scientific inquiry leads to both epistemological change and the acquisition of higher-order reasoning skills.

These hypotheses, along with the underlying detail distilled from the interviews, guided the further development of the educational research on this project (and on a subsequent project, currently supported by REPP). We designed a classroom observation instrument to assess empirically the instructional methods that were being used in support of reasoning skills and epistemological change, and we developed an essay-style assessment of general scientific reasoning skills that could be administered to beginning college students in any field in pre-post designs. These instruments were put into use in the fall semester of 1998.

Classroom Observation Research: Description & Methods

In the fall of 1998 we began research to characterize how the faculty conception of inquiry-oriented instruction is actually implemented in the classroom. Four introductory science classes at Hampshire were selected for observation throughout the term: Evolution of the Earth, a geology course oriented toward field research projects; Musical Acoustics, a topic-oriented physics course oriented toward acoustic measurement and fabrication projects; Human Biology, a biology course that employs medical-diagnosis case studies; and Human Evolution, a biology course that emphasizes readings in the primary literature on human evolution. Observation of the Human Biology course involved collaboration with the NSF’s STEMTEC grant to the University of Massachusetts and Five Colleges, Inc.

An observer from our research team was present in each course for at least half of the class meetings. Observers coded each class meeting in terms of a goals-by-methods matrix that was developed from the faculty interview research. The reliability of the coding was assessed by having two independent observers code a number of classes. Observers also took anecdotal notes on each class meeting.

Classroom Observation Research: Results

The results of the observation research were used within this project primarily to inform the software development described below. We inventoried the instructional methods that were actually being used by faculty and identified methods that could potentially be significantly enhanced by educational technology. A general qualitative presentation of the results was made at the Spring 2000 AERA convention. We reported three main findings. First, inquiry-oriented teachers devote significant amounts of time to classroom activities that actually are oriented
toward inquiry objectives. Second, there is more variability in the classroom practices of inquiry-instructors than there is in their learning objectives for students. That is, there is uniformity in objectives and diversity in practices. Third, inquiry-oriented instructors inject inquiry goals into traditional methods of instruction (e.g. lectures), making it important to distinguish between instructional methods and goals. Further work with this instrument is currently being supported by our REPP project.

Student Assessment Research: Description & Methods

In 1998 we began a project to study cognitive change and learning outcomes in students. The learning goals derived from the faculty interviews were used to design a series of open-ended questions that probe students’ ability to think scientifically about novel scientific research scenarios. The questions ask students to generate hypotheses and research designs, interpret and critique empirical data, and reason about hypothetical scientific controversies. They are specifically designed to emphasize general scientific reasoning skills rather than field-specific knowledge. The questionnaire and scoring protocol were pilot-tested and revised during the spring and summer of 1998.

In 1998-99 the reasoning instrument was used in a pre-post study of first-semester college students, comparing four groups of students: (1) Hampshire students enrolled in science courses; (2) Hampshire students not enrolled in science courses; (3) Students in an innovative interdisciplinary general science course at a comparison college; (4) Students in a traditional introductory biology course at the comparison college. The two Hampshire groups were comparable, because students' needs to fulfill the college's first-year distribution requirement produces a largely arbitrary pattern of enrollment or non-enrollment in science courses during their first and second terms. The Hampshire groups and the comparison college groups were well matched in general academic ability, although other possible differences in background or personality were not assessed.

In a collaborative project, Laura Wenk, a doctoral candidate in education at UMass, interviewed students pre and post-semester in groups 1, 3, and 4 above, using an epistemological assessment instrument that she developed. Wenk's scheme of epistemological development was based on the literature concerning college students' epistemology rather than on our faculty interviews. Wenk's interview sample was not a subset of the scientific reasoning sample above. The number of students who participated in both studies was too small to support any conclusions concerning direct relationships between gains in scientific reasoning and epistemology.

Student Assessment Research: Results

We found a significant positive changes in scientific reasoning skills in the students enrolled in inquiry-oriented science courses (Group 1) and no change in the other three groups, lending support to hypotheses 1 and 3 above. The paper reporting these results in available at http://helios.hampshire.edu/lspector/NSF-LIS/ed-research.html.
Our collaborator, Laura Wenk, found significant increases in epistemological sophistication in Group 1 and no changes in groups 3 and 4 above, also lending support to hypotheses 1 and 3 above.

The increases in both reasoning skills and epistemology in samples from common groups of students supports hypothesis 2 above, in the sense that the hypothesis predicts that gains in the two areas should be correlated. The support is weak, however, because the correlation could not be assessed within students.

Software Development Research

The project interleaved classroom research and software development, each informing the other. A major goal of the project was to produce three mature prototypes of inquiry-oriented software applications for beginning college-level instruction. This goal was achieved. It was also exceeded in two respects. First, all of the prototypes received more extensive classroom testing than was proposed under the final budget for the project. Second, with new sources of support, two of the projects are receiving extensive further development and are on track for widespread distribution. Funds have been requested to support further development of the third project as well.

Four software development projects were initiated in the Spring of 1998 with the intention of selecting three for further development under the terms of the grant. The three projects selected were CHAT, FOREST, and Geology Observer. The fourth project, Sound, Music, and Mind, was eliminated because cross-platform development of an interactive sound-intensive application was too expensive for the project budget. The selected projects will be described separately below followed by some more general remarks.

Linguistics: CHAT Project

The CHAT project reflects the team’s interest in developing an application in the cognitive sciences. Co-PI Steve Weisler completed an interactive multimedia CD-ROM-based introduction to linguistics for the MIT Press in Fall 1998. One origin of our project is that Weisler found that he had neither the technology nor the development resources to include true inquiry-based interactivity in his CD-ROM. CHAT is an inquiry-based environment within which students can explore theories of the syntax of natural languages. It includes significant resources for internet-based student interaction, and it implements the agent-based architecture for scaffolding student inquiry that we outlined in our proposal. This aspect of the project was directed by co-PI Lee Spector. A mature Java-based prototype of the program has been completed and tested in the classroom. With some further development, for which we plan to seek funds, it will be a stand-alone product suitable for national distribution. The MIT Press has indicated its willingness to publish the final version.

In syntactic research linguists propose sets of rules (grammars) that are intended to generate the (infinite set of) sentences of the language under investigation. A grammar is thus a scientific
hypothesis (or model) that is tested by assessing the degree to which it generates all and only sentences of the target language. CHAT is designed to make it easy for students to propose and test grammars and to share their results with other students and the instructor. Professional syntactic research on particular languages takes place within a theoretical framework. CHAT is based on the current Chomskyan minimalist theory of universal grammar.

CHAT provides a rule editor in which students propose and change syntactic categories and rules and a lexical editor in which they specify or edit the syntactic properties of words that they add to their dictionaries. It includes a generator, which, at the user's command, generates random sentences using the current grammar and lexicon. The program will also display the phrase-structure tree diagram for any sentence it has generated, which allows students to see the structural consequences of the sets of rules they have proposed. The rule editor allows the student to mark rules as disabled or preferred during sentence generation, to facilitate testing hypotheses about the effects of particular rules. Grammars and lexicons can be saved and loaded, allowing students to save their work and continue later at another machine or to begin their work with a basic grammar provided by the instructor. The program is thus a powerful, flexible, interactive model builder and simulator for natural-language syntactic theory. It should perhaps be noted here that the implementation of these features was a significant linguistic research and software development project. Natural-language syntax engines are generally the province of professional researchers who are intimately familiar with the underlying programs. Finding a way to express a current theory of universal grammar in a modeling environment that is accessible to the average undergraduate was a major research accomplishment.

CHAT also includes features for student-student and student-instructor interaction. Within the program users can log onto a server, allowing them to communicate via the internet. All connected users can see the Chat window. Users can post sentences that they have generated to the Chat window, and they can inspect the phrase-structure trees for sentences posted by other users, allowing them to compare their grammars to the grammars of other users. Logged on as a user, the instructor can inspect sentences posted by the students and comment on them or post sentences that scaffold student work.

Another innovative feature of CHAT is a collection of agents that observe the student's work and issue reports and suggestions that the student might want to act on. The agents are based on a set of considerations outlined in our proposal, where we argued that in an open-ended inquiry situation a program would not be able to scaffold students' work by reference to a known solution. Thus, intelligent software-based interactivity would depend on critiquing characteristics of the student's inquiry activities or general characteristics of their partial solutions. We proposed an architecture for inquiry tutors in which students conduct inquiry in a simulation environment. In the architecture monitoring agents observe the student's use of the simulator's tools and the resulting output and post observations to a blackboard. Reporting agents take input from the blackboard and possibly from databases that are external to the simulation and issue reports and suggestions to the student. The architecture is flexible in the sense that the agents are independent of the simulation engine and of each other and thus can be added incrementally to a system.
During the project we implemented a collection of agents that illustrate the range of complexity that we believe is useful and feasible in this type of architecture. Examples of simple agents are those that notice and report to the student that a certain syntactic rule is rarely or never used in generation, or that a previously generated sentence can no longer be generated. An example complex pair of agents automatically sends randomly-generated sentences over the internet to Temperley, Sleator, and Lafferty's natural-language parser at Carnegie-Mellon and reports parsing failures back to the student.

The development of CHAT extended from the spring of 1998 through the early fall of 2000. We followed a rapid-prototyping, incremental approach to design and development in which versions of the design were critiqued by students and the project research team and tested in the classroom several times. The development process proved to be a very successful site for undergraduate training. An entire semester-long course was devoted to the early design and prototyping. Three undergraduate assistants worked on the second prototype, in Macromedia Director. One of these students, Roger Bellin, eventually became the lead programmer on the project, finishing the Director version and then porting it to Java, which allowed better data representations, a more modular design, and a clean implementation of the agent architecture.

Stable, mature versions CHAT were tested in the classroom three times. We found that the program is very successful in an electronic classroom in which students are seated individually or in pairs at computers with an instructor present. Students seated individually frequently interacted with their neighbors. Instruction is facilitated if the instructor's display including the Chat window is projected onto a large screen visible to the entire class. In this setting, students work on inquiry problems posed by the instructor, who takes questions from the class and monitors their activity. As students exhaust each inquiry problem, the instructor displays and comments on students' alternative solutions, summarizes, and sets a new inquiry problem. Students moved through two to four inquiry cycles in an eighty minute class meeting. Student morale and engagement in these classroom tests was extraordinarily high, relative to standard college classrooms and particularly to syntax classes, which most student find dull. Videotapes of four pairs of students in the first assessment showed that they were using all features of the program and were engaging in inquiry cycles, in which they made changes to their grammars and lexicons and tested them by using the generator and by inspecting alternative approaches being pursued by other students. In a small-scale classroom research project in the Fall of 2000 we confirmed through student questionnaires and inspection of students' grammars that engagement was very high, that students actively sought solutions to the problems that were posed, and that they mastered discipline-specific material that underlay the inquiry assignments. Although there was no control group for this study, experienced instructors felt that the level of student mastery was much higher than what is typical in introductory linguistics courses.

Significant institutional funds outside of the grant were expended to complete and stabilize the Java version of CHAT and to complete the Fall 2000 classroom test. We are seeking funds to prepare an instructor's manual and polish the interface of the program, prior to national distribution, which we will undertake jointly with the MIT Press. We may also seek funds for further development and evaluation of the agent architecture.
CHAT and further information about it can be downloaded at http://helios.hampshire.edu/lspector/chat/.

**Biology & Environmental Science: FOREST**

The FOREST project was led by Co-PIs Tom Murray and Larry Winship. It represents a second expression of the ideas about simulation software that were presented in our proposal and the fulfillment of our proposal to develop one application in the biological sciences. A mature Director-based prototype of FOREST was completed with grant funds and underwent several preliminary classroom tests. Co-PIs Murray, Winship, and Stillings secured further funding from NSF-CCLI (Grant No. 9972486) to develop a Java-based port of the program, to develop a curriculum to support its use, and to disseminate the software and curriculum at the college and middle school levels through summer faculty development workshops with classroom follow-up. That project is now in its second year, having overlapped with the extension of the LIS project.

FOREST addresses two issues for college-level inquiry-oriented ecology and environmental science courses that we identified in our early planning work. First, although courses in this field offer abundant opportunities for students to do significant field work, appreciating the context of the field observations often requires an understanding of complex models and of processes that extend over long periods of time. Simulation programs offer students a concrete introduction to complex dynamic models and to the unfolding of the modeled processes over arbitrary time periods. We chose to build a forest-growth simulator in part because Winship had experience with research-level simulators and in part because the simulations model complex phenomena in terms of parameters that are easy for beginning students to understand and measure. At Hampshire students can collect measurements on a woodlot a few hundred yards from the science building. When the data are fed to the simulator they immediately raise interesting questions about the past and future and about the implications of the simulation model itself. Some of the questions can be explored by making further field measurement at other local sites that involve different conditions and younger or older (including "old growth") trees.

The second issue addressed by FOREST emerged fully during our planning and design work and therefore was not featured in the proposal. The agent architecture described in our proposal and implemented in CHAT addresses the issue of scaffolding open-ended inquiry within a simulation, but it does not address another concern expressed by faculty members, which is the opacity, or black box nature, of student-oriented simulations. This was not an issue in CHAT because the "equations" in the simulation are syntactic rules that students write themselves. In that sense, CHAT is not a black box. FOREST was potentially a black box because, like many simulations, it could be written in such a way that students could see the results of their parameter settings but could not see the equations that the parameters are fed into. Professional scientists are typically explicitly aware of the model underlying a simulation, and research involves tinkering with both the parametric consequences of the model and the model (set of equations) itself. We therefore set out to make FOREST a "glass box" that would allow students to inspect and alter its underlying model. We believe that implementing this kind of transparency in a way that is accessible to beginning students without programming skills and with average backgrounds in high school mathematics will be a major development in educational software. It
is also a very stiff challenge that could not be met fully within the project, and that was part of our motivation for seeking the CCLI funding.

Like CHAT, FOREST was developed incrementally in Director. Successive versions tested in the classroom during four successive terms beginning in Spring 1999, with the most recent classroom tests in Fall 2000 funded by the CCLI grant. FOREST's user interface is fairly complex, and changes were made in every version in response to student comments and to classroom observations by the research team. Overall, students enjoy the program and show a very high level of engagement with running field data backward and forward in time under varying assumptions and with trying to create forests with varying characteristics.

On several occasions we videotaped pairs of students using FOREST, and two student interns coded the tapes for inquiry cycles, in which students posed, attempted to test, and revised hypotheses. These analyses confirmed that students were using the software for inquiry, and that the software promoted inquiry in the sense that many more inquiry cycles occurred in a class meeting than occurred in classes with traditional discussions of readings and field data. As with CHAT, we found that systematic inquiry using a simulation must be scaffolded. College students, like younger students, often fail to apply good principles of research design in examining their hypotheses. In an electronic classroom setting, this scaffolding was supplied spontaneously by the instructor, Winship, who is an experienced inquiry-oriented teacher. In the CCLI project we are studying the needed supports further, developing curriculum and training faculty to promote inquiry. The current plan is leave FOREST a relatively pure simulator, with the glass box feature and to embed the supports for its use in the surrounding curriculum and learning environment. One reason for this decision was our belief that the simulator would be useful at several different levels of instruction for several different purposes, each of which would have different curricula. That belief has been born out rather spectacularly. Although the project focus was inquiry-oriented science at the upper high school and beginning college levels, FOREST has drawn interest from faculty members who teach advanced courses in forestry, and it is now in use in five Western Massachusetts middle schools at the eighth grade level.

Current information about the Forest Simulation project and software downloads are available at http://ddc.hampshire.edu/simforest/index.html.

**Geology: From Observation to Hypothesis**

The Geology Observer (sometimes known as Geo Observer or Dr. Geo) program is the purest expression of the project's commitment to base its software development on the experiences and insights of inquiry-oriented college faculty members. It implements themes from the field-based introductory course Evolution of the Earth, developed by Professor John Reid at Hampshire. Although not a PI on the project, Reid became a full member of the development team and worked closely with PI Beverly Woolf on the development of Geo Observer. We completed and classroom tested a prototype of Geo Observer that contains one fully developed inquiry unit. We have received funding from the FIPSE Comprehensive Program to add several more inquiry units, which must be added before the software can be fully tested and widely disseminated.
One of the goals of Evolution of the Earth is helping students to learn to make geologically relevant observations about natural scenes and to generate initial hypotheses and research ideas from their observations. Reid developed a classroom exercise in which students are shown a high-quality photograph of a natural scene and initially asked to make relatively uninterpreted, or atheoretical, observations about the scene. For example, one picture is of a bend in the Tuolumne River, flowing through a meadow in the Sierra Mountains. Beginning students generate lists of observations that include features such as, "there’s gravel on the right bank and grass on the left bank," or, "it looks like there is more turbulence in the water on the left than on the right." Reid refers to the training and exercise of observational curiosity as "Page 1."

Page 1 observations lead to questions that require further measurement and to initial hypotheses. In the Tuolumne example, students might decide that they wish to know the depth of the river at various points or its rate of flow. This transition to further observation and measurement and to initial, relatively shallow hypotheses is referred to as Page 2. Page 3 involves a further transition to more comprehensive hypotheses that invoke causal mechanisms or theoretical frameworks. For example, it might be hypothesized that the location of the bend in the river has moved in a particular direction over the years.

Geo Observer underwent incremental development, again using Director, throughout the project, beginning with a design phase in Spring 1998 and semester-long observations of Reid's course in Fall 1998. Our initial conclusion was that a relatively simple software application could enhance the classroom experience considerably. Software could allow students to zoom in on parts of the image, to inspect multiple views of the same scene, and to tie their observation to pointers to particular spots in a scene. These facilities would encourage each student to independently generate multiple observations prior to group discussion. As students moved toward hypotheses, they could explore further information about the scene stored in a database. In the Tuolumne River scene, for example, data about the rate of flow and depth of the river at various locations could be made available to the student at an appropriate point. Students could be asked, for example, to make predictions about these measurements prior to accessing them.

In the final prototype the movement from observation to hypothesis is tied together by a sequence of notebooks, in which students first enter their observations and are then asked to make preliminary hypotheses and predictions about certain aspects of the scene, such as the depth profile of the river scene. They are finally led to develop relatively complete models of some aspect of the scene, e.g. a qualitative explanation of why the location of the bend in the river may have moved over a long period of time.

Recent classroom observations have confirmed that the software succeeds in getting all students to participate in the observation and model building process and that the notebooks help connect the skills of observation and explanation by forcing students to explicitly record their hypotheses and predictions before accessing data that is not immediately observable. In principle, it would be possible to accomplish most of this functionality with a carefully-orchestrated set of printed handouts, but the interplay in the software between pictorial information, field data, and the notebooks is more efficient and compelling.
The FIPSE grant, awarded in August 2001 to LIS co-PI Beverly Woolf and participant John Reid, will support the development of several more inquiry exercises, comparable to the river exercise, that illustrate other major geological phenomena, such as glaciers and volcanoes. With these additions the program will be a suitable accompaniment to a range of introductory courses in geology and environmental science or engineering.

General points about the software research

Each of the software projects satisfies three objectives that were central to our research program.

1. The software builds on, and attempts to improve, an existing, successful inquiry-oriented innovation in the college classroom.

2. The software explores an issue in educational technology that has significance beyond the discipline represented. CHAT explores the inquiry-agent architecture; FOREST explores the glass-box metaphor for simulations; Geology Observer explores the potential for software to scaffold students' observations and their moves from observation to hypothesis.

3. The software can be flexibly incorporated into a range of courses for a range of in-class or out-of-class uses. For example, CHAT can be used for a 2-4 week introductory syntax unit in a general linguistics course or throughout the term in a syntax class, supporting the exploration of advanced topics. It can be used for in-class group work facilitated by the instructor or for homework assignments with networked interaction among students.
Inquiry-Based Learning: Cognitive Measures & Systems Support
NSF Grant REC-9720363

Screen shots of the three educational software applications developed for the project

Figure 1: CHAT, a simulation environment in which students explore the syntax of natural languages by designing and testing their own grammars and sharing their work with other students over the network.
Figure 2: Two screen shots of FOREST, a simulation environment in which students can explore, inspect, and alter a model of forest growth and ecology.
Figure 3: Geo Observer, a learning environment in which students make observations of natural scenes, develop hypotheses to explain their observations, and test their hypotheses against geological data collected at the sites.