Evaluating Learning about the Nature of Science in a Research Experiences Program

Pamela C. Burnley
Department of Geology, Georgia State University

(Condensed from "A Comparison of Approaches and Instruments for Evaluating a Geological Sciences Research Experiences Program" by Pamela C. Burnley, William Evans and Olga S. Jarrett, Journal of Geoscience Education, volume 50:1, p15-24, 2002)

Laboratory or field research experience offers us an alternative pedagogical model to classroom instruction. Generally, research experience is reserved for graduate students who have mastered much of their subject and are prepared to embark on self-guided learning. However, undergraduates can also participate in research and in some cases can gain a great deal from it. Research experiences are widely believed to be an important mechanism for recruiting undergraduates into science careers and for giving students an opportunity to test their interest in research (NSF, 1996; Mervis, 2001). Student learning in a research experience is different from a classroom experience many ways. The content knowledge that that student must master is generally more in depth than typical course work but much more limited in scope because it must be highly focussed on the research project. In addition, students must master a particular set of research skills that usually are in some ways independent of the content knowledge that they are learning. Research experiences tend to be highly personalized and unstructured compared to the classroom environment; expectations regarding the final outcome (poster presentation, paper, data, samples etc.) are particular to each program. Therefore, assessing in some standardized way, what content knowledge students have learned from research experiences would be very difficult.

From our own early research experiences, we know that we learned much more than what was explicitly taught to us. A research experience involves an enculturation (Kardash, 2000) or socialization process (Hogan and Maglienti, 2001) where students pick up styles of speaking, the structure of explanation, and attitudes towards science from their mentors (Bleicher et al., 1996). In other words, the common element that students learn in a research experience is what research is and how it is done. We hypothesize that examining changes in knowledge of the nature of science could form a basis upon which to measure and compare the effects of different kinds of research experiences on undergraduates.

As part of the Atlanta Consortium for Research in the Earth Sciences (ACRES), we run an NSF funded summer research experiences program at Georgia State University in Atlanta. The program takes 12-15 undergraduates as teachers for 8-week period each summer. The participants work in teams on four different geoscience research projects. The summer program also serves as a test bed for our efforts to understand and learn to quantify the impact of research experiences on undergraduates.

Survey Instruments

Knowledge of the nature of sciences has typically been treated, like knowledge of other subject matter, as something that can be measured by objective instruments (Hogan, 2000). Many tests and inventories have been developed that compare respondents understanding of the nature of science with the nature of science as understood by those developing the instruments. In order to ensure that instruments are valid, instruments are written such that there will be agreement among science educators and in some cases scientists about what the "correct" answer is. However, there does not appear to be uniform agreement as to what the nature of science actually is. Philosophy of science presents us with widely divergent views regarding the nature of science. Although the extent of disagreement is debated (Elfin et al., 1999), modern philosophers hold somewhat different beliefs about of the nature of science than those held by science educators (Alters, 1997). The views of practicing scientists are different from those of both philosophers and science educators (Pomeroy, 1993). In addition, there is good evidence to suggest that geoscience is not identical in nature to other sciences such as physics and chemistry (Ault, 1998; Frodeman, 1995; Peters, 1996). Certainly, one can easily determine via a casual conversation with one's colleagues that geoscientists hold a range of opinions about the nature of geoscience. Therefore, the task of developing an instrument with a key of "correct" answers is fraught with questions about the validity of the questions as well as their answers. The result is that many of items in existing instruments are very general, capturing science in its broadest form. On the instruments that we experimented with, we found that adults could pick the "correct" answers independent of their science experience.

Our approach to the "no right answer" problem is to stop thinking of the instrument as a test but instead think of it as an instrument like an oscilloscope that measures a signal. We do not expect an oscilloscope to render an exact replica of what it measures but rely instead on calibration with known input signals. In this case, our "signal" is the distribution of opinions about the nature of geoscience that a given population holds. We reason that we should therefore compare the "signal" that we get from undergraduates engaged in research to the "signal" we get from a population of geoscientists.

We have been working to develop our own instrument (Statements About Science Instrument (SASI)) for measuring undergraduates' (as well as science teachers involved in our program) understanding of the nature of science. The new survey instrument is based on clusters of statements representing a variety of philosophical positions, from which respondents must pick one statement.

For example:

a) Science is a collection of true facts.
b) Science is a procedure.
c) Science is a world view.

a) When examining data, logic is more important than creativity.
b) When examining data, creativity is more important than logic.
c) Examining data requires only logical thought.
d) Examining data requires only creative thought.

a) Science assumes cause and effect.

b) Science assumes nothing.

We compare the distribution of the choices made by a group of respondents with the distribution of choices made by geoscience faculty. The first version of the instrument was able to differentiate between three different groups of students with different science backgrounds. Some of the statement clusters detected changes in our research experience participant's attitudes over the course of a summer. For example, the percentage of participants who considered science a world view (cluster 1 above) and the percentage who considered science independent of culture became more like the faculty by the end of the summer. We believe that with further modification, an instrument can be developed that will detect changes induced by participation in a research experience.

Open ended-questions

We have also experimented with the use of the following open-ended questions:

1. What does it mean to study something scientifically?
2. What is a theory?
3. How can one distinguish good science from bad science?

Question 1 is borrowed from the National Science Board's Science and Engineering Indicators project, which has occasionally asked this question of a random sample of American adults (National Science Board 1993, 1998, 2000). Questions 2 and 3 were developed by us. Responses to the questions were analyzed using WordStat, a software package for text analysis.

Responses to question 1 were coded according to the same criteria used by the National Science Board (1993, 1998, 2000). A response was coded as adequate if it touched on the role of theory-building or -testing, the use of experiments, or the application of rigorous comparison. In pre- and post- testing of participants in our research experience program we found that the number of individuals providing an adequate answer increased from 62.5% to 88.9%; a statistically significant increase ($X^2 = 4.50$, $df = 1$, $p < .05$). In comparison, only 20% of a group of education graduate students that we tested provided adequate answers to this question.

Responses to question 2 were coded in terms of whether or not a respondent reported that theory was more than a guess or an opinion. Only 12.5% of participants in our research experience program initially responded that a theory is no more than a guess or an opinion and only one participant, or 5.6%, responded this way at the end of the experience. In comparison 42.1% of a group of education graduate students indicated that a theory was little more than a guess or an opinion.

Responses to question 3 were coded in terms of whether or not the respondent made reference to scientific method, the need for objectivity, or the application of peer review. Responses were also coded to assess whether or not the participant made any reference to social or ethical factors. In pre- and post- testing of participants in our research experience program we found that there were no changes at the end of the program; 87.5 % percent of participants offered adequate answers. In comparison only 42.1% of education graduate students answered this way. 16.7% of participants in the research program mentioned at least one social or ethical factor, which increased slightly to 21.1% at the end of the program. No education graduate students mentioned social or ethical factors although some mentioned other factors including inclusiveness and "hands-on".

The use of open-ended questions to probe participants' understandings of scientific processes holds promise. We find that those who have chosen to participate in a research experience program are well-prepared to grapple with such questions, and that their open-ended responses provide potentially rich data regarding their cognitive models of science. Although most of the changes we observed were not statistically significant, many of the differences between the education students and the participants in the research program were significant. The changes in responses to question 1 between pre- and post- testing were also statistically significant. We are currently experimenting with additional open-ended questions and plan to examine the resulting textual data for evidence of particular beliefs and the use of particular terms. Hopefully, this will allow us to determine how and why a research experience may be cultivating specific views regarding science.

References

Alters, B.J., 1997, Whose nature of science?: Journal of Research in Science Teaching, 34:1, 39-55.

Ault, C. R, Jr., 1998, Criteria of excellence for geological inquiry: The necessity of ambiguity.:Journal of Research in Science Teaching, 35:2, 189-212.

Bleicher, R.E., 1996, High school students learning science in university research laboratories: Journal of Research in Science Teaching, 33:10, 1115-1133.

Elfin, J.T., Glennan, S. and Reisch, G., 1999, Nature of science: a perspective from the philosophy of science: Journal of Research in Science Teaching, 36:1, 107-116.

Frodeman, R., 1995, Geological Reasoning: Geology as an Interpretive and Historical Science: Geological Society of America Bulletin, 107, 960-968

Hogan, K., 2000, Exploring a process view about students' knowledge about the nature of science: Science Education, 84:1, 51-70.

Hogan, K. and Maglienti, M., 2001, Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions: Journal of Research in Science Teaching, 38:6, 663-687.

Kardash, C. M., 2000, Evaluation of and undergraduate research experience: perceptions of undergraduate interns and their faculty mentors: Journal of Educational Psychology, 92:1, 191-201.

Mervis, J. 2001, Student Research: What is it good for?: Science, 293:5535, 1614-1615.

National Science Board, 1993, Science & Engineer Indicators: 1993, Arlington, VA, National Science Foundation.

National Science Board, 1998, Science & Engineer Indicators: 1998, Arlington, VA: National Science Foundation.

National Science Board, 2000, Science & Engineer Indicators: 2000, Arlington, VA: National Science Foundation.

National Science Foundation, 1996, Shaping the Future: New expectations for Undergraduate Education in Science, Mathematics, Engineering and Technology. (NSF 96-139)

Peters, E. K., 1996, No stone unturned: reasoning about rocks and fossils: New York , W.H. Freeman and Company.
Pomeroy, D.H., 1993, Implications of teachers' beliefs about the nature of science: Comparison of the beliefs of scientists, secondary science teachers, and elementary teachers: Science Education, 77:3, 261-278.

Freshmen Learning in A Web-based Chemistry Course[à]

Yehudit Judy Dori
Department of Education in Technology and Science
Technion, Israel Institute of Technology
Haifa 32000, Israel
yjdori@tx.technion.ac.il
Center for Educational Computing Initiatives
Massachusetts Institute of Technology
Cambridge, MA 02139-4307, USA
dori@ceci.mit.edu

Abstract

Chemistry courses in higher education have traditionally been composed of lectures, problem solving sessions and laboratories. This study describes a Web-based chemistry course and the learning outcomes of freshmen that used it. Chemistry faculty and teaching assistants were interviewed regarding their views about Web-based teaching and learning. Students who took part in a Web-based general chemistry course were divided into two groups based on their preference of participating in a Computerized Molecular Modeling (CMM) project. The experimental group students carried out an individualized project using CMM software to represent a complex molecule in three model types, compute its molecular weight, and construct hybridization and electrical charge distribution for each of the carbon atoms in the molecule. Pre- and post-tests along with final examination grades served for assessing the students achievements. The 95 experimental students achieved significantly higher grades than their 120 control group peers in both the post-test and the final examination. The experimental students were able to switch from 1-D to 2- and 3-D molecular representations, argue for selecting an appropriate substance for a particular purpose and transfer between the four levels of understanding in chemistry better than their control counterparts.

Introduction

Simulations, graphing, and microcomputer-based laboratories have been used in the last two decades as effective teaching methods in science education at both college and high school levels (*1-5*). Scientists, engineers and science educators use models to concretize, simplify and clarify abstract concepts, as well as to develop and explain theories, phenomena and rules. Researchers underscored the need for models as enablers of students' mental transformation from two-dimensional to three-dimensional representations (*6-8*). Virtual models enhance teaching and learning of various topics in chemistry. Studies have shown that when teaching topics, such as chemical bonding and organic compounds, is aided by 3D computerized models, students' understanding improves *(9-11)*.

During the past decade, science educators have been engaged in experimental projects that focus on the integration of the Internet and World Wide Web as an additional medium for teaching and learning. The Internet and the WWW are used as a source of scientific data and theoretical information (*12-14*), a tool for designing learning environments (*11*, *15-17*), integrating virtual models (*18*), and creating learning communities (*19-25*).

[◊] This paper is a short version of a paper submitted to the Journal of Chemical Education with Miri Barak and Noam Adir

While teaching the properties of substances and how they react, chemistry educators identified three levels of understanding: macroscopic, microscopic, symbolic (*26-29*). Dori and Hameiri (*30*) suggested additional fourth level – the process level, at which the substance is formed, decomposed, or reacts with other substances. Mastering this process level often requires higher order thinking skills as well as at least two of the previous three chemistry understanding levels. Researchers have shown that plastic and virtual models, such as Computerized Molecular Modeling (CMM), help students develop conceptual understanding (*31, 32)* as well as the ability to transfer across the various levels (*26-28*).

Methodology

Chemistry courses in higher education have traditionally been composed of lectures, problem solving sessions, and laboratories. This study, which took place at the Technion, Israel Institute of Technology, was aimed at developing a freshmen Web-based chemistry course and investigating the performance of the students who use it.

The research objective was to investigate the learning process in a Web-based environment. The research questions were:

(I) How did chemistry faculty, teaching assistants, and students view Web-based teaching and learning?

(II) How did an individual optional computerized molecular modeling project affect the students' learning outcomes?

Research population and setting

The research population consisted of 13 instructors and 53 students, who participated in a survey, and 215 students who participated in three Web-based chemistry courses. The instructors included seven chemistry faculty and six chemistry teaching assistants. Based on students' preference of participating in the optional computerized molecular modeling project, the 215 freshmen were divided into experimental (N = 95) and control (N = 120) groups. Only students who responded to the pre-test, post-test, and final examination were included in the research.

To validate the assumption that the baseline of the two groups (experimental and control) is identical we compared the entry-level grades (SAT and GPA equivalents). These grades are a combination of the high school matriculation examinations and a battery of psychometric tests in mathematics, English and Hebrew of the students in both groups. We found no significant difference between the two research groups regarding their entry-level grades. We also compared the two research groups in terms of their prior knowledge in chemistry and found no significant difference between the two groups.

Students in the two research groups studied in the same class with the same instructor and teaching assistants. Hence, the difference between the two research groups was that the experimental group carried out an individual project, which involved an intensive use of the Web and CMM and credited them with extra 5 points to their course's final grade. The project was handed out at the 6^{th} week, after the students had studied chemical bonding and molecular orbitals and was due for the last week of the semester.

Each student received a different complex molecule, which he or she had to download from the course website. We assigned each student in the experimental group with a certain molecule from a list of substances that are used on a daily basis, including Vitamin A, B, and C, Nicotine, Caffeine, Adrenaline, TNT, and DDT. The project required downloading two shareware programs (*33, 34*), one for writing the structural formula of the molecule, and the other for viewing and manipulating it in three representation forms: framework, ball-and-stick, and space-filling. The student was required to build virtual models of the molecule in three representation modes, compute its molecular weight, construct hybridization and electrical charge distribution for each of the carbon atoms in the molecule, and seek information on the Web about the daily use or applications of the substance.

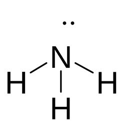
Students carried out the project voluntarily in their free time in addition to the regular course load. The control group students elected not to participate in this activity. All the students in the three courses, regardless of whether or not they elected to undertake the optional individual project, were exposed during lectures to examples of molecules represented by the same CMM software tools (*33, 34*). In addition, two recitation sessions were devoted to practice building molecules with those packages.

Research Tools

Research tools included semi-structured personal interviews with faculty, teaching assistants and experimental students, a students' survey, and pre- and post-tests. The faculty and teaching assistant's interviews, and the students' survey were administered prior to the development of the Web-based chemistry courses. The results served as guidelines for constructing the Internet sites and the CMM project that were used in the courses.

To investigate students' learning outcomes we used chemistry understanding pre- and post-tests, entry-level grades, and final examination scores. The pre- and post-tests were similar and included three questions with images of models that appear in general chemistry textbooks. The tests were aimed at assessing students' chemistry understanding. The first question investigated students' ability to apply transformation between the four levels of chemistry understanding: macroscopic, microscopic, symbolic and process (*11, 30*). The second question, presented in Table 1, investigated students' ability to apply transformation from one-dimensional molecular representation to two- and three-dimensional representations, and vice versa. This question was developed and validated by Dori and Barak (*11*). The third question, which was developed and validated by Reid (*35, 36*) investigated students' ability to answer a higher order thinking skills question.

Table 1. Question 2 of the pre- and post-test

Compound	Molecular formula	Structural formula	Spatial structure	Hybridization (sp, sp^2, sp^3)	Model
Ethanol	C_2H_6O				
				sp^3	
			Triangular pyramid		

Results

Attitudes toward Using Web and IT in Chemistry Courses

Interviews with faculty and teaching assistants indicated that none of them had used information technology (IT) for teaching a general chemistry freshmen course.

Their attitudes towards the use of computers and the Internet in teaching and learning chemistry were mixed and ambivalent. Responses were classified into three categories: (1) Interested in Web-based teaching; (2) Not interested in Web-based teaching; and (3) Undecided. Faculty and the teaching assistants who expressed interest in using the Web, wanted to use it for various purposes, which are listed below along with interviewee responses.

- Information extracting and problem solving: *"I can refer interested students to the Web, so they can find enriching information."*
- Modeling: *"If I had a big screen in the class, I could show the students computerized demonstrations. Even showing one picture or a video clip of an experiment is important."*
- Assessment: *"Students can take a computerized test, and the teacher gets a summary of the results."*

The instructors who were not interested in using information technology indicated that they did not want to change their teaching methods. Some comments were: *"It is fine for a young lecturer who is starting his career," "It is difficult to change old habit,"* and *"I am not familiar with the Internet."* Some were concerned about losing the personal contact with the students: *"I am against the use of computers because I believe we need to work more intimately with the students… to allow students who do not understand the learning material to raise their hands, stop me during the lecture and ask a question."*

The interviewees who were interested in Web-based teaching, expressed reservations regarding the time required for preparing a Web-based course, incorrect information presented on the WWW, technical problems, and the lack of computers in the lecture halls. Conversely, teachers who were not interested in IT-enhanced teaching, mentioned positive aspects, such as the variety of teaching methods, students' motivation, and the ability to visualize abstract concepts.

Analyzing the students' survey, we found that 95% responded positively to the open question, which was "Would you like to learn chemistry in a Web-based and Computerized Molecular Modeling environment? If so, specify the preferred chemistry topics." This indicates that the vast majority of students were interested in learning chemistry in a Web-based environment. More than half of the students chose organic compounds and stereochemistry, and almost one third chose atom structure and chemical bonding. These topics, which are taught in freshmen general chemistry courses, were indeed found in other studies to be best taught with computerized molecular modeling (*9, 11, 31*).

Students who studied in a Web-based environment were asked to specify the number of times and purposes for entering the course Web site. The differences between the experimental group (students who carried out the CMM project) and control group are presented in Figure 1. The site was mainly used for accessing homework assignments, getting their solutions, and reading course summaries. Students who elected to carry out the project were also engaged in reading peer's projects, linking to other chemistry sites and downloading computerized molecular modeling programs.

Only a few students used the forum to contact the teaching assistants and ask them questions. The individual project required an intensive use of the Web and computerized molecular modeling software. Figure 2 shows an example of part of a CMM project.

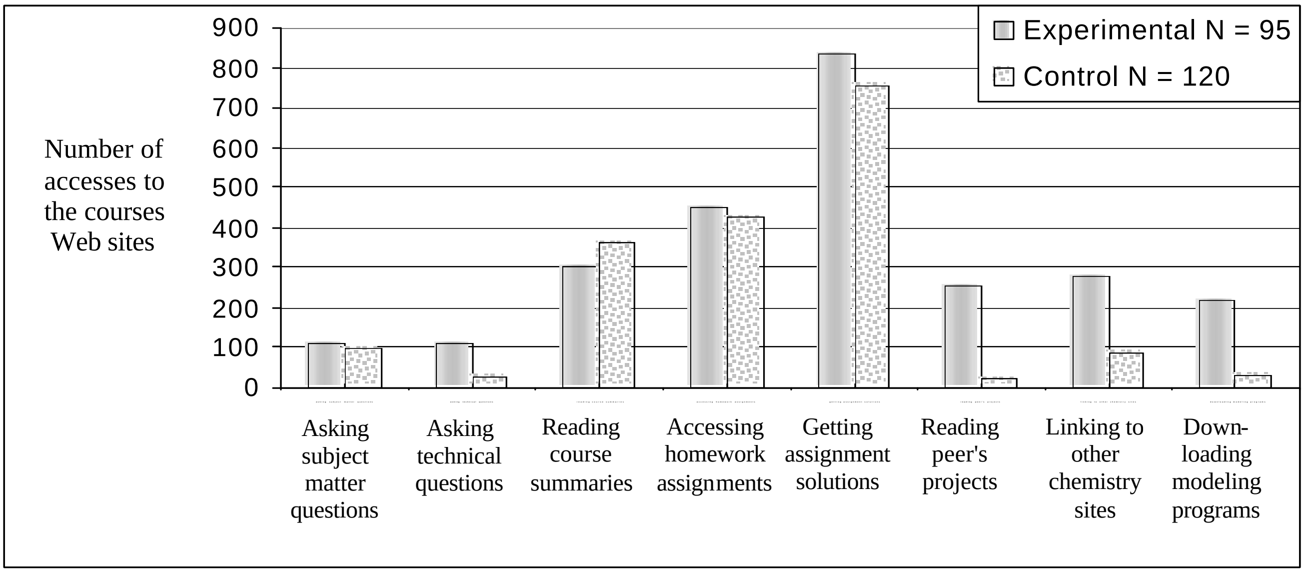
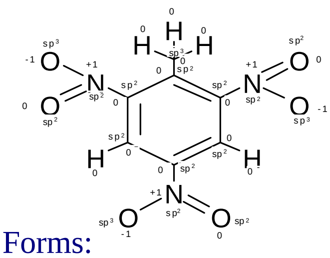


Figure 1. Comparison of frequency and purpose of accessing the courses Web-sites between the research groups

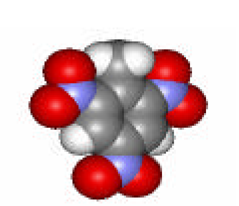
Trinitrotoluene (TNT) molecule

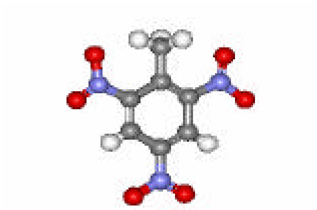
Molecular Formula: $CH_3C_6H_2(NO_2)_3$
Description: pale yellow crystals
Melting point: 82° C (180° F). Its low melting point allows it to be melted and poured into artillery shells and other explosive devices.
Density: 1.65 gr/cm^3
Burns at: 295° C (563° F), but it may explode if confined.
Hybridization and formal charge:

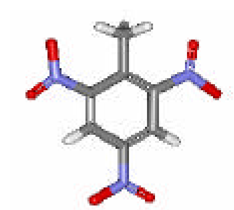


Computerized Molecular Models in Three Representation Forms:

Space-filling **Ball-and-stick** **Framework**







Trinitrotoluene (TNT) is prepared by the nitration of toluene. Trinitrotoluene is highly explosive, but, unlike nitroglycerin, it is unaffected by ordinary shocks and jarring, and must be set off by a detonator. Because it does not react with metals, it can be used in filling metal shells. It is often mixed with other explosives, e.g., with ammonium nitrate to form amatol.

Figure 2. A student's CMM Project

Students' Achievements in the Web-based Chemistry Course

To analyze the effect of this project on students' achievements, we used analysis of covariance (ANCOVA). Although the pre-test average scores of the experimental and control group students were very close (30.14 and 31.82 respectively) the pre-test scores were used as the covariant for the post-test analysis. The entry-level grade and the pre-test scores were used as the covariant for the final examination analysis. As noted, no significant difference was found between the research groups regarding their entry-level grades and their prior knowledge in chemistry.

Table 2. Analysis of covariance of the post-test and the final examination scores

Dependent variable	Research group	N	$\overline{X}$	SD	F	p <
Post-test score	Experimental	95	72.65	17.56	57.49	0.001
	Control	120	53.52	19.38		
Final exam score	Experimental	95	70.28	18.90	5.19	0.02
	Control	120	62.02	25.23		

Table 2 shows that the experimental group students received significantly higher scores on both the post-test and the on the course final exam. We assumed that the extra activities that experimental students carried out while studying the general chemistry course improved their chemistry understanding and higher order thinking skills to a larger extent than their control group peers.

We analyzed students' responses to each of two questions individually. Question 2, presented in Table 1, tested students' ability to apply transformations to and from one-dimensional molecule representation to two- and three-dimensional representations. To analyze the effect of the CMM project on students' ability to apply transformations, we perform an analysis of covariance (ANCOVA), using the pre-test scores as the covariant (see Table 3). We found that the integration of the CMM project into the general chemistry course was the main source for the difference in the students' ability to apply transformations (scores of question 2).

Table 3. Analysis of covariance of the transformation abilities between three, two and one molecule representation modes in the post-test

Source of variant	SS	DF	MS	F	p <
Learning method (integrating the CMM project)	86.61	1	86.61	26.68	0.001

Analyzing the models students had drawn in Question 2, we found that the experimental group students filled 73% of the blank cells with models (see Table 1), while the control group students filled 51% of the blank cells. Students' drawings of NH_3 (Figure 3) and CH_3CH_2OH molecule models depict typical differences between the two research groups. Space-filling model was the most popular molecule representation among the experimental group, and accounted for 70% of the drawings. Among the control group, the ball-and-stick model was the most popular molecule representation, accounting for 46% of the drawings. As Figure 3 demonstrates, most experimental group students - 83% (as opposed only 5% of the control student) drew the non-bonding electrons in the ammonia molecule model, and some of them drew tetrahedrons models.

Figure 3. Drawings of an ammonia molecule

a. ***Experimental group students*** *drew a space-filling model or a tetrahedron, both including the non-bonding electron pair.*

b. ***Control group students*** *drew ball-and-stick or space-filling models without the non-bonding electrons.*

Other differences were reveled in drawing of a C_2H_5OH molecule model. Experimental group students were thorough and detailed when drawing 3D molecular models. They showed the tetrahedral angle (109.5^0) and drew atoms in front and behind the central atom. In contrast, most control group students drew the models as if the atoms were connected at 90^0 angles. Experimental group students used size and color to differentiate between the carbon, oxygen, and hydrogen atoms in the molecule. Models that control group students sketched were less meticulous about these aspects.

Question 3 in the pre- and post-tests, which we evaluated in detail, required higher order thinking skills. It tested students' ability to analyze information about six compounds, select the best anaesthetic substance and provide argument for their choice. Given that ether is flammable and chloroform is known to cause liver damage, the students were asked to select the best alternative anaesthetic and provide arguments for their choice.

The focus of our analysis in this question was the level of students' arguments and their ability to transfer between four understanding levels in chemistry: macroscopic, microscopic, symbolic and process. The correct answer should be $CF_3CHClBr$ and is based on experimentation (*35, 36*), which cannot be expected of chemistry students. Therefore, we based our evaluation on the quality as well as the quantity of the arguments provided. Students were expected to refer in their arguments to the substance physical and chemical properties: structural formula, molecular mass, boiling point, AD_{50} (anaesthetic dose), LD_{50} (lethal dose), anaesthetic index, and halogen percentage.

The responses were categorized into three groups: (1) high level arguments, (2) partial or insufficient arguments, and (3) no argument. An example of an experimental group student's response from the post-test follows. Interleaved within the student's response in italics are our interpretations of the transformations between the four understanding levels.

"$CF_3CH_2CF_3$ *is a good possibility…"* – reference to the symbol level.

"Due to its high boiling point, it will not evaporate in room temperature or in the patient's body. It can be injected in low concentration (we do not need a lot of the substance). Its lethal dose is very low. On the other hand, its anesthetic index is high…"– reference to the macro and micro levels.

"Also, since its halogen percentage is high, there is little chance that the carbon compound will burn when mixed with air." – reference to transfer from the micro (*halogen percentage)* to the process (*will burn*) level.

This well-founded response was categorized as being at the high level. Conversely, a post-test example of a partial, insufficient response, given by a control group student, was: "$CF_3CH_2CF_3$ *is best because the anesthetic index is the highest."*

Analyzing the students score in this question we found a significant difference between the experimental and control grads ($F = 31.08$, $p < 0.001$)

In the pre-test, 65% of both research group students provided no argument whatsoever to support their choice and the remaining responses contained partial or insufficient arguments. As Figure 4 shows, in the post-test the two research groups differed in their argument level.

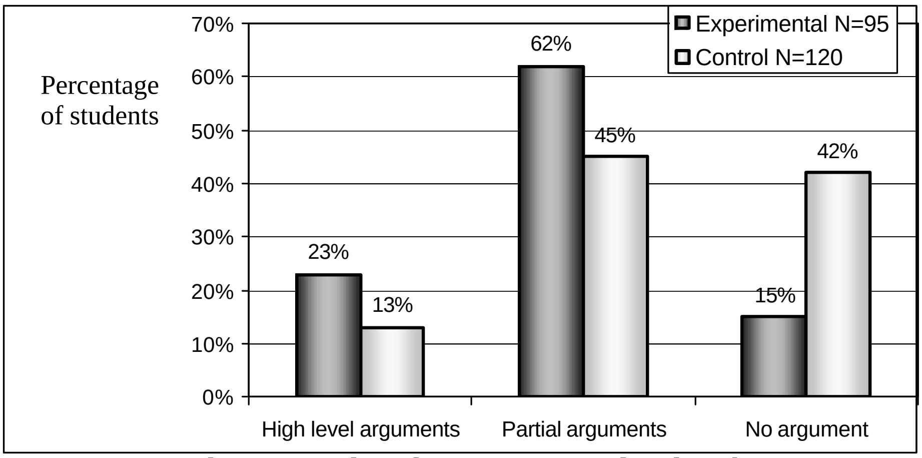


Figure 4. Experimental vs. control students' argument level in the post-test

The percentage of students who provided high level arguments in the experimental group was nearly twice as much as that of their control group peers, while for partial arguments it was 1.4 times as much. Conversely, the percentage of students who gave no argument in the experimental group was one third of the corresponding percentage in the control group. As these results show, experimental students demonstrated better argumentation skills as well as better ability to transfer between the four chemistry understanding levels.

One limitation of our research is that the experimental students were not chosen randomly but based on their willingness to take on the extra project. This may indicate that they were inclined to spend the extra effort and time required, some in order to achieve a higher grade and others because they were more motivated.

Discussion and Summary

As Bunce and Robinson (*37*) have noted, the chemical education community encompasses three intertwined activities: instruction, practice, and research. Many of the chemical educators are involved in at least two of these activities. Indeed, our study was feasible thanks to collaboration among chemistry and chemical education faculty and instructors. We have been actively engaged in Web-based instruction, practicing with chemistry undergraduate and graduate students. One author investigates three-dimensional structures of biological macromolecules (*38*, *39*), while the others study learning processes that employ computerized molecular modeling (*2*, *8*, *9*, *11*).

Based on students' interviews and our observations in class, the use of the Web as a source of a variety of molecular modeling software, inspired students in our research, as well as in the research described in (*40*), and created an enthusiastic learning environment. We found that students were in favor of Web-based chemistry courses despite the fact that chemistry faculty had various reservations as to their readiness to apply IT-enhanced teaching in their classroom. Students noted that access to Web-based learning materials and assignments was valuable, as it contributed to their learning experience. In the interviews with students during their work on the project, some indicated that they had started the project (and the course in general) with low motivation and gained motivation to study chemistry as a result of working on the project. It thus appears that the project enhanced students' motivation to study chemistry.

Incorporating Web-based assignments and computerized molecular modeling into the chemistry courses has been found to foster understanding of molecular 3-D structure and related properties (*9, 11, 15*). Williamson and Abraham (*31*) found that engaging in dynamic animations of molecules promote deeper encoding of information than that of static pictures. Our research aimed at improving and promoting higher education chemistry teaching through the development, implementation, and assessment of a Web-based freshmen general chemistry course. Our findings indicate that IT-enhanced teaching positively affects students' achievements provided the students are actively engaged in constructing computerized models of molecules. These results are in line with the findings of Kantardjieff et al. (*40*), and of Donovan and Nakhleh (*15*). Kantardjieff et al. found that sophomore students, who engaged in exploration activities, learned to apply modern chemistry software packages, and acquired skills needed to become practitioners of their discipline. Donovan and Nakhleh concluded that the Web site used in their general chemistry course was instrumental in visualizing and understanding chemistry.

The level of students' engagement with Web-based activities depended on the assignments they were required to deliver as part of the course. In study (*15*), students could succeed in the course without using the Web and in fact, low academic level students accessed the Web more frequently than high academic level ones because they viewed it as a supplementary source of help. In our study, all the students who elected to undertake the Web-based computerized molecular modeling project (the experimental group) performed significantly better in both the post-test and the final examination than those who elected not to carry out the project (the control group). We found that low academic level students of the experimental group made the greatest progress in chemistry understanding.

Experimental students at all academic levels applied transformations from one-dimensional molecule representation, to two- and three-dimensional representations, and vice versa better than their control group peers. The differences in drawings of molecular models between the two research groups indicated that experimental group students understood the geometric structure of molecules and their related physical and chemical properties better than the control group students.

Harrison and Treagust (*41*) noted that students who were encouraged to use multiple models demonstrated understandings of particles and their interactions better than students who searched for one best model. In our research, the experimental students carried out an individualized project using computerized molecular modeling software to represent a complex molecule in three model types, compute its molecular weight and construct hybridization and electrical charge distribution for each of the carbon atoms in the molecule. As a result of their interaction with the software to execute their project, they were better prepared to argue for selecting an appropriate substance for a particular purpose and could carry out transformation between the four levels of understanding in chemistry.

While other means, such as plastic models and extra recitations hours, might have replaced the Web-based learning environment, in the long run, technology-rich environment is less labor-intensive and provides for asynchronous, interactive learning. Indeed, our Web-based chemistry course has proven to be an effective means to foster freshmen learning and should therefore be further practiced and investigated with the objective of establishing the elements that contribute the most to enhancing students' higher order thinking.

References

1. DeBard, R.; Guidera, S. *J. Educ. Tech. Systems*, **2000**, *28*, 219-230.
2. Dori, Y. J. *J. Sci. Educ. and Tech.* **1995**, *4*, 163-170.
3. Lazarowitz, R.; Tamir, P. In Gable, D.L. (Ed.), *Handbook of Research on Science Teaching and Learning*. Macmillan Publishing Company, N. Y. **1994**, pp. 94-128.
4. Nakhleh, M.B. *J. Comp. in Math. Sci. Teach.* **1994**, 13, 368-381
5. Nakhleh, M.B.; Krajcik, J.S. *J. Res. Sci. Teach.* **1994**, 31, 1077-1096.
6. Baker, R.S.; Talley, H.L. . *J. Res. Sci. Teach.* **1974**, 11, 95-97.
7. Eliot, J.K.; Hauptmam, A. *Studies in Sci. Educ.* **1981**, 8, 45-66.
8. Dori, Y.J.; Barak, M. *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference*, C. Hoadley; J. Roschelle (Eds.) Dec. 12-15, **1999**, Stanford University, Palo Alto, California. Mahwah, NJ: Lawrence Erlbaum Associates. http://kn.cilt.org/cscl99/A16/A16.HTM
9. Barnea, N.; Dori, Y. J. *J. Sci. Educ. and Tech.* **1999**, *8*, 257-271.
10. Copolo, C.F.; Hounsell, P.B. *J. Sci. Educ. and Tech.* **1995**, *4*, 295-305.
11. Dori, Y. J.; Barak, M. *Educ. Tech. and Soc.* **2001**, *4*, 61-74. http://ifets.ieee.org/periodical/vol_1_2001/dori.pdf
12. Merlic, C. A.; Fam, B. C.; Miller, M. M. *J. Chem. Educ.* **2001**, *78*, 118-120.
13. McGrowan, C.; Sendall, P. *J. Chem. Educ.* **1997**, *74*, 391.
14. Waldow, D. A.; Fryhle, C. B.; Brock, J. C. *J. Chem. Educ.* **1997**, *74*, 441-442.
15. Donovan, W.J.; Nakhleh, M.B. *J. Chem. Educ.* **2001**, *78*, 975-980.
16. Foust, R. D.; Cruickshank, B.; Stringer, M.; Olander, J.; Curley, A. O. *The Chem. Educ.* **1999**,*4*, 108-111. http://journals.springer-ny.com/chedr
17. Lui, D.; Walter, J.; Brooks, D. *J. Chem. Educ.* **1998**, *75*, 123-125.
18. Kozma, R.B.; Russell, J. *J. Res. Sci. Teach.* **1997**, *34*, 949-968.
19. Glaser, R. E.; Poole, M.J. *J. Chem. Educ.* **1999**, *76*, 699-703.
20. Gordin, D. N.; Gomez, L. M.; Pea, R. D.; Fishman, B. J. *J. Comp. Mediated Commun.* **1997**, *2*. http://cwis.usc.edu/dept/annenberg/vol2/issue3/gordin.html#foot1.
21. Hughes, C.; Hewson, L. *Educ. Tech.* **1998**, July-August 48-55.
22. Judd, C. S. *J. Chem. Educ.* **2000**, *77*, 808.
23. Paulisse, K.W; Polik, W.F. *J. Chem. Educ.* **1999**, *76*, 704-708.
24. Schwartz, A. T.; Bunce, D. M.; Silberman, R. G.; Stanitski, C. L.; Stratton, W. J.; Zipp, A. P. *J. Chem. Educ.* **1994**, *71*, 1041.
25. Smith, S. G.; Stovall, I. K. *J. Chem. Educ.* **1996**, *73*, 911-915.
26. Gabel, D.L. In Fraser, B.; Tobin, K. Eds. *International Handbook of Science Education*, London: Kluwer Academic Pub. **1998**, pp. 233-248.
27. Gabel, D.L. & Bunce, D.M. In: Gable, D. L. Ed. *Handbook of Research on Science Teaching and Learning*, Macmillan: N.Y. **1994**, pp. 301-326.
28. Herron, J. D. *The Chemistry Classroom: Formulas for Successful Teaching*; American chemical Society: Washington, D.C. **1996**.
29. Johnstone, A. H. *J. of Comp. Assisted Learn.* **1991**,*7*, 75-83.
30. Dori, Y. J.; Hameiri, M. *Int. J. Sci. Educ.* **1998**, *20*, 317-333.;
31. Williamson, V.M.; Abraham, M.R. *J. Res. Sci. Teach.* **1995**, *32*, 521-534.
32. Harrison, A.G.; Treagust, D.F. *Sci. Educ.* **1996**, *80*, 509-534.
33. WEBLAB - http://www.msi.com/life/products/weblab/viewer/index.html
34. ISIS DRAW - http://www.mdli.com/cgi/dynamic/welcome.html
35. Reid, N. *Chem. Educ.: Res. and Prac. in Europe*, **2000**, *1*, 381-392. http://www.uoi.gr/conf_sem/cerapie/2000_October/09.html
36. Reid, N. (1999). Bringing work experience to undergraduates. *Education in Chemistry*, **36**(1), 23-24
37. Bunce, D. M.; Robinson, W. R. *J. Chem. Educ.* **1997**, *74*, 1076.
38. Adir, N. *Acta Crystalographica*. **1999**, *D55*, 891-894.
39. Asojo, O.; Friedman,J.; Adir, N.; Belakhov, V.; Shoham, Y.; Baasov, T. *Biochem.* **2001**, *40*, 6326-6334.
40. Kantardjieff, K. A.; Hardinger, S. A.; Van Willis, W. *J. Chem. Educ.* **1999**, *76*, 694-697.
41. Harrison, A.G.; Treagust, D.F. *Sci. Educ.* **2000**, *84*, 352-381.

Learning-for-Use in Earth Science

Daniel C. Edelson
School of Education and Social Policy
Northwestern University

Prepared for a workshop entitled "Bringing Research on Learning to the Geosciences" sponsored by NSF and the Johnson Foundation. July 8-10, Racine, WI.

Abstract

The Learning-for-Use design framework provides guidance for curriculum designers based on principles from cognitive science research. The design framework offers design strategies for the three stages of learning that are necessary to acquire knowledge that will be accessible and usable in situation where it is useful. I briefly describe the Learning-for-Use design framework and Planetary Forecaster, a middle school Earth systems science unit that was developed based on the framework.

1. Learning for Use

In this paper, we are concerned with the design of learning activities that develop what we call *useful knowledge* Our primary motive is that address two significant challenges to teaching that are often overlooked in the design of learning activities: fostering engagement and ensuring that learners develop knowledge that they can access and apply when it is relevant. Regardless of the nature of the learning activities that students participate in, if they are not sufficiently and appropriately engaged, they will not attend to those activities in ways that will foster learning. Likewise, if students do not construct knowledge in a manner that supports subsequent re-use of that knowledge, it remains inert (Whitehead, 1929). To address these problems, we have developed the Learning-for-Use model and design framework based on contemporary research in cognitive science (Edelson, 2001). The Learning-for-Use model is a model of the learning process that describes how learners can develop useful knowledge. The Learning-for-Use design framework provides guidance to instructors and curriculum developers on how to design learning activities that foster engagement and useful understanding.

The Learning-for-Use model describes the learning process that results in useful knowledge. It builds on fundamental theories of learning with the express aim of supporting designers in the development of learning activities. The model is based on four principles that are shared by many contemporary theories of learning. The four principles are[1]:

1. Learning takes place through the construction and modification of knowledge structures.
2. Knowledge construction is a goal-directed process that is guided by a combination of conscious and unconscious understanding goals.
3. The circumstances in which knowledge is constructed and subsequently used determine its accessibility for future use.
4. Knowledge must be constructed in a form that supports use before it can be applied.

The Learning-for-Use model incorporates these four principles and their implications into a description of learning. The Learning-for-Use model characterizes the development of useful understanding as a three-step process consisting of (1) motivation, (2) knowledge construction, and (3) knowledge refinement.

Motivate: Experiencing the need for new knowledge[2]. The first step in learning for use is recognizing the need for new knowledge. The *motivate* step in the learning-for-use model creates a need for specific content understanding. In this context, *motivate* is being used in a very specific sense. It describes the motive to learn specific content or skills, not a general attitude or disposition to learn in the particular context. Understanding the usefulness of what they are learning for tasks that are meaningful to learners, provides a motivation for students to engage in learning activities and to construct understanding in a useful form.

Knowledge Construction: Building new knowledge structures. The second step in learning for use is the development of new understanding. This step results in the construction of new knowledge structures in memory that can be linked to existing knowledge. An individual constructs new knowledge as the result of experiences that enable him or her to add new concepts to memory, subdivide existing concepts, or make new connections between concepts. The "raw material" from which a learner constructs new knowledge can be firsthand experience, communication from others, or a combination of the two. This step in the Learning-for-Use model recognizes incremental knowledge construction as the fundamental process of learning.

Knowledge Refinement: Organizing and connecting knowledge structures. The third step in learning-for-use is refinement, which responds to the need for accessibility and applicability of knowledge. In the refinement step, knowledge is re-organized, connected to other knowledge, and reinforced in order to support its future retrieval and use. To be useful, declarative knowledge must be reorganized into a procedural form that

---

[1] From (Edelson, 2001)

2 While the first step in the Learning-for-Use model is called *motivate*, this phase is only concerned with a small portion of what is normally thought of as *motivation* in education. In this context, I am using *motivate* to refer to a specific type of motivation—the motivation to acquire specific skills or knowledge within a setting in which the student is already reasonably engaged. Addressing the broader motivational challenges of engaging students in schooling are critical to, but beyond the scope of, the Learning-for-Use model.

supports the application of that knowledge (Anderson, 1983). Useful knowledge must also have connections to other knowledge structures that describe situations in which that knowledge applies (Chi, Peltovich, & Glaser, 1981; Glaser, 1992; Kolodner, 1993; Schank, 1982; Simon, 1980). Refinement of knowledge can also take the form of reinforcement, which increases the strength of connections to other knowledge structures through the traversal of those structures and increases the likelihood that those connections between knowledge structures will be found in the future.

While there is an inherent ordering among these three steps, the ordering does not preclude overlaps or cycles. For example, knowledge construction and revision may be interleaved, and knowledge construction or revision can create new motivation. Because of the incremental nature of knowledge construction, it can require several cycles through various combinations of the steps to develop an understanding of complex content. Even with this cyclical nature, the order of steps is important. To create the appropriate context for learning, motivation must precede construction, and to insure accessibility and applicability, refinement must follow construction.

2. The Learning-for-Use Design Framework

Based on this model of learning for use, we have developed the *Learning-for-Use Design Framework*. This framework provides guidelines for the design of activities that will contribute to the development of robust, useful understanding. The design framework articulates the requirements that a set of learning activities must meet to achieve particular learning objectives. The Learning-for-Use model poses the hypothesis that for each learning objective a designer must create activities that effectively achieve all three steps in the learning for use model.

The Learning-for-Use design framework describes different design strategies that meet the requirements of each step (Table 1). The different design strategies for each step can be treated as alternative or complementary ways to complete the steps. In the case of rich content, however, several learning activities at each step involving both of the processes for that step may be necessary.

Table 1: Overview of the Learning-for-Use Design Framework.

Step	Design Strategy	Student Experience
Motivate	Activities *create a demand* for knowledge when they require that learners apply that knowledge to complete them successfully.	Perceive need for understanding
	Activities can *elicit curiosity* by revealing a problematic gap or limitation in a learner's understanding.	Experience curiosity
Construct	Activities that provide learners with *direct experience* of novel phenomena can enable them to *observe* relationships that they encode in new knowledge structures.	Experience or observe phenomena
	Activities in which learners receive direct or indirect *communication* from others allow them to build new knowledge structures based on that communication.	Hear, view, or read about phenomena
Refine	Activities that enable learners to *apply* their knowledge in meaningful ways help to reinforce and reorganize understanding so that it is useful.	Apply understanding
	Activities that provide opportunities for learners to retrospectively *reflect* upon their knowledge and experiences retrospectively, provide the opportunity to reorganize and reindex their knowledge.	Reflect upon experiences or understanding

Although it was designed to describe learning in general, when applied to inquiry-based science learning, the Learning-for-Use design framework represents a variant of the Learning Cycle (Abraham, 1998; Karplus & Thier, 1967; Lawson, 1995; Renner & Stafford, 1972). While they are similar in many ways, the two frameworks were developed with different goals. The Learning Cycle was developed as way to bring the process of learning from inquiry that scientists engage in to students. The Learning-for-Use design framework has been developed to highlight the need for motivation based on usefulness and the need to develop knowledge that is organized to support access and application (the *motivate* and *refine* phases in the framework), because they are too often overlooked.

3. Planetary Forecaster[3]

Planetary Forecaster is a middle school curriculum unit for Earth systems science that we have developed using the Learning-for-Use design framework. It combines computer-supported investigations of geospatial data with hands-on laboratory activities in which students observe and measure the phenomena under study. The Planetary

[3] Planetary Forecaster is a revised version of the Create-A-World activity, which is described in greater detail in Edelson (2001)

Forecaster curriculum unit is the product of an ongoing iterative development effort that involves teachers both directly as members of design teams and indirectly as implementers who are observed or provide feedback. The curriculum has been through three revision cycles based on three cycles of classroom implementation.

3.1 Unit Scope and Sequence

The content goal for the unit is for students to understand how physical geography influences temperature at a climatic timescale. The premise of the curriculum unit is that students have been asked by a fictional space agency to identify the portions of a newly discovered planet that are habitable given information about the planet's topography, water cover, and the tilt of its axis. For simplicity, the planet has the same atmospheric make-up as Earth, is orbiting around a star with the same intensity as the sun, and has an orbit with the same radius as Earth's. This mission is designed to create a demand for understanding of the curriculum's target content.

There are four major relationships that students must understand to complete the task. They are:

Curvature—The effect of a planet's curved surface on the intensity of the solar radiation received at each point. Temperatures decreases toward the poles because of the planet's curvature. Where the plane of the ecliptic cuts through Earth, sunlight hits the surface at a ninety-degree angle; as you move toward the poles, the angle at which sunlight hits the Earth decreases, which in turn increases the amount of surface area covered by the same quantity of light and decreases the intensity of the light in any unit area. This decreased intensity has a smaller heating effect.

Tilt—The effect of the tilt of a planet's axis of rotation on temperatures at different times of year. Because the Earth is tilted on an axis, the angle at which the sunlight hits the Earth's surface at each latitude changes on a seasonal cycle. Between the March and September equinoxes, the location of the most direct sunlight is north of the equator (northern hemisphere spring and summer), and between September and March, the location of the most direct sunlight is south of the equator (southern hemisphere spring and summer).

Land/Water heat capacity—The effect of surface cover (land vs. water) on the temperatures at different locations. Water takes longer than land to heat/release heat. This causes temperature differences between water and land depending on the time of year. Generally, air over water is cooler than over land in summer and warmer in winter.

Topography—The effect of surface elevation on the temperatures at different locations. Temperatures decrease as elevation increases. This results from air pressure decreasing as elevation increases.

Understanding these relationships requires an understanding of fundamental scientific concepts that are commonly found in national, state, and local standards documents, such as the Earth-sun relationship, radiative energy transfer, conservation of energy, heat and temperature, specific heat, and the ideal gas law.

The curriculum is divided into seven sections that take from 1-5 class periods each:

1. Setting the stage. In this section, students conduct an exercise in articulating prior conceptions in which they draw color maps showing their current conceptions of

global temperatures. They then compare their maps with actual data from Earth and formulate initial hypotheses about the factors that influence temperature.

2. Getting the task. Students learn about their mission of identifying habitable regions on a newly discovered planet, *Planet X*. They do an exploration of habitable regions on Earth. (For the purposes of this unit, habitable is defined as having minimum temperatures above 25F and maximum temperatures below 80F.) Students are assigned to investigate the four factors listed above (shape, tilt, surface cover, and elevation), to investigate for their influence on temperature. They are told that they will receive data about the shape, tilt, surface cover, and topography of Planet X that will help them to develop a map predicting the distribution of temperature on Planet X.
3. Investigating shape. Students investigate the effect of angle of incidence of solar energy on surface temperature through hands-on labs and explorations of global incoming solar energy data for Earth. They create an initial temperature map for Planet X that shows variation of temperature with latitude.
4. Investigating tilt. Students investigate the effect of a tilted axis of rotation on temperature at different times of year, through explorations of incoming solar energy data for Earth. They observe how the bands of incoming solar energy shift with seasons. They modify their temperature map for Planet X to account for seasonal differences.
5. Investigating surface cover. Students investigate the effect of land versus water on temperatures through hands-on labs looking at specific heat of water and soil and explorations of global surface temperature data for Earth. They modify their temperature map for Planet X to account for differences in temperature over land and water.
6. Investigating elevation. Students investigate the effect of elevation on temperature through explorations of global surface temperature data for Earth. They modify their temperature map for Planet X to account for differences in temperature at different elevations.
7. Final Recommendations. Students identify habitable areas by looking at maximum and minimum temperature values in their temperature maps for Planet X. They present their findings and their recommendations for colonization.

The curriculum materials place a special emphasis on forming and revising hypotheses and includes journaling activities that ask students to record their hypotheses together with evidence and rationales. Teachers have the option of using a computer-based inquiry-support tool called the Progress Portfolio (Loh et al., 2001), which gives students a place to store visual records of their work and provides prompts students to structure students' journaling. At each stage of the curriculum, students are asked to describe the factors that they believe affect temperature, how they affect temperature (i.e., the direction of the effect), and why (i.e, the underlying causes). They are also asked to provide any evidence they might have for these hypotheses and any open questions. They first record their hypotheses about the factors that affect temperature during the initial "setting the stage activity". During the portions of the unit where they investigate individual factors, they record their initial hypothesis about how each factor affects temperature before they do their investigations, and then they record their revised understanding following the investigation. It is this revised description of the relationship between a particular factor and temperature that they use when they construct their temperature maps for Planet X.

3.2 Planetary Forecaster as an Example of Learning-for-Use

Planetary Forecaster incorporates all six design strategies in the Learning-for-Use design framework to achieve all three steps of learning.

Motivate. The curriculum *creates a demand for understanding* through the mission of determining habitable areas on Planet X. This mission requires that they model temperatures for Planet X based on the provided data about the planet, which in turn demands that students understand the relationships between physical geography and temperature that comprise the content learning objectives for the unit. It also *elicits curiosity* through the stage-setting activities which ask students to articulate there prior conceptions and confront them with the limitations of their understanding. After trying to create temperature maps for Earth off the tops of their heads, students become curious about what the actual temperature patterns are and why they are that way.

Construct. Students learn about the relationships between physical geography and temperature through a combination of hands-on labs, computer-based investigations of Earth science data, readings, lectures, and discussions. The hands-on labs provide them with direct, concrete *experiences* with the phenomena and relationships they are learning about. The computer-based investigations provide them with *observations* of these same relationships at a scale that they cannot experience directly. The readings, lectures, and discussions help to *communicate* information about the phenomena and relationships from which they can construct understanding.

Refine. The process of constructing temperature maps for Planet X gives students the opportunity to *apply* their understanding of the relationships between physical geography and temperature as they are developing it. Classroom discussions and the journaling activities where students record their hypotheses encourage students to *reflect* upon their developing understanding.

4. Acknowledgment

This is an excerpt from a paper presented at the annual meeting of the National Association for Research on Science Teaching, New Orleans, LA, April 2002. It draws heavily from an article that appeared in the Journal of Research on Science Teaching.

The author would like to express their gratitude to the many teachers and researcher in the Center for Learning Technologies in Urban Schools who contributed to this research. This material is based upon work supported by the National Science Foundation under Grants No. REC-9730377, REC-9720663 and REC-0087751. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Abraham, M. R. (1998). The learning cycle approach as a strategy for instruction in science. In B. J. Fraser & K. G. Tobin (Eds.), International Handbook of Science Education (pp. 513-524). Dordrecht, Netherlands: Kluwer.

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Atwood, R. K., & Atwood, V. A. (1996). Preservice elementary teachers' conceptions of the causes of seasons. Journal of Research in Science Teaching, 33(5), 553-563.

Chi, M. T. H., Peltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Dove, J. (1998). Alternative conceptions about the weather. The School Science Review, 79, 65-69.

Edelson, D. C. (2001). Learning-For-Use: A framework for the design of technology-supported inquiry activities. Journal of Research in Science Teaching, 38(3), 355-385.

Glaser, R. (1992). Expert knowledge and process of thinking. In D. F. Halpern (Ed.), Enhancing Thinking Skills in the Sciences and Mathematics. Hillsdale, NJ: Erlbaum.

Jones, B., Lynch, P., & Reesink, C. (1987). Children's conceptions of the earth, sun, and moon. International Journal of Science Education, 9, 43-53.

Karplus, R., & Thier, H. D. (1967). A new look at elementary school science. Chicago: Rand McNally.

Kolodner, J. L. (1993). Case-based reasoning. San Mateo, CA: Morgan Kaufmann.

Lawson, A. E. (1995). Science teaching and the development of thinking. Belmont, CA: Wadsworth.

Loh, B., Reiser, B. J., Radinsky, J., Edelson, D. C., Gomez, L. M., & Marshall, S. (2001). Developing reflective inquiry practices: A case study of software, the teacher, and students. In K. Crowley & C. Schunn & T. Okada (Eds.), Designing for Science: Implications from Everyday, Classroom, and Professional Settings (pp. 279-323). Mahwah, NJ: Erlbaum.

Philips, W. C. (1991). Earth science misconceptions. Science Teacher, 58(2), 21-23.

Piaget, J. (1929). The Child's Conception of the World. New York: Harcourt Brace.

Renner, J. W., & Stafford, D. G. (1972). Teaching science in the secondary school. New York: Harper & Row.

Russell, T., Bell, D., Longden, K., & McGuigan, L. (1993). Rocks, Soil, and Weather. Primary SPACE Project Research Report. Liverpool, UK: Liverpool University Press.

Schank, R. C. (1982). Dynamic Memory. Cambridge: Cambridge University Press.

Simon, H. A. (1980). Problem solving and educaton. In D. T. Tuma & R. Reif (Eds.), Problem Solving and Education: Issues in Teaching and Research (pp. 81-96). Hilldale, NJ: Erlbaum.

Whitehead, A. H. (1929). The Aims of Education. Cambridge: Cambridge University Press.

Inquiry Learning Through Students' East-West Coast On Line Collaboration about Plate Tectonics [1]

Janice Gobert, Ph.D.
Senior Research Scientist
The Concord Consortium
Concord, MA 01742
jgobert@concord.org

Introduction

This paper will briefly describe a framework, curriculum, and large scale design study involving a total of 1100 middle and high school students from California and Massachusetts who collaborated on-line about plate tectonic activity in their respective location. The students, drawn from demographically diverse schools, collaborated on-line using WISE (Web-based Science Environment, Linn & Hsi, 2000). WISE is an integrated set of software resources to engage students in many types of scientific inquiry, including prompted reflection, electronic discussions, evidence sorting and argument mapping, collaborative search for evidence, collaborative design, and analysis (Linn, 1998; Linn & Hsi, 2000).

The theoretical framework employed in this research draws principally from Model-based Teaching & Learning, put forth in a special issue of the International Journal of Science Education (Gobert & Buckley, 2000). Modelling fits within a current vein of science education which seeks to promote integrated understanding by use of model-based tasks such as, presenting students with models to learn with (Raghavan & Glaser, 1995; White & Frederiksen, 1990), or engaging them in model-building tasks (Gobert, & Clement 1994, 1999; Gobert, 1998; 1999; Penner et al., 1997; Jackson, et al., 1994). Having students critique each others' models, as in the work described here, is a novel approach to both deepening their understanding of the content (so that they may critique others' work) as well as fostering an understanding of what models are and how they are used in science (Gobert et al, 2002). It is believed that having students construct, reason with, and critique each others' models engages them in authentic scientific inquiry, and can significantly impact lifelong learning and scientific literacy (Linn & Muilenberg, 1996) by developing generative knowledge that can be intergrated across science topics and applied to real world problems, such as understanding scientific findings described by the media (Linn, 1999). Since being scientifically literate includes understanding the nature of science, as well as understanding science content and having inquiry skills (Perkins, 1986), the model-based approach here can promote all three types of science knowledge.

Domain Studied

The domain Plate Tectonics was chosen for two reasons. First, it is an excellent domain in which to investigate students' modeling skills because of the important role that model building and causal reasoning play in understanding the hidden mechanisms, e.g., convection underlying

[1] Making Thinking Visible is funded by the National Science Foundation under grant No. REC-9980600 awarded to Janice Gobert. The WISE project is funded by the National Science Foundation by grants awarded to Marcia Linn. Any opinions, findings, and conclusions expressed herein are those of the author and do not necessarily reflect the views of the National Science Foundation.

continental drift, earthquakes, volcanoes, mountain formation, and sea floor spreading[2]. Secondly, it is an excellent context in which to foster students' understanding of science and of models both because there are many excellent models in the domain with which to engage learners in model-based tasks, and theory of plate tectonics is a good example of the dynamic nature of science, how scientific inquiry proceeds, and how a hypothesis can be proposed, discarded, modified, and then redefined.

Plate tectonics, which is typically covered in fifth or sixth grade and then again in eighth or ninth grade is representative of a difficult school science topic. It is difficult to learn for many reasons: 1) the earth's internal layers are outside our direct experience, 2) the size scale and the unobserved processes, e.g., convection, are difficult to understand (Ault, 1984; Gobert & Clement, 1994; 1999), 3) the time scale of geological processes is difficult for people to conceptualize since it surpasses our reference of a human lifetime (Jacobi et al., 1996), and 4) it involves the comprehension and integration of several different types of information, such as, spatial, causal, and dynamic (Gobert & Clement, 1994; 1999).

We designed a curriculum unit called "What's on your plate?" around two WISE pedagogical principles, namely, Make Thinking Visible and Help Students Learn from Each Other.

Make Thinking Visible. Here, we: 1) engage students in drawing tasks to make their models explicit and use these as knowledge artifacts for both model revision as well as peer critique, and 2) provide students with a set of **dynamic, runnable models** of plate tectonic phenomena. Here, students use the runnable prototypes to visualize dynamic, causal, and temporal processes in order to test, critique, and revise their own models. WISE prompts students to justify and explain their changes in order to reify learning. Prompts to be designed include: "What does your new model include that it didn't before?", and "What does your new model describe or explain that it didn't before?"

Help Students Learn From One Another. In terms of **helping students learn from one another,** we engaged students in tasks in which they critiqued their learning partners' models from the opposite coast. We did this to provide students with an opportunity to both think deeply about the domain in order to do the critiques, as well as think about how models are used as tools for communication in science.

The "What's on your plate?" unit the students are engaged in model-based inquiry activities and tasks to learn from one another in the following ways:

1. ***Students' Model Building & Explanation of their Models***. Students were asked to construct in WISE visual models of plate tectonic-related phenomena; that is, each pair of students drew a model of how mountains are formed (East coast only) while students on the West coasts drew models of earthquake or volcanic eruption. Students were then asked to write in WISE a short explanation for their models with the following prompt "Now that you have drawn your model, write an explanation of what happens to each of the layers of the earth when an earthquake erupts (or a mountain is formed, a volcano erupts)". Once students had done these two steps, they posted their models and explanations for their learning partners on the opposite coast.

[2] The theory of plate tectonics states that the outer layer of the earth (the crust) is broken up into slabs (the plates) which move on the partially molten layer of the earth (the mantle) due to the convective movement of hot magma in the mantle (Feather, Snyder, & Hesser, 1995; Plummer & McGeary, 1996).

2.Students' Evaluation and Critique of the Learning Partners' Models. Students read two pieces of text in WISE called "What is a Scientific Model? And "How to evaluate a model?" in order to give them some basic knowledge with which to evaluate their leaning partners' models. Then students were prompted to critique learning partners' models using prompts that were presented in WISE. The prompts include:

- 1. Are the most important features in terms of what causes this geologic process depicted in this model?
- 2. Would this model be useful to teach someone who had never studied this geologic process before?
- 3. What important features are included in this model? Explain why you gave the model this rating.
- 4. What do you think should be added to this model in order to make it better for someone who had never studied this geologic process before?

These prompts were designed to focus students' thinking about models in two general ways: the causal mechanisms/processes depicted (items 1 and 3), and the model as a communication tool to learn or reason with (items 2,and 4). Prompts similar to the latter have been successful in getting students to generate rich explanations (Gobert, 1997b), and it was believed that they might be successful here as well in getting students to think about how useful a model is as a tool for communication purposes. Once students discussed the evaluation with their in class partner (computer partner), they then posted their evaluation for their opposite coast learning partners to evaluate.

*3**. Students'Model Revision&Justification***. Students read the evaluation that was written and posted by their learning partners on the opposite coast. They were the asked to revise their models based on the critique from their learning partners as well as the content knowledge they had learned from the unit (the model-based content activities will be discussed next). They were also asked to write a revised explanation for their new models. Lastly, here students were asked to justify their changes to their models in WISE in order to engage students in reflection about how their understanding had changed. Prompts here include:

- I changed my original model of.... because it did not explain or include...."
- "My model now includes or helps explain…"
- "My model is now more useful for someone to learn from because it now includes…."
- "I revised this on the basis of my learning partners' critique in the following ways….
- **"I revised this on the basis of the activities in these WISE units….. ".**

4. ***Geology Websites***. As part of the unit students do an on-line field trip and are guided to visit multiple USGS websites with current data in order to the differences between the coasts in terms of their mountains, volcanoes, and earthquakes. After each "site visit", students write a relfection note for their learning partners on the oppoiste coast about what they have learned about earthquakes, volcanoes, and mountains on their coast. This reflection note is posted for the learning partners to read and reflect on in terms of how the data observed differ from that of their own coast.

Students also visit a Plate boundaries website in order to speculate about how the location, frequency, and magnitude of geological events (mountains, earthquakes, and volcanoes) "observed in Activity 2 are related to plate boundaries in the earth's crust. After visiting the plate boundaries website, students are asked to write a Reflection Note with the following prompt: Write one (or two) question(s) you have about plate boundaries or plate movement that will help you better understand why the geologic processes on the West and East coasts are different. Students revisit these questions in a Discussion Forum later in the unit.

*5. **Dynamic-runnable models***. These models were designed in line with previous research which has shown that visualization facilitates the understanding of dynamic phenomena (Monaghan & Clement, 1995) and that middle and high school students can understand rich dynamic concepts if provided with the appropriate scaffolds and tools (Jackson, et al., 1994; Ploger & DellaVedova, 1999; Frederiksen, White, & Gutwill, 1999).

Students view and read about the different types of plate boundaries, namely, collisional, divergent, convergent, and transform boundaries in order to begin to think about how the location of and type of plate boundary are related to geological occurences on the earth's crust. Students reify their learning by writing reflection notes about what types of geological events are typical of specific types of plate boundaries.

Students also visit a model of mantle convection which is accompanied by a text which scaffolds their understanding of the dynamic and causal features of the model by directing their processing of the causal and dynamic information in the model as it "runs". Students write a reflection note to explain how processes inside the earth relate to plate movement.

Lastly, students visit a series of dynamic models which depict different types of plate convergence, namely, oceanic-oceanic convergence, oceanic-continental convergence, and continental-continental convergence. Again, students' understanding is scaffolded via a text which directs their processing of the causal and dynamic information in each model as it "runs.

To view "What's on your Plate?"—you can either start an account for yourself, or go to an account that has already been set up (but it may have others' work in it that cannot be changed) on the computer provided. To get your own account for this unit, go to the WISE new student registration page http://wise.berkeley.edu/pages/newStudent.php Fill in with your: First name, Last name, for PERIOD, put 10, enter a password of your choice, for your student registration code, type SZP87G. Click on "go to the student portal."Or to go to an account that is already set up, go to wise.berkeley.edu, click on Member entrance, and for login enter "AnonyM1" and "try" as your password.Click on "Plate Tectonics: What's on Your Plate?".

Summary

This research utilized a state-of the art science learning environment, WISE, to promote deep learning of subject-matter in plate tectonics and model-based inquiry skills involving model critiquing and revision. Data from this large scal research project has yielded significant learning gains both in terms of students' content knowledge of Plate Tectonics as well as their understanding of the nature of models in science. As such, from these data, it appears that model-based tasks, students' critquing each others models, and students' collaboration are useful approaches to promoting learning in this domain and scientific literacy in general.

REFERENCES

Feather, R., Snyder, S., & Hesser, D. (1995). *Earth Science*. Westerville, Ohio: Glencoe/McGraw-Hill.

Frederiksen, J., White, B., & Gutwill, J. (1999). Dynamic mental models in learning science: The importance of constructing derivational linkages among models. To appear to Journal of Research in Science Teaching.

Gobert, J. (1997b). *Summarizing, Explaining, and Diagramming: The Differential Effects on Text-Base Representations and Mental Models.* Presented at the Nineteenth Annual Meeting of the Cognitive Science Society. Stanford University, August 7-10. Palo Alto, CA.

Gobert, J. (2000). A typology of models for plate tectonics: Inferential power and barriers to understanding. *International Journal of Science Education, 22(9),* 937-977.

Gobert, J. & Buckley, B. (2000). Special issue editorial: Introduction to model-based teaching and learning. *International Journal of Science Education*, *22(9),* 891-894.

Gobert, J. & Clement, J. (1994). *Promoting causal model construction in science through student-generated diagrams*. Presented at the Annual Meeting of the American Educational Research Association, April 4-8. New Orleans, LA.

Gobert, J. & Clement, J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching,* 36(1), 39-53.

Gobert, J. Slotta, J. & Pallant, A., Nagy, S. & Targum, E. (2002). A WISE Inquiry Project for Students' East-West Coast Collaboration, Presented at the Annual Meeting of the American Educational Research Association, New Orleans, LO, April 1-5.

Jackson, S., Stratford, S., Krajcik, J., & Soloway, E. (1994). Making dynamic modeling accessible to pre-college science students. *Interactive Learning Environments*, 4(3), 233-257.

Jacobi, D., Bergeron, A., & Malvesy, T. (1996). The popularization of plate tectonics: presenting the concepts of dynamics and time, *Public Understanding in Science,* 5, 75-100.

Linn, M. C. (1998). *Supporting teachers and encouraging lifelong learning: A web-based integrated science environment (WISE).* Proposal funded by the National Science Foundation.

Linn, M. C. (1999). Designing the knowledge integration environment: The partnership inquiry process. Created for *International Journal of Science Education*.

Linn, M. C., & Hsi, S. (2000). *Computers, Teachers, Peers: Science Learning Partners.* Hillsdale, NJ: Erlbaum.

Linn, M. C., & Muilenberg, L. (1996). Creating lifelong science learners: What models form a firm foundation? *Educational Researcher,* 25 (5), 18-24.

Penner, D. E., Giles, N. D., Lehrer, R., Schauble, L. (1997). Building functional models: designing an elbow. *Journal of Research in Science Teaching,* 34(2), 125-143.

Perkins, D. (1986). *Knowledge as design*. Hillsdale, NJ: Erlbaum.

Ploger, D. and Della Vedova, T. (1999). Dynamic Charts in the Elementary Classroom. *Learning and Leading with Technology* 26, 38-41.

Plummer, C. & McGeary, D. (1996). *Physical Geology*. Wm. C. Brown Publishers.

Raghavan, K. & Glaser, R. (1995). Model-based analysis and reasoning in science: The MARS curriculum. *Science Education,* 79, 37-61.

White, B. & Frederiksen, J. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, 24, 99-157.

Learning Chemistry by Design

Loretta L. Jones

Chemistry can be a difficult subject to teach because not only is it a molecular science in which many of the concepts and processes are not visible to the eye, but students often enter the course with negative feelings toward the subject. A number of strategies have been developed to address these problems, including the following:

1. Visualizing the molecular level of matter/moving among representations.
2. Presenting content within an interesting context
3. Involving students in design activities
4. Promoting cooperative learning and active learning in the classroom
5. Addressing student misconceptions
6. Building problem solving skills
7. Engaging students in inquiry-based learning and authentic science activities

In the *ChemDiscovery* (formerly *ChemQuest*) curriculum (Agapova, Jones, & Ushakov, 2002), we designed a pedagogy for a technology-based curriculum that would incorporate all of these strategies. *ChemDiscovery* is an inquiry-oriented environment for learning chemistry (along with some topics in physics, biology, astronomy, geology, and environmental science) through the design of a virtual world. The curriculum uses linked web pages organized around a series of active learning scenarios with multiple entry points.

While completing the 10 Quests (projects) that comprise the curriculum, students design a virtual world. Each Quest has two context shells, Design of the Universe and Living in the Universe, that allow students to enter the world of chemistry from environmental, scientific, and humanistic perspectives. Once inside a Quest, students work in self-directed pairs, choosing their own starting point for each Quest and, together with their teacher, designing individual pathways through the learning environment.

Because introductory science courses are usually focused primarily on analysis and categorization, design activities are not ordinarily encountered until more advanced courses (Jones, 1999). Yet design processes play important roles in our daily lives as well as in science and engineering. The design processes at the heart of *ChemDiscovery's* pedagogy involve students in understanding needs and responsibilities; selecting raw materials; planning, designing, and predicting the properties of objects (nuclei, atoms, molecules, crystals, and larger systems); constructing the objects; and evaluating their predictions.

In addition to its motivational role, context is important for reinforcing learning. An interactive field guide shows students how the principles they are learning relate to their lives as they map their communities---not only their roads and houses, but maps of power grids, the route water takes as it flows from its source through their municipal water system, and the rock formations underlying their homes.

New types of assessments had to be developed for this approach. The assessments require students to design, model, and construct new systems. They also evaluate student ability to use comprehensive databases of information---the same type of databases used by professional scientists.

We measured the impact of *ChemDiscovery* on classroom interactions by using a series of systematic observations combined with field notes (Schoenfeld-Tacher, Madden, Pentecost, Mecklin, & Jones,1999). We found that when *ChemDiscovery* was used in the chemistry classrooms of two high school teachers, the focus of classroom interactions shifted from a teacher-led lecture format to one in which students spent a significant amount of time working cooperatively in pairs. Teachers spent more time acting as facilitators and resources rather than as lecturers.

Because chemistry deals with the molecular level of matter, much of *ChemDiscovery* deals with modeling atoms and molecules. This work has been influenced by the research that has been conducted on learning from models and simulations (Jones, Jordan, & Stillings, 2001). A real concern of this research is how what we present to students affects their understanding and their ability to form accurate mental models. Representations of all types, whether physical models or animated simulations, can induce misconceptions.

The issues that I am struggling with and that I hope to confront with others at this conference are (1) investigating the impact of various kinds of representations on student learning, (2) finding out what characteristics make simulations and modeling activities appropriate and effective, (3) redesigning instruction to help students move from fragmented knowledge to coherent mental models, (4) finding a balance between the simplification of material that helps beginners grasp complex topics and the truth as scientists know it.

References

Agapova, O., Jones, L, & Ushakov, A. (2002). *ChemDiscovery*, Dubuque, IA: Kendall-Hunt (See http://www.chemdiscovery.com).

Jones, L. L. (1999). Learning Chemistry through Design and Construction, *Uniserve News*, 14, November, 1999 (http://science.uniserve.edu.au/newsletter/vol14/jones.html).

Jones, L, Jordan, K., and Stillings, N. (2001). Molecular Visualization in Science Education: Report from the Molecular Visualization in Science Education Workshop, http://pro3.chem.pitt.edu/workshop/workshop_report_180701.pdf

Schoenfeld-Tacher, R.. Madden, S. Pentecost, T. Mecklin, C. and Jones, L. (1999). A Systematic Comparison of Technology-Based and Traditional High School Chemistry Classrooms, *National Meeting of the National Association for Research in Science Teaching*, Boston, MA, March 30, 1999.

My Top Ten Topics in Geoscience Education Research

Kim Kastens, June 15, 2002

for Conference on "Bringing Research on Learning to the Geosciences"

Many of the most exciting advances within Geosciences are now coming from the study of interfaces within the Earth System, for example the ocean-atmosphere interface, the core-mantle boundary, the coastal zone, outposts of the biosphere within the solid earth. One interface remains understudied: the interface between the Earth and the human mind. How do geoscientists and geoscience students “get our minds around” some aspect of the Earth or environment? I am looking forward to thinking about this at the upcoming conference.

Here is my personal list of topics that I think would benefit from research on learning,, and from dissemination of existing research results to geoscience curriculum developers and front-line geoscience educators.

(1) How do people learn to comprehend vast expanses of space, using a mind that evolved to cope with spaces one could see across or walk across?

(2) How do people learn to comprehend vast expanses of time, using a mind that evolved to cope with time scales bounded by the human lifespan?

(3) How do people learn to comprehend, and mentally manipulate, objects, processes or phenomena in three-dimensions (or four dimensions including time)?

(4) How do people learn to make and interpreting spatial representations, including maps, cross-sections, block-diagrams, stereonets, etc.

(5) How do people comprehend, and learn to make predictions about, systems in which there are multiple interacting causality chains, and/or circular causality chains (feedback loops)

(6) Most geoscientists are trained to think about the past. But society is now asking us questions about the future, about earthquakes, global warming, and so on. What mental processes are involved in thinking rigorously about the future, and how can we foster this ability in our students?

(7) Most students today learn most things from human beings or from human artifacts such as books, computer programs, videos, etc. For a student who has grown up in this situation, what does it take to enable that student to learn directly from Nature, by direct observation, in the field, of rocks, organisms, etc.?

(8) How can we foster students' ability to learn from data?

(9) How can we foster students' ability to learn from models, either by building models or manipulating models made by others?

(10) How can we foster equation literacy in our students, i.e. the ability to translate from a holistic understanding of a situation into a quantitative equation-based description, and vice versa.

GEOSCIENCE LEARNING RESEARCH IN UK HIGHER EDUCATION

Dr Helen L. King

Introduction

This paper discusses a collaborative programme of learning research funded by the Learning and Teaching Support Network Subject Centre for Geography, Earth & Environmental Sciences (LTSN-GEES), and a cross-disciplinary project to explore the linkage between teaching and research funded by the LTSN Generic Centre. In order to place this discussion in context, a brief background to the current climate in UK Higher Education is also provided.

Background – the current climate in UK Higher Education (HE)

HE is going through an extensive period of change in the UK. Approximately five years ago students were required to pay their own tuition fees, and maintenance grants were gradually phased out to be replaced by loans. As a result, students are becoming more demanding of the quality of their learning experience and are also looking to undertake more vocational courses. This has contributed towards a decline in student recruitment in the sciences, and this, together with the effects of a reduction of geography and geology in the K-12 curriculum has added to the problems faced by the geosciences in HE.

As in many countries, there is a conflict of interest between teaching and research in HE. Research is funded, in part, by the HE Funding Councils and the level of this funding is dependent on departments' performance in the Research Assessment Exercise (RAE). Consequently, smaller departments and those that have more of a teaching focus are less likely to obtain enough funding to sustain their activities. Recently, several geology departments have had to close and others are being merged with other disciplines. Not only does this trend reduce the amount of geoscience research but affects the provision of geoscience teaching in HE. Until recently, learning research was only acknowledged by the RAE where it was conducted in Departments of Education and, consequently, there was no formal motivation to undertake discipline-based learning research. This has recently changed but, in order for departments to optimise their RAE results, individuals are still discouraged from undertaking any research outside their own discipline area.

"The Government has said that 50 per cent of young people should have the opportunity to benefit from higher education by 2010. This target has both an economic and a social purpose. More graduates are needed to enable the UK to sustain and develop a knowledge economy able to compete globally. And fair access for those from disadvantaged backgrounds to all forms of education, including higher education, is an essential part of addressing social exclusion." (HEFCE, 2002). The following areas of national priority have been identified:

- Widening participation,
- Ensuring fair access to HE (including students with special education needs and disabilities),
- Maintaining and improving retention rates,
- Enhancing employability of graduates, and
- Encouraging and disseminating good and innovative practice in support of high quality learning and teaching.

It is within this climate that the Learning and Teaching Support Network (LTSN) was launched in January 2000 and funded by the four HE councils in the UK (one each for England, Scotland, Wales and Northern Ireland). The LTSN consists of 24 Subject Centres which offer subject-specific expertise and information, a Generic Centre which addresses learning and teaching issues that cross subject boundaries, and a Technologies Centre and Technology for Disabilities Information Service. The LTSN aims to:

- Promote and transfer high quality learning and teaching practices in all subject disciplines,
- Provide a 'one-stop-shop' of learning and teaching resources and information for the whole HE community,
- Develop and support networks of practitioners.

The LTSN Subject Centre for Geography, Earth & Environmental Sciences (LTSN-GEES) is based at the University of Plymouth. Further information can be found at:
LTSN: http://www.ltsn.ac.uk LTSN-GEES: http://www.gees.ac.uk

Learning Research into the Educational Effectiveness of Fieldwork

Until recently, discipline-based learning research has mainly been undertaken only by individuals on a relatively small scale, for their own interest, fuelled by their enthusiasm for their discipline and their commitment to the learning experience of their students, and with little or no external recognition or reward. However, with the increasing need for high quality learning and teaching, the declining recruitment in geosciences despite the Government's widening participation agenda, and a relative reduction in resources to support learning and teaching, learning research is starting to become recognised as a necessary aspect of scholarship and professionalism in UK HE.

In 2001, LTSN-GEES obtained £62, 000 to fund a programme of work to develop discipline-based learning research capacity through supporting collaborative research into the educational effectiveness of fieldwork in Geography, Earth and Environmental Sciences. The programme began in June 2001 and continues until April 2003. In addition to the LTSN-GEES team, it involves a total of around 30 academics from over 15 different geoscience departments throughout the UK, and two advisors with experience and expertise in generic learning research.

The programme arose as a result of four main factors:
- The need to develop the capacity of the discipline-based communities to undertake learning research,
- The central role of fieldwork to student learning in Geography, Earth and Environmental Sciences,
- The fact that fieldwork is, as yet, little theorised,
- Fieldwork lends itself well to a range of learning research methodologies.

To support these needs, the programme has the following aims:
- To build capacity for undertaking geoscience learning research,
- To undertake research into learning and teaching issues associated with fieldwork,
- To disseminate and embed the results of this research,
- To disseminate the research methodologies.

The programme consists of 5 mini-projects (each funded with up to £5000) supported by a series of staff development workshops. The first workshop, in June 2001, attracted nearly 60 interested individuals and helped to define the topics for the projects. Once the project teams were established, a two day event was held in September 2001 to discuss research methodologies and to refine the project plans. LTSN-GEES was extremely fortunate to have the support of Prof John Carpenter from the University of South Carolina to help out with event and to share his experience and expertise. Once the mini-projects were underway, a third workshop was held in May 2002 to provide an opportunity to discuss the issues involved in qualitative data analysis. A final event will be held in January 2003, to bring together all the project teams and to discuss their findings. The research will be written up in relevant journals and presented at the 2003 GeoSciEd IV conference in Calgary, as well as being outlined in LTSN-GEES's publication, Planet (available to download free from our web-site). The five projects are:

1. The Role of Fieldwork in the Curriculum
Initial research question: To what extent does fieldwork relate to Biggs' theory of constructive alignment (Biggs, 1999)?

2. The Impact on the Learning and Teaching Experience of the Removal of Fieldwork from Academic Programmes
Initial research question: At particular institutions, did the absence of fieldwork [due to the 2001 Foot & Mouth epidemic] impact on a) the module grades, b) on students perceptions of the learning environment, c) on staff perceptions of the learning environment?

3. Fieldwork is good? – the student view
Initial research question: What is the effect of fieldwork on the students' affective domains [processes that deal with emotions, feelings and values](Kern & Carpenter, 1984)?

4. Fieldwork Education and Technology
Initial research question: What is the relationship between C&IT [communication and information technology] and fieldwork education?

5. Learning to do Pedagogic Research
Initial research question: How far and in what ways has the programme succeeded in developing capacity to undertake learning research?

All the projects are currently in the data gathering and analysis part of the process, so, unfortunately, few results can be reported at this stage. However, the programme has a very positive feel and the participants seem to be enjoying being involved. For the geosciences community, the initial results from the second project listed above, shortly to be submitted to the Journal of Geography in Higher Education (JGHE), are extremely encouraging as illustrated by extracts from the draft paper copied:

"Internationally, fieldwork is generally seen as intrinsic to the very nature of geographical education. However, the educational objectives and actual outcomes of fieldwork have rarely been examined. During 2001 in the UK, fieldwork was withdrawn from many university degree programmes as Foot and Mouth Disease led to restrictions on access to the countryside. This restriction provided an unexpected opportunity to assess student perceptions of fieldwork in the light of its absence and to review those alternative learning strategies which were put in its place (where appropriate). To this end, Nominal Group Technique (NGT) was applied to five groups of students from five separate UK Universities to obtain information on the groups' perceptions of the value of fieldwork. NGT elicited almost 300 responses from 33 students representing a high level of group consensus on the issues involved.

"Results demonstrate that student evaluation of fieldwork supports many of the academic / practitioner views of the role of fieldwork in student learning. As such, the educational objectives of fieldwork (Gold et al, 1991) are being achieved from these students' perspectives. Students perceive the greatest strengths of fieldwork to lie in the experience of reality, enhancing understanding of the topic, using specific equipment and techniques, developing holistic and transferable skills and in developing interpersonal skills, both between peers as well as between students and lecturers. However, this paper has also shown that fieldwork may not provide a completely positive learning experience for students. Issues which most concern students in this study include the time spent on fieldwork, teaching / delivery (e.g. missing lectures), increases in assessment and workloads, concerns about understanding of subject-specific knowledge (e.g. less detailed notes complicating understanding), technical concerns to do with equipment, and financial concerns (e.g. too expensive). However, the results in this study show that the positive attributes of fieldwork far exceed the negative attributes reported by the students' groups, demonstrating fundamentally that overall fieldwork is a much valued student learning experience." (Fuller, Gaskin & Scott, in preparation).

Linking Teaching and Research

LTSN-GEES is involved in a project funded by the LTSN Generic Centre to explore the linkages between teaching and research. LTSN-GEES' contribution aims to identify, record and disseminate case studies of the way individuals, course teams and departments within Geography, Earth and Environmental Sciences enhance the learning of their students by developing the links with their research and to promote ways in which individuals and departments can maximise the benefits for students of these linkages. The project began in March 2002 and case studies are currently being identified.

It is recognised that students may benefit from research in a variety of ways including, where:

- The content of courses in informed by staff research,
- Students learn about research methods,
- Teaching methods adopt a research-based approach, e.g. through problem-based learning,
- They undertake their own research projects, whether individually or in teams,
- They participate in staff research projects as subject, as in, for example, perception studies,
- They assist staff with their research projects,
- Staff undertake learning research which benefits the quality of their teaching.

Research in this context is interpreted widely to include RAE-level research, consultancy for clients, and action research aimed primarily at improving practice. It is recognised that there are also potential negative impacts from staff involvement in research, such as staff absences and lower priority being given to teaching. For the benefits to be maximised and the disadvantages to be minimised, the relationship between teaching and research needs to be effectively managed.

~~/~~

Dr Helen L. King
Manager, LTSN-GEES
University of Plymouth, UK
Email: h.king@plymouth.ac.uk
Web-site: http://www.gees.ac.uk

References

Biggs, J. 1999. *Teaching for Quality Learning at University*. Society for Research into Higher Education and Open University Press pp 11-32

Fuller, I., S. Gaskin & I. Scott. In preparation. An assessment of student perception of geography and environmental science fieldwork in the light of restricted access to the field, caused by Foot and Mouth Disease in the UK in 2001.

Gold, J.R., A. Jenkins, R. Lee, J. Monk, J. Riley, I.D.H. Shepherd & D.J.Unwin. 1991. *Teaching Geography in Higher Education*. Blackwell, Oxford.

Higher Education Funding Council for England (HEFCE), 2002. *Teaching Quality Enhancement Fund: Funding arrangements 2002-2003 to 2004-2005.* Available at http://www.hefce.ac.uk/

Kern, E. & J.Carpenter. 1984. Enhancement of student values, interests and attitudes in Earth Science through a field-orientated approach. *Journal of Geological Education*, Vol. 21 No.3, pp313-332.

Geoscience Education: Building a Research Paradigm*

Julie C. Libarkin

Science Education Department
Harvard-Smithsonian Center for Astrophysics
60 Garden St. MS-71
Cambridge MA 02140

Prepared for: "Bringing Research on Learning to the Geosciences", July 8-10, 2002, Racine, Wisconsin

* Some excerpts and ideas borrowed from Research Methodologies in Science Education, column prepared for the Journal of Geoscience Education (May 2001 through present), co-authored by J. Kurdziel and J. Libarkin.

Synopsis. Science education is currently undergoing a transformation, especially at the collegiate level. Teaching reforms are being implemented across the board by faculty, departments, and institutional policy committees interested in improving upon the scientific literacy of the voting public. With this transformation, education research within scientific disciplines is finding new respect, and more and more science faculty are devoting themselves to studying the classroom environment. Traditionally, however, scientific and education disciplines have remained unconnected, with limited transfer of information and new knowledge between the two fields. As a result, disciplinary-specific science education research is often far removed from research paradigms that have long been established in the behavioral sciences. Similarly, the field of science education, while grounded in established methodologies, rarely has a direct impact on the teaching and learning actually occurring in classrooms. Indeed, this disparity exists in our disciplinary journals; researchers writing for the Journal of Research in Science Teaching rarely engage in science classroom instruction themselves, especially at the college level, and studies presented in the Journal of Geoscience Education are often presented by teaching faculty, although with little focus on the suitability of research design. Geoscience education today is primarily embedded in classroom, or localized, research, and it is to methodology and structure that researchers must now turn their attention. To ensure that research findings are useful and applicable outside of the immediately studied environment, the geoscience community must come to a consensus about the form education research should take, the methods for acquiring and analyzing data, and the means for establishing research validity. With this consensus, education researchers can ultimately implement a common research paradigm that will open the lines of communication across all disciplines.

Geoscience education has been an active focus of research within the education community for at least fifty years, as evidenced by the long publication record of a journal specifically designed to document and promote geoscience education. The Journal of Geoscience Education (previously known as the Journal of Geological Education) has provided a venue for the presentation of innovative ideas in instruction and research findings since 1951. In the ensuing half a century, interest in geoscience education has fluctuated, although the past decade has seen a resurgence in policy interest and funding for educational initiatives, especially at a national level. Faculty are now being asked to focus on the successes, and failures, of their teaching endeavors.

Faculty expectations for student learning usually cover a wide range of possibilities. In general, a number of student outcomes can result from engagement in a course, including content knowledge acquisition, skills development, changes in attitudes/values/beliefs, and long-term behavioral outcomes (Ewell, 1987). At the introductory and non-major levels, goals for geosciences courses can vary widely, although they typically include several of the following: deep conceptual understanding of fundamental principles, improved understanding of the processes of science, improved attitudes toward science, and skills development (critical thinking, synthesis, and communication). Apart from attitudinal shifts, many faculty would agree that beginning geosciences courses should provide students with the knowledge and skills necessary for complex decision-making about their own interactions with the Earth. Upper division courses are usually focused on learning the "language" of geology, acquiring the higher order thinking skills necessary for interpreting geologic data, and developing skills for evaluating the scientific literature. What needs to be done to further our understanding of how students learn in the geosciences? What techniques are most effective at facilitating learning and do additional factors influence learning? What protocols should be established to ensure valid and reliable research endeavors in the future? *Ultimately, how do we know if students have learned?*

University professors typically have many theories about the most effective methods for promoting student learning. Some faculty swear by traditional lecture, while others insist that alternative techniques must be used if students are going to become scientifically literate. Phrases like "minds-on, hands-on", "inquiry", "collaborative and cooperative learning", "instructional technology", "problem solving instruction", and "peer-peer interactions" are being used increasingly by university faculty to describe their latest innovations in the classroom. The number of education sessions at annual geology and geophysics meetings is increasing, resulting in wide dissemination of a variety of teaching techniques, curricula, technologies, and course structures. In fact, professional geoscience organizations are actively sponsoring workshops advancing these reforms. For instance, Geological Society of America and American Geophysical Union have both sponsored workshops on collaborative learning strategies, technology, and/or inquiry-based instruction in recent years. Although these alternative methodologies are adopted from studies that have shown their effectiveness at the K-12 level (e.g., Johnson and Johnson, 1994) or in other disciplines (e.g., Zietsman and Hewson, 1986), relatively little research has addressed the effectiveness of these techniques for college-level education, and very few studies have focused on geoscience education at any grade level. Those studies that do consider geoscience education are primarily grounded in classroom, rather than educational, research (Angelo and Cross, 1993).

A survey of the articles contained within the Journal of Geoscience Education exemplifies the predominance of classroom research in geoscience education. Classroom research refers to practices that assist a teacher in determining student learning gains and instructional effects.

Student grades, exams, in-class writing assignments, and even focus groups all provide some insight into course outcomes. Teachers can use classroom assessment to inform their own instruction, but are usually not able to generalize their findings to other classrooms or educational settings. Educational assessment uses similar data types; however, the validity and reliability of the research design must also be established. For example, classroom exams are usually created by instructors based on their individual beliefs about what students should have learned and how questions should be asked. Outside opinion is rarely consulted, and qualitative and/or quantitative methods for validation are never required to prove that an exam score is a reliable measure of student learning. The development of a test for educational assessment, however, does require strict adherence to validity and reliability guidelines. This added effort helps ensure that educational research findings can be generalized to multiple educational environments.

The Journal of Geoscience Education issues from 1991-2001 contain 610 articles. Most of these articles persuasively describe a course, curriculum, program, or useful classroom technique, with about 30% of the articles explicitly mentioning the effectiveness of the educational initiative (Fig. 2a). Although many of the studies presented in this journal describe innovative and potentially effective teaching methods, it is difficult to determine the extent of student outcomes using only classroom assessment methodologies. In fact, many authors acknowledge this difficulty, stating about their own research:

> *This new approach seems more successful, although it is difficult to evaluate...*[1]
> *How can we evaluate the success or failure of a course such as this one?* [2]
> *It has not been feasible to demonstrate statistically the effect (if any) of...* [3]

In all only 19 articles use a quantitative instrument or established qualitative techniques, with just eight studies using a previously published (4 articles) or partially-normed (4 articles) assessment tool (Fig. 2). We as a community must encourage geoscience education researchers to move beyond classroom research and begin to adopt methodologies that have long been used to study human interactions in the fields of education, psychology, and nursing. This adaptation of research methodologies will help ensure that future studies will be applicable to a wide range of learning environments.

Personal Perspective. I would argue that a discussion of learning in the geosciences will have only modest impact on the geoscience education community at large until we address the twin issues of validity and reliability in research. Geoscientists actively engaged in education research have by and large focused on classroom research and overlooked existing paradigms for educational research methodologies. Although classroom research is actively pursued by many geoscience educators to inform their own classroom instruction, few of these studies have evolved into educational research that can be used to inform the community at large. In particular, a variety of methodologies are currently being used to acquire, analyze, interpret, and report data. Without a consensus within the community as to what does and does not constitute a valid and reliable research paradigm, research into learning in the geosciences can never evolve beyond classroom specific conclusions. In particular, we as educators should be concerned with the generalizability of our research findings to all classrooms and learning environments.

---

[1] Wiswall and Srogi, 1995; [2] Picard, 1993; [3] Lutz and Srogi, 2000

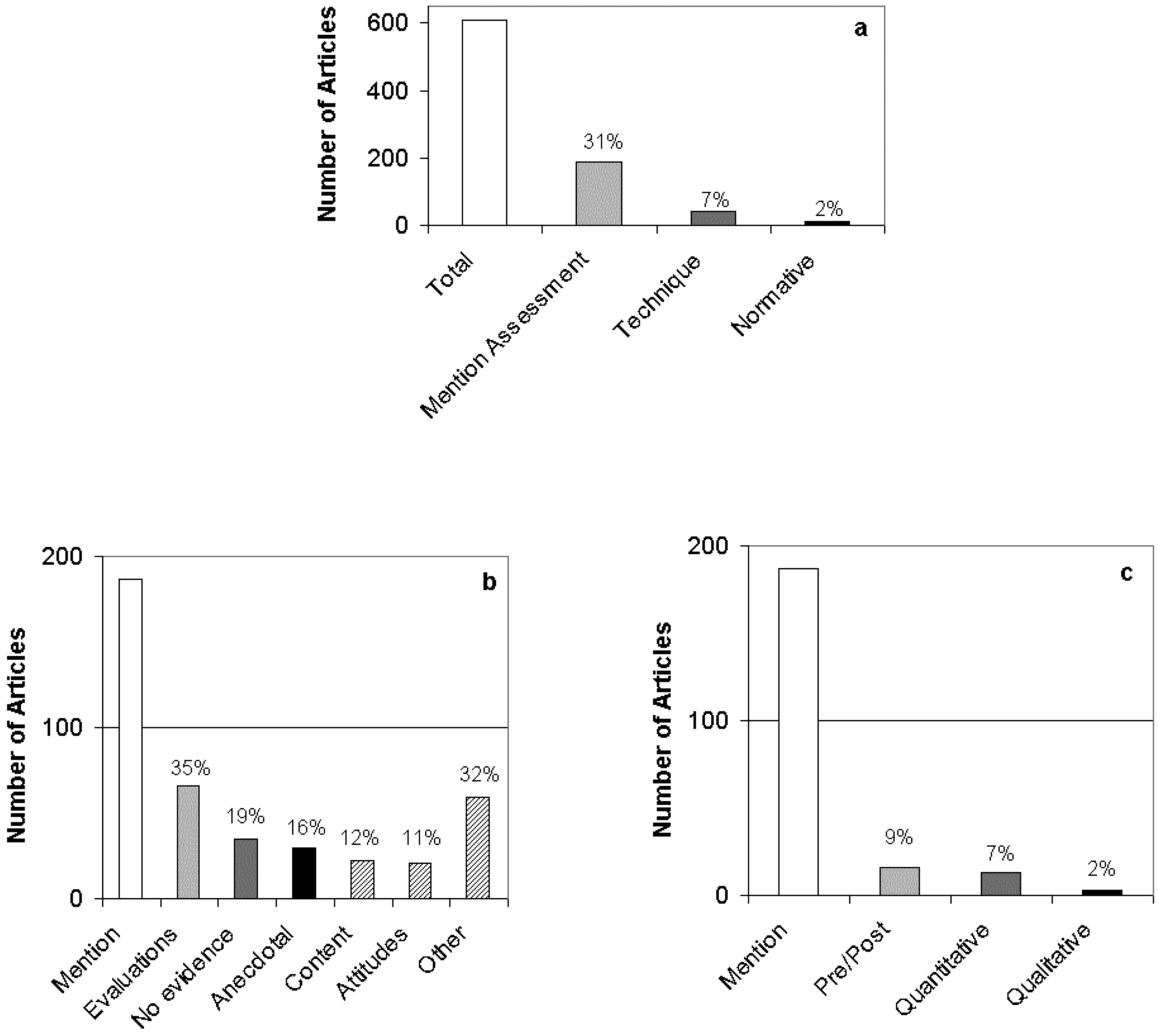


Figure 1. Review of Assessment Practice, Journal of Geoscience Education, 1991-March 2001. Percentages are with reference to total number of articles (**a**) and those articles which explicitly mention assessment or effectiveness (**b**, **c**). **a.** 610 articles appear, with 187 referring to assessment or effectiveness in some form. Of these, only 10% use established educational research methods and only 2% (11 articles) use some type of normative measure that would allow comparison with other studies. **b.** 35% of the studies provide no evidence to support claims of effectiveness, or use anecdotal evidence only. Remaining studies rely overwhelmingly upon course or activity evaluations, and/or one or more of the following: content knowledge, attitudes, faculty-generated survey instruments, grades, participant self-assessments, interviews, focus groups, attendance rates, and student work. **c.** Of the 187 studies that mention assessment, only 10% utilize pre/ post-testing methods and/or quantitative statistical techniques. Only three studies make use of established qualitative or quasi-qualitative methods.

References

Angelo, T.A., and Cross, K.P., 1993, Classroom assessment techniques: A handbook for college teachers: San Fransisco, Jossey-Bass, pp. 427.

Ewell, P.T., 1987, Establishing a campus-based assessment program, *in* Halpern, D.F., editor, Student outcomes assessment: What institutions stand to gain: Jossey-Bass, San Francisco, p. 9-24.

Johnson, D.W., and Johnson, R.T., 1994, Learning Together and Alone. Cooperative, Competitive, and Individualistic Learning (4^{th} Ed.): Massachusetts, Allyn and Bacon, pp. 292.

Lutz, T.M, and Srogi, L., 2000, Role of a shadow course in improving the mathematics skills of geoscience majors, Journal of Geoscience Education, v.48, p.474-478.

Picard, M., 1993, Teaching the unknown in geology, Journal of Geological Education, v. 41, p.204-207.

Wiswall, C.G., and Srogi, L., 1995, Using writing in small groups to enhance learning, Journal of Geological Education, v. 43, p.334-340.

Zietsman, A.I., and Hewson, P.W., 1986, Effects of instruction using microcomputer simulations and conceptual change learning strategies in science learning, Journal of Research in Science Teaching, v. 23, p. 27-39.

THE LEARNING SCIENCES AND GEOSCIENCE

Marcia C. Linn
4523 Tolman Hall #1670
Graduate School of Education, University of California at Berkeley
Berkeley, CA 94720-1670
e-mail: mclinn@socrates.berkeley.edu
fax: (510) 643-0520

This material is based upon work supported by the National Science Foundation under grant Nos. 9873180, 9805420, 0087832, and 9720384. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

This paper was partially prepared while Linn was a Fellow at the Center for Advanced Study in the Behavioral Sciences with support from the Spencer Foundation, Grant 2000 00206.

I particularly appreciate the help and suggestions from the Web-based Integrated Science Environment Research Group. This research reflects the contributions of the WISE Partnership and builds on the work of individuals involved in the Computer as Learning Partner and the Knowledge Integration Environment Project. Thanks also to Lisa Safley, Scott, Hsu, and David Crowell for help in the production of this manuscript

THE LEARNING SCIENCES AND GEOSCIENCE

Too often, science instruction in the United States results in memorized details rather than linked or connected ideas. High stakes assessments can inadvertently reinforce this form of instruction with multiple choice items on isolated topics. In contrast students who conduct inquiry projects develop more cohesive, robust and coherent accounts of complex science and continue to develop their ideas after completing science classes. These students perform well on tests that require them to integrate their ideas into coherent arguments. To capture the excitement of science and stimulate knowledge integration by students, teachers, and district leaders, we advocate inquiry instruction which is well suited to learning in the geosciences.

Recent analyses of American textbooks conclude that students study "heavy books—light on learning" (AAAS, 1999) and that the United States curriculum is "a mile wide and an inch deep, with more topics covered than most other nations, but less time devoted to making sense of science" (Schmidt, Raisen, Britton, Bianchi, & Wolfe, 1997). As research on memory would predict (e.g., Bjork, 1994, 1999; Baddeley & Longman, 1978), this form of instruction leads to little cumulative learning and rapid forgetting. National assessments (O'Sullivan, Reese, & Mazzeo, 1997), international comparisons (Schmidt, McKnight, & Raizen, 1997), and state assessments (Blank, 2000) report dismal American performance in science. Most teachers and administrators learned science from courses that neglected the integration of knowledge and accumulation of understanding. Our proof-of-concepts investigations document that, as the result of conducting an inquiry project, students develop more cohesive, robust and coherent accounts of complex scientific topics and continue to develop their ideas after completing classes (Linn, Bell, & Davis, in press; Linn & Hsi, 2000; Linn & Slotta, 2000; Slotta & Linn, 2000).

The ***Web-based Inquiry Science Environment (WISE***) projects address aspects of geoscience [see interface below as well as http://wise.berkeley.edu]. WISE projects (see list below) have improved knowledge integration in studies of over 10 thousand students in varied educational contexts. These projects leverage modern technologies to flexibly adapt to new student populations as well as to local weather, geological features, or waterways. Flexibly adaptive projects also support customization by embedding assessment in the software to track student learning and teacher activities. Teachers, as part of professional development can modify the projects and their curriculum based on student progress.

WISE can promote knowledge integration in the geoscience curriculum by engaging students in inquiry: the intentional process of diagnosing problems, critiquing experiments, distinguishing alternatives, planning investigations, researching conjectures, searching for information, constructing models, debating with peers, and forming coherent arguments. Inquiry projects promote knowledge integration by introducing new, normative ideas and by helping students link, evaluate, connect, critique, sort out, and test all of their ideas. Most science standards (NRC, 1996; AAAS, 1994; NRC, 1999; AAUW, 2000) mandate teaching science and technology as inquiry, yet 90% of teachers primarily use other methods (Becker, Ravitz, & Wong, 1999; Alberts, 2001; Horizon, 2001). Detractors argue that inquiry projects take time away from the numerous topics in the science standards. Advocates contend that technology-enhanced inquiry can help students become self-motivated learners of science and technology who continue to deepen their understanding even after completing the science curriculum.

WISE has established a mentored professional development program that has four features crucial for teacher knowledge integration. First, analysis of student ideas; second, reflection, which enables teachers to regularly review and enhance their ideas about a particular teaching strategy; third, pivotal cases that introduce new ideas for teachers to consider; and fourth, customization, that enables teachers to implement their ideas about effective teaching and redesign of WISE projects.

WISE responds to research on scaling by (1) creating a multidisciplinary partnership with a common vision for reform; (2) dynamically connecting curriculum, professional development, assessment, technology, and administrative policies; (3) developing technology-enhanced, flexibly adaptive curriculum materials and regularly refining them; (4) designing professional development that supports customization of materials by taking advantage of the local knowledge and creativity of teachers, administrators, and students; and (5) carrying out a research program with multiple indicators of success, opportunities to refine curriculum, instruction, and assessment based on performance as well as rigorous comparisons and longitudinal investigations.

THE SCAFFOLDED KNOWLEDGE INTEGRATION FRAMEWORK.

WISE takes advantage of experimental research on science learning synthesized in the Scaffolded Knowledge Integration framework. Research on how individuals make sense of the natural world clarifies why inquiry instruction succeeds (Bransford, Brown, & Cocking, 1999; Piaget, 1971; Inhelder & Piaget, 1972; Scardamalia & Bereiter, 1996; Vygotsky, 1962). Learners develop scientific expertise by interpreting the facts, processes, and inquiry skills they encounter in light of their own ideas and experiences. Typically, students hold a repertoire of ideas about scientific phenomena and investigations (Driver, 1985; Driver, Leach, Millar, & Scott, 1996; Pfundt & Duit, 1991; Eylon & Linn, 1988; Slotta, Chi, & Joram, 1995). Inquiry instruction succeeds when learners engage in diverse ideas and engage in a process we call *knowledge integration*. This is a process where students make connections between their existing ideas, information from science class, observations, or alternative perspectives suggested by peers or experiments with the goal of developing more coherent, robust, and generative science knowledge (Piaget, 1971; Vygotsky, 1962; Scardamalia & Bereiter, 1991; diSessa, 2000; Linn & Hsi, 2000).

Knowledge integration responds to research documenting nonnormative ideas that students, teachers, and many adults develop from observing the world. Although often called misconceptions, these ideas reflect intellectual effort and keen observation. Early efforts to contradict misconceptions failed because students can easily hold conflicting views: One physics student explained that objects might remain in motion in class but they certainly come to rest on the playground. A knowledge integration approach requires students to find a mechanism to connect observations—like friction. WISE seeks to promote knowledge integration among all partners in education: students, teachers, policy makers at the school and district level, curriculum and technology designers, and professional development teams.

To promote knowledge integration, first WISE projects add powerful, normative ideas including pivotal cases to the views held by learners. Second, WISE guides learners to link, connect, sort out, reflect, critique, analyze, and organize knowledge such that it becomes more cohesive, generative, and useful. Knowledge integration around science occurs not only in science classes, but continuously in everyday situations as learners respond to news articles about science, personal dilemmas such as health decisions, and policy issues such as environmental stewardship. For example, the WISE project on genetically modified food spurs students to seriously consider complex topics like gene flow, critique persuasive messages, sort out alternative perspectives and make sense of conflicting information.

Knowledge integration has interpretive, cultural and deliberate aspects (Linn, 2001). Learners interpret new material in light of their own ideas and experiences, frequently relying on personal perspectives rather than instructed ideas. To take advantage of the interpretive nature of learning, WISE projects add *pivotal cases* that help students organize their ideas (Linn, in press). For example, students contrast genetic modification of the Hawaiian papaya in the 1900s with methods of crossing varieties of Irish potatoes in the 1800s. Learning happens in a *cultural context* where group norms, expectations, and social supports shape learner activity (Dewey, 1900, 1901; Vygotsky, 1962; Cole, 1996; Lave & Wenger, 1992) and impact views of who should be scientists in the future. WISE projects help students understand scientific advance by showcasing controversial aspects of science, engaging students in constructing arguments using evidence, and supporting debates where students negotiate norms and reach conclusions. WISE promotes equity by supporting diverse learners and ensuring participation of all students. Individuals make *deliberate* decisions about their own science learning, future course selection, and career choice. WISE projects direct energy towards knowledge integration by asking students to predict outcomes, test their ideas, and reflect on their progress to increase learning (e.g., White and Frederikson, 1998; Chi, 1996).

WISE PROJECTS

WISE library projects (see the projects at http://wise.berkeley.edu)

Investigation Projects:
Awful Waste of Space... This project incorporates data collected by scientists to support students' exploration of planets found outside our solar system. Students think about, discuss, and model relationships between conditions that are necessary for life to begin on these newly discovered planets. Students also compare two methods that are currently in use to look for other life in the universe.
Creek Detectives. This project introduces Pine Creek, its location in the community, and its

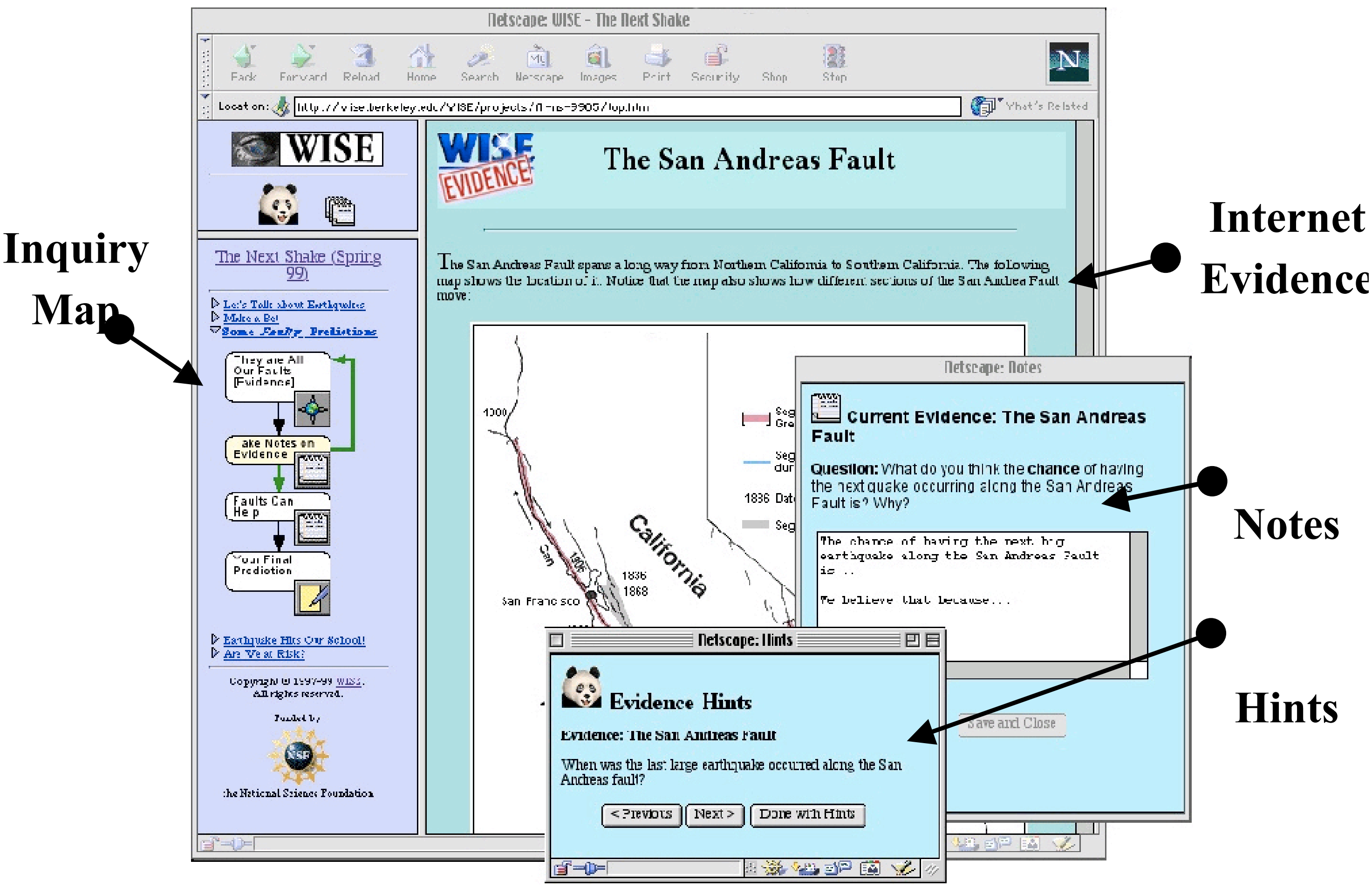
watershed. The project asks students to compare and contrast the creek at different points along the water path and at different seasons. Students learn about watersheds, what is carried in them, and how to make careful observations and predictions based on their observations at the local creek and online images.
Drink or Swim? Students learn about water quality by trying to answer the question about beach water; Would you drink or swim? Students read a story about two children who get ill from swimming in water, learn about water contaminants, and have a class discussion (both online and in the classroom) about water uses. The main goal is to teach students that depending on how water is used it can be safe or unsafe.
How do Earth and Space Plants Grow? In this project, students investigate different conditions for growing plants in space and growing plants on the earth. After thinking about the differences, they predict which plants are regular earth plants and which plants are NASA space plants. This will involve observing plant growth and development daily, collecting, and analyzing qualitative and quantitative data.
Pine Creek - Introduction. Students are invited to become detectives as they explore a local creek, its environment and ongoing status. Students participate in field trips, acquisition of data through water testing and observations, application of data to tables and charts, and interpretation of data for planning future trips and jobs at the creek. Students also upgrade the quality of the environment around the creek.
Probing Your Surroundings. Students explore thermal equilibrium in the context of the temperature of objects around them. After making predictions, and gathering data, students create and electronically discuss principles to explain that data. Students then go on to explore why objects feel hot or cold.
Rainforest Interactions. How might deforestation affect the endangered rainforest animal I have studied? This project explores trophic level interactions among species in a rainforest. It will be part of a multi-project rainforest study involving understanding some of the basic processes of ecosystems, analyzing some of the statistical data concerning deforestation, and developing viable conservation plans.
The Next Shake Project. In this project, students critically examine earthquake predictions made by others, and then come up with their own prediction for "the next big shake." They explore evidence from the World Wide Web that illustrates the effects of earthquakes on buildings and other structures. Using this evidence, they then evaluate how safe their own school would be during an earthquake.

What Makes Plants Grow? In this project, students will explore the factors needed to sustain plant life on earth such as soil, water, nutrients and light. They will utilize the World Wide Web to investigate the above factors required for optimal plant growth.
Yellow Starthistle: Briones Park. Yellow Starthistle is an invasive exotic plant pest throughout the western United States. In this project students first learn a little about the history and biology of the plant. Students study the results of a five year study. In the final activity students assume the role of one of the people impacted by the control plan in a presentation to a decision making board.
Controversy Projects:
California Flora - Native or Alien? In the "California Flora - Native or Alien?" project students learn about invasive non-native (alien) plants and three strategies for controlling or eliminating their impact. Students first learn to identify non-native plants in the area where they live and the major methods of intervention to control their spread. Students develop a plan which they present.
Controversy in Space. This project serves to introduce students to the role of controversy in advancing scientific discovery. Students investigate how scientists use evidence to support their claims.
Deformed Frogs - The Chemical Hypothesis. The Environmental Chemical Hypothesis investigates in more detail the argument that frog deformities are being caused by an environmental chemical that stimulates growth.
Deformed Frogs - The Parasite Hypothesis. This project gives more explicit information about the mechanism of the parasite hypothesis: observations and experiments by scientists; additional information about the complex life cycle of the trematode, some of which is spent in a tadpole; and Lefty the Frog, an important example that the parasite hypothesis has difficulty explaining.
Genetically Modified Foods in Perspective. The unit was designed with the goal of improving students' understanding of genetically modified foods: both their science content knowledge and their understanding of the complexity of this controversy. This requires students to think about the advantages and disadvantages of genetic engineering of foodstuffs and organic versus intensive farming.
How Far Does Light Go? Can light travel forever until absorbed, or does it eventually die out? Students are introduced to several pieces of 'evidence' which focus on different aspects of the physics of light. Students critique and organize this evidence in an attempt to answer the dilemma for themselves.
Malaria Introduction. In the "Malaria Controversy" project, students learn about three different

strategies for controlling the spread of malaria. Students analyze and examine evidence from the World Wide Web related to the malaria controversy. Students investigate the three suggested strategies for controlling the spread of malaria.
Origins. How did the universe come to be? This question serves as the entry point into students' exploration of sound and light waves, doppler effect, etc. Students use these concepts to explore the current debate between big bang and steady state theory. Students also explore creation stories from around the world in order to think about the role of religion and science in various cultures.
The DDT-Malaria Controversy. In this project, students critique the scientific evidence related to the productive uses and harmful side-effects of DDT. Based on what they learn about this pesticide and what they already know about malaria, they create an argument about the proposed global ban of DDT and present this argument during a classroom debate.
The Deformed Frogs Mystery. This project lays the foundation for the investigation of the nature and cause of frog deformities. This project can provide an introduction for in-depth investigation of the competing hypotheses involved in the controversy.
Wolves in your Backyard. This project first introduces students to the basic biology of wolves, addresses some frequently asked questions, as well as the nature of wolves. The project then presents some biology of predator-prey relations, and asks students to think about their own model for the food chain. Students explore the different perspectives of the wolf control controversy.
Critique Projects:
New Tabloid Trash or Serious Science Debate. Students study and apply a methodology for evaluating Internet materials to several different articles. Students then discuss and critique the way each group evaluated the articles.
Sunlight SunHEAT. Students learn about the topic of passive solar energy. Students also develop and apply criteria in the process of critiquing information found on the World Wide Web. Who wrote it and why? Are claims supported by evidence? What questions do you have after reading through the information?
Design Projects:
Ocean Stewards. This project teaches students about the ocean environment and the reasons for conducting expeditions within this environment. Students can explore six different National Marine Sanctuaries (NMS) in order to learn about the different marine habitats and the flora and fauna. Students will then prepare a proposal for an expedition within the chosen sanctuary.
What's in a House? In this project students design a house which would be energy efficient in a

desert environment. Their design is based on evidence which compares desert weather with their own local weather and how plants have adapted to the extremes of the desert climate.

The WISE environment



R E F E R E N C E S

Almond, R.G., Steinberg, L.S., & Mislevy, R.J. (1999). A sample assessment using the four process framework. White paper prepared for the IMS Working Group on Question and Test Inter-Operability. Princeton, NJ: Educational Testing Service. [pdf format]

American Association for the Advancement of Science (1994). Benchmarks for science literacy: Project 2061. New York: Oxford.

American Association of University Women (AAUW) Educational Foundation (2000). Tech-Savvy: Educating girls in the new computer age. Washington, D.C.: AAUW.

Anderson, J. R. & Schunn, C. D. (2000) Perspectives on learning, thinking, and activity. Educational Researcher, 29, 11-13. 227

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. Educational Researcher, 25(4), 5-11.

Anderson, J. R., Reder, L. M., & Simon, H. A. (1997). Situative versus cognitive perspectives: Form versus substance. Educational researcher, 26(1), 18-21.

Anderson, R. D., & Mitchner, C. P. (1994). Research on science teacher education. In D. Gabel (Ed.), Handbook of Research in Science Teaching and Learning . New York: Macmillan.

Aronson, Elliot (1978). The jigsaw classroom. Beverly Hills: Sage.

Baddeley, A.D., & Longman, D.J.A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21,* 627-635.

Bahrick, H.P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General, 108,* 296-308.

Ball, D.L. (1996). Teacher learning and the mathematics reforms: What we think we know and what we need to learn. Phi Delta Kappan, 500-508.

Ball, D.L. (1998). Unlearning to teach mathematics. *Journal for the Learning of Mathematics* 8(1): 40-48.

Becker, H., Ravitz, J. & Wong, Y. (1999). Teacher and Teacher-directed Student Use of Computers and Software. Report #3. Teaching Learning and Computing: 1998 National Survey. Irvine, CA: University of California at Irvine, Center for Research in Information Technology in Organization.

Becker, H.J., and Anderson, R. (1998). Teaching, Learning, and Computing: Available at http://www.crito.uci.edu/tlc/html/findings.html.

Becker, Henry Jay. (February 1999). "Internet Use by Teachers: Conditions of Professional Use and Teacher-Directed Student Use." Center for Research on Information Technology and Organizations, The University of California, Irvine and University of Minnesota.

Bell, P. (1997). Using argument representations to make thinking visible. In R. Hall, N. Miyake, & N. Enyedy (Ed.), Proceedings of CSCL '97: The Second International Conference on Computer Support for Collaborative Learning (pp. 10-19). Toronto, Canada: University of Toronto Press.

Bell, P. (1998). Designing for students' conceptual change in science using argumentation and classroom debate. Unpublished doctoral dissertation, University of California, Berkeley.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the Web with KIE. International Journal of Science Education [Special Issue], 22, 797-817.

Bell, P., Davis, E. A., & Linn, M. C. (1995). The Knowledge Integration Environment: Theory and design. In J. L. Schnase & E. L. Cunnius (Eds.), Proceedings of the Computer Supported Collaborative Learning Conference, CSCL '95, Bloomington, IN (pp. 14–21). Mahwah: Erlbaum.

Bell, P., Linn, M. C. & Hoadley, C. (in press). Design Research. In Linn, M. C., Bell, P., & Davis, E. A. (Eds.), Internet Environments for Science Education. Mahwah, New Jersey, Lawrence Erlbaum Associates.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127,* 55-68.

Bjork, R.A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp.185-205). Cambridge, MA: MIT Press.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher and A. Koriat (Eds*.),* *Attention and peformance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 435-459). Cambridge, MA: MIT Press.

Blank, R., Kim, J., & Smithson, J. (2000, June, 2000). Survey Results of Urban School Classroom Practices in Mathematics and Science: 1999 Report – Using the Survey of Enacted Curriculum Conducted during Four USI Site Visits. Study Monograph No. 2, Systemic Research, Inc.

Blumenfeld, P.C., Fishman, B.J., Krajcik, J.S., Marx, R. (2000). Creating usable innovations in systemic reform: Scaling up technology-embedded project-based science in urban schools. Educational Psychologist, 35, 149-164.

Blumenfeld, P., Marx, R. W., Krajcik, J., Fishman, B., & Soloway, E. (in press). Creating research on practice: Scaling inquiry supported by technology in urban middle schools. Educational Psychologist.

Blumenfeld, P.C., Soloway, E., Marx, R., Krajcik, J.S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing; supporting the learning. Educational Psychologist, 26, 369-398.

Bowyer, J. (1990). Scientific and Technological Literacy: Education for Change. UNESCO (United Nations Educational, Scientific, and Cultural Organization), Paris.

Bowyer, J. & Ridgeway, C. (2000). Partnerships and Systemic Reform in Science Education. Journal of Staff Development. 21: 10-15.

Bransford, J. D. (1979). Human cognition: Learning, understanding, and remembering. Belmont, CA, Wadsworth.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). How people learn: Brain, Mind, Experience, and School. Washington, DC: National Research Council.

Bransford, J.D., Sherwood, R.D., Hasselbring, T.S., Kinzer, C.K. & Williams, S.M. (1990). Anchored instruction: Why we need it and how technology can help. In Nix, D. & Spiro, R. (Eds.), Cognition, education, and multimedia. Hillsdale: Erlbaum.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. Journal of the Learning Sciences, 2, 141-178.

Brown, A.L. & Campione, J.C. (1994). Guided discovery in a community of learners. In McGilly, K. (Ed.), Classroom lessons: Integrating cognitive theory and classroom practice. Cambridge: MIT/Bradford.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(1), 12-42.

Bruer, J. T. (1993). Schools for thought: A science of learning in the classroom. Cambridge, Mass.: MIT Press.

Bruner, J. S. (1979). On knowing: Essays for the left hand. Cambridge, Belknap Press of Harvard University Press.

Case, R., Griffin, S., & Kelly, W. (1999). "Socioeconomic gradients in mathematical ability and their responsiveness to intervention during early childhood." In Keating, D.P. & Hertzman, C. (Eds.), Developmental Health and the Wealth of Nations: Social Biological, and Educational Dynamics. New York: Guilford Press.

Cazden, C. B. (1988). Classroom discourse : the language of teaching and learning. Portsmouth, NH: Heinemann.

Champagne, A. B., Klopfer, L. E., & Anderson, J. H. (1980). Factors influencing the learning of classical mechanics. American Journal of Physics, 48, 1074-1079.

Chi, M. T. H. & Slotta, J. D. (1993). The ontological coherence of intuitive physics. Commentary on A. diSessa's "Toward an epistemology of physics." Cognition and Instruction, 10, 249-260.

Chi, M. T. H. (1996). "Constructing self-explanations and scaffolded explanations in tutoring." Applied Cognitive Psychology **10**: S33-S49.

Christina, R. W., & Bjork, R. A. (1991). Optimizing long-term retention and transfer. In D. Druckman & R. A. Bjork (Eds.), *In the mind's eye: Enhancing human performance* (pp. 23-56). Washington, DC: National Academy Press.

Clancey, W. J. (1999). Conceptual coordination: How the mind orders experience in time. Mahwah, N.J., L. Erlbaum Associates.

Clark, D. (2000). Scaffolding Knowledge Integration Through Curricular Depth. Unpublished Doctoral Dissertation. University of California, Berkeley.

Clark, H. (1996). Design of performance - Based assessments as contributors to Student Knowledge Integration. Unpublished doctoral dissertation, University of California, Berkeley.

Clement, J. (1991). Non-formal reasoning in science: The use of analogies, extreme cases, and physical intuition. Mahwah: Erlbaum.

Clement, J. (1993). "Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics." Journal of Research in Science Teaching, **30**(10): 1241 - 1257.

Clifford, G. J. (1973). A history of the impact of research on teaching. In R. M. W. Travers (Ed.), Second handbook of research on teaching (pp. 1-46). Chicago: Rand McNally.

Cobb, P. & Bowers, J. (1999). "Cognitive and situated learning perspectives in theory and practice." Educational Researcher, **28**(2): 4-15.

Cognition and Technology Group at Vanderbilt. (1997). The Jasper Project: Lessons in Curriculum, Instruction, Assessment, and Professional Development. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. Review of Educational Research, 64(1), 1-35.

Cole, M. (1996). Cultural psychology: A once and future discipline. Cambridge, MA: Belknap/Harvard.

Collins, A. (1992). Toward a design science of education. In Scanlon, E. & O'Shea, T. (Eds.), New directions in educational technology. New York: Springer-Verlag.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. American Educator, 15, 6-11; 38-39.

Confrey, J. (1996). The role of new technologies in designing mathematics education. In C. Fisher, D. Dwyer, and K. Yocam (Eds.), Education & Technology: Reflections on Computing in Classrooms. San Francisco: Apple Press/Jossey-Bass, pp.129-149.

Confrey, J. (in press). Improving research and systemic reform: Toward equity and quality. In Handbook of methods for research in learning and teaching science and mathematics.

Cronbach, L. J. (1982). Designing evaluations of educational and social programs. San Francisco, CA: Jossey-Bass.

Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington.

Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. International Journal of Science Education [Special Issue], 22, 819-837.

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

Dewey, J. (1900). Psychology and social practice. Psychological Review, 7, 105-124.

Dewey, J. (1901). Psychology and social practice. Chicago, University of Chicago Press.

Dewey, J. (1916). Democracy and education: An introduction to the philosophy of education. New York: MacMillan Publishing Co.

diSessa, A. (1988). Knowledge in pieces. In Forman, G. et. al. (Eds.), Constructivism in the computer age. Hillsdale: Erlbaum.

diSessa, A. (1992). Images of Learning. In De Corte, E., Linn, M.C., Mandel, H. & Verschaffel, L. (Eds.), Computer-based learning environments and problem solving. Berlin: Springer-Verlag.

diSessa, A. (2000). Changing minds: Computers, learning, and literacy. Cambridge, MA, MIT Press.

diSessa, A. A. & Minstrell, J. (1998). Cultivating conceptual change with benchmark lessons. In Greeno, J. G. & Goldman, S. (Eds.). Thinking practices. Mahwah, NJ, Lawrence Erlbaum Associates.

diSessa, A. A. (1991). Local sciences: Viewing the design of human-computer systems as cognitive science. In J. M. Carroll (Ed.), Designing Interaction: Psychology at the Human-Computer Interface (pp. 162-202). Cambridge, England.

diSessa, A. A. (1993). Toward an epistemology of physics. Cognition and Instruction, 10, 105-225.

diSessa, A., Hammer, D., Sherin, B. & Kolpakowski, T. (1991). Inventing graphing: Metarepresentational expertise in children. Journal of Mathematical Behavior, 10, 117-160.

Driver, R. (1985). Changing perspectives on science lessons. In N. Bennett, & C. Desforges (Eds.), Recent advances in classroom research. British Journal of Psychology Monograph.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). Young People's Images of Science. Buckingham, UK: Open University Press.

Duffy, T. M. (1997). Strategic Teaching Framework: An instructional model for learning complex, interactive skills. In C. R. Dills & A. J. Romiszowski (Eds.), Instructional development paradigms (pp. 571-592). Englewood Cliffs, NJ: Educational Technology Publications.

Dwyer, D. C., Ringstaff, C., & Sandholtz, J. H. (1991, May). Changes in teachers' beliefs and practices in technology-rich classrooms. Educational Leadership, 45-52.

Edelson, D. C. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. Journal of the Learning Sciences, 8(3/4), 391-450.

Edelson, D. C., & O'Neill, D. K. (1994). The CoVis Collaboratory Notebook: Computer support for scientific inquiry. Paper presented at the New Orleans, LA.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences, 8*(3/4), 391-450.

Eisenhower Professional Development Program. (2000). Does Professional Development Change Teaching Practice? Results from a Three Year Study. Evaluation report.

Erickson, F. (1992). Ethnographic microanalysis of interaction. In M.D. LeCompte, Millroy, W.L. & J. Preissle, Handbook of qualitative research in education. San Diego: Academic.

Eylon, B. S. and M. C. Linn (1988). Learning and instruction: An examination of four research perspectives in science education. Review of Educational Research 58(3): 251–301.

Feurzeig, W., & Roberts, N. (Eds.). (1999). Modeling and simulation in science and mathematics education. New York: Springer.

Fine, M. (Ed.). (1994). Chartering urban school reform: Reflections on public high schools in the midst of change. New York: Teachers College Press.

Foley, B. (2000). Visualization Tools: Models, Representations and Knowledge Integration. Unpublished doctoral dissertation, University of California, Berkeley, CA.

Frase, L.T., Chudorow, M., Almond, R.G., Burstein, J., Kukich, K., Mislevy, R.J., Steinberg, L.S., & Singley, K. (in press). Technology and assessment. In H.F. O'Neil & R. Perez (Eds.), Technology applications in assessment: A learning view. [html document]

Gardner, M., Greeno, J. G., Reif, F., Schoenfeld, A. H., diSessa, A., & Stage, E. (Eds.). (1990). Towards a scientific practice of science education. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gik, M. L. & Holyoak, K. J. (1980). "Analogical problem solving." Cognitive Psychology, **12**: 306-355.

Glaser, R. (1976). Cognitive psychology and instructional design. In D. Klahr (Ed.), Cognition and Instruction, p. 303-316. Hillsdale: Erlbaum.

Gordin, D. & Pea, R.D. (1995). Prospects for scientific visualization as an educational technology. Journal of the Learning Sciences, 4, 249-279.

Greeno, J. G. (1989). Situations, mental models, and generative knowledge. In Klahr, D. & Kotovsky, K. (Eds.), Complex information processing: The impact of Herbert A. Simon. Hillsdale, NJ, Lawrence Erlbaum Associates.

Griffen, P. & Cole, M. Contextual factors in education improving science and mathematics education for minorities and women. Madison: Wisconsin Center for Education Research.

Grillmeyer, O. (2000). Functional composition, recursion, and applicative operators: An exploration into the more difficult areas of functional programming. Unpublished doctoral dissertation, University of California, Berkeley, CA.

Gunstone, R. F. (1988). Learners in science education. In P. J. Fensham, Ed., Development and dilemmas in science education . London: Falmer Press.

Halloun, I. A. and Hestenes, D. (1985)The initial knowledge state of college physics students, American Journal of Physics 53, 1043-1056.

Hawkins, J., Frederiksen, J., Collins, A., Bennett, D. & Collins, E. (1993). Assessment and technology. Communications of the ACM, 36, 74-76.

Heath, S.B. (1983). Ways with words : Language, life, and work in communities and classrooms. Cambridge: Cambridge.

Hegarty, M., Quilici, J., Narayanan, N.H., et al. (1999). "Multimedia instruction: Lessons from evaluation of a theory-based design." Journal of Educational Multimedia and Hypermedia, **8**(2): 119-150.

Heubert, J., & Hauser, R. (Eds.). (1998) High Stakes: Testing for Tracking, Promotion, and Graduation. Washington, D.C., National Academy Press.

Hilgard, E. (1996). History of educational psychology. Handbook of Educational Psychology, 990-1004.

Hoadley (1999). Scaffolding scientific discussion using socially relevant representations in networked multimedia. Unpublished doctoral dissertation, University of California, Berkeley, CA.

Hoadley, C. M., & Hsi, S. (1993). A multimedia interface for knowledge building and collaborative learning, The Adjunct Proceedings of InterCHI '93, (International Computer-Human Interaction Conference) (pp. 103–104). Amsterdam, The Netherlands: Association for Computing Machinery.

Hoadley, C. M., Mann, L. M., Linn, M. C., & Clancy, M. J. (1996). When, why, and how do novice programmers reuse code? In W. Gray & D. Boehm-Davis (Eds.), Empirical studies of programmers: Sixth workshop (pp. 109–129). Norwood, NJ: Ablex.

Hoadley, C., & Linn, M. C. (2000). Teaching science through on-line peer discussions: SpeakEasy in the Knowledge Integration Environment. International Journal of Science Education [Special Issue], 22, 839-857.

Hodas, S. (1993). Technology refusal and the organizational culture of schools. Education Policy Analysis Archives, 1(10).

Holland and Thayer (1998). Differential item performance and the Mantel-Haenzsel procedure. In H. Wainer and H.I. Braun (Eds.), *Text Validity* (pp. 129-145). Hillsdale, NH: Lawrence Erlbaum Associates.

Hsi, S. (1997). Facilitating knowledge integration in science through electronic discussion: The Multimedia Forum Kiosk. Unpublished doctoral dissertation, University of California, Berkeley, CA.

Hsi, S., Linn, M. C., and Bell, J. (1997) The role of spatial reasoning in engineering and the design of spatial instruction. Journal of Engineering Education. Vol 86, No. 2, pp.151-158

Huberman, A. M. (1989). Research on teachers' professional lives. Oxford ; New York: Pergamon Press.

Inhelder, B. & Piaget, J. (1958/1972). The growth of logical thinking from childhood to adolescence. New York: Basic Books.

Issacs, E. A., & Tang, J. C. (1996). Technology transfer: So much research, so few good products. Communications of the ACM, 23-36.

Jackson, S. L., Stratford, S. J., Krajcik, J., & Soloway, E. (1994). Making system dynamics modeling accessible to pre-college science students. Interactive Learning Environments, 4, 233-257.

Jacoby, L. L., Bjork, R. A., & Kelley, C. M. (1994). Illusions of comprehensions and competence. In D. Druckman and R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance.* (pp.57-80) Washington, DC: National Academy Press.

Jones, S. (1998). Information, Internet, and Community: Notes toward an understanding of a community in the Information Age. In S. Jones (Ed.), Cybersociety 2.0: Revisiting computer-mediated communication and community : Sage Publications.

Kafai, Y. B., & Resnick, M. (Eds.). (1996). Constructionism in practice: Designing, thinking, and learning in a digital world. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kali, Y., Orion, N., & Eylon, B. (in press) "A Situative Approach for Assessing the Effect of an Earth-Science Learning Program on Student's Scientific Thinking Skills." Journal of Research in Science Teaching.

Keller, E. F. (1983). Gender and Science. San Francisco, CA, W.H. Freeman.

Kennedy, M. M. (1997). The connection between research and practice. Educational Researcher, 26(7), 4-12.

Kim, J., Crasco, L., Blank, R., & Smithson, J. (2001, April, 2001). Survey Results of Urban School Classroom Practices in Mathematics and Science: 2000 Report – Using the Survey of Enacted Curriculum Conducted during Eight USI Site Visits. Study Monograph No. 3, Systemic Research, Inc.

Kim, J., Crasco, L., Lee, D., Karantonis, A., & Leavitt, D. (2000, September, 2000). Urban School Key Indicators of Science and Mathematics Education: Based on Key Indicator Data System (KIDS-1999), Volme I, II, III, IV, and Appendix. Systemic Research, Inc.

Kintsch, W. (1998). Comprehension: a paradigm for cognition. Cambridge, Cambridge University Press.

Kolodner, J. (1993). Case-Based Reasoning. San Mateo, CA, Morgan Kaufmann Publishers, Inc.

Krajcik, J. S., Blumenfeld, P. C., & Starr, M. L. (1993). Integrating knowledge bases: An upper elementary teacher preparation program emphasizing the teaching of science. In P. A. Rubba, L. M. Campbell, & T. M. Dana (Eds.), Excellence in educating teachers of science (pp. 37-54). Colombus, OH: ERIC Clearinghouse.

Krajcik, J. S., Blumenfeld, P. C., Marx, R. W., & Soloway, E. (1994). A collaborative model for helping middle grade science teachers learn project-based instruction. The Elementary School Journal, 94(5), 483-497.

Krajcik, J., & Layman, J. (1993, March, 1993). Microcomputer-based laboratories in the science classroom. Research Matters to the Science Teacher, 31.

Krajcik, J., Marx, R.W., Blumenfeld, P.C., Soloway, E. & Fishman, B. (2000). Inquiry based science supported by technology: Inquiry and motivation among urban middle school students. Conference proceedings of the annual meeting of the American Educational Research Association, 2000, New Orleans LA.

Kuhn, D. (1993). "Science as argument: Implications for teaching and learning scientific thinking." Science Education, **77**(3): 319-337.

Lampert, M. and Blunch, M. L. (1998). Talking mathematics in school : studies of teaching and learning. New York, NY, Cambridge University Press.

Lampert, M., & Ball, D. (1998). Teaching, multimedia, and mathematics: Investigations of real practice. New York: Teachers College Press.

Lampert, M., Heaton, R. & Ball, D. (1994). Using technology to support a new pedagogy of mathematics education. Mathematics Teacher Education, 12, 276-289.

Landay, J. A., & Davis, R. C. (1999). Making sharing pervasive: Ubiquitous computing for shared note taking. IBM Systems Journal, 38(4), 531.

Lave, J. (May 1988). The culture of acquisition and the practice of understanding. Report IRL88-0007 to Institute for Research on Learning.

Lave, J. and Wenger, E. (1990). Situated Learning: Legitimate Peripheral Participation. Cambridge, UK, Cambridge University Press.

Leinhardt, G., Young, K. M., & Merriman, J. (1995). Integrating professional knowledge: The theory of practice and the practice of theory. Learning and Instruction, 5, 401-408.

Levey, B. Designing instruction for teachers using technology in their science classes. Unpublished masters project.

Lewis, C. Can you lift 100 kilograms?, The secret of magnets, & The secret of trapezes. Video series.

Lewis, C. & Tsuchida, I. (1997). Planned educational change in Japan: The case of elementary science instruction. Journal of Educational Policy, 12, 313-331.

Lewis, C. & Tsuchida, I. (1998). A lesson is like a swiftly flowing river: Research lessons and the improvement of Japanese education. American Educator, Winter, 14-17; 50-52.

Lewis, C. (1995). Educating hearts and minds: Reflections on Japanese preschool and elementary education. New York: Cambridge University Press.

Lewis, E. (1991). The process of scientific knowledge acquisition among middle school students learning thermodynamics. Unpublished doctoral dissertation, University of California, Berkeley, CA.

Lewis, E. (1996). Conceptual change among middle school students studying elementary thermodynamics. Journal of Science Education and Technology, 5, 3-31.

Linn, M. C. (1992). The Computer as Learning Partner: Can Computer Tools Teach Science? In K. Sheingold, L. G. Roberts, and S. M. Malcolm (Eds.), Technology for Teaching and Learning. Washington, DC: American Association for the Advancement of Science.

Linn, M. C. (1995). "Designing computer learning environments for engineering and computer science: The Scaffolded Knowledge Integration framework. " Journal of Science Education and Technology, 4 (2), 103-126.

Linn, M. C. (1996). From separation to partnership in science education: Students, laboratories, and the curriculum. In R. F. Tinker (Ed.), Microcomputer based labs: Educational Research and standards (pp. 13-46) (Vol. 156). Berlin: Springer-Verlag.

Linn, M. C. (2000). "Designing the Knowledge Integration Environment: The partnership inquiry process." International Journal of Science Education, **Special Issue**(22): 781-796.

Linn, M. C. (2002). Science education: Preparing lifelong learners. In Smelser, N. J. & Baltes, P.B. (Eds.), International Encyclopedia of the Social and Behavioral Sciences (Vol. 3). New York: Pergamon.

Linn, M. C. (in press). WISE design for lifelong learning—Pivotal Cases. In Gärdenfors, P. & Johansson, P. (Eds.), Cognition, Education and Communication Technology. Mahwah, New Jersey, Lawrence Erlbaum Associates.

Linn, M. C. & Clancy, M. J. (1992). The case for case studies of programming problems. Communications of the ACM, 35(3): 121-132.

Linn, M. C., Clark, D., & Slotta, J. D. (in press). WISE Design for Knowledge Integration. In Barab, S. (Ed.), Building Sustainable Science Curriculum: Acknowledging and Accommodating Local Adaptation, Science Education.

Linn, M. C., Clement, C., & Pulos, S. (1983). Is it formal if it's not physics? Journal of Research in Science Teaching,20(8), 755-770.

Linn, M. C., & Hsi, S. (2000). Computers, Teachers, Peers: Science Learning Partners. Mahwah: Erlbaum.

Linn, M.C., Lewis, C., Tsuchida, I. & Songer, N.B. (2000). Beyond fourth grade science: Why do U.S. and Japanese students diverge? Educational Researcher, 29, 4-14.

Linn, M. C., & Muilenburg, L. (1996). Creating lifelong science learners: What models form a firm foundation? Educational Researcher, 25(5), 18-24.

Linn, M. C., Shear, L., Bell, P., & Slotta, J. D. (1999). Organizing principles for science education partnerships: Case studies of students' learning about 'Rats in Space' and 'Deformed Frogs'. Educational Technology Research & Development, 47(2), 61-84.

Linn, M.C. & Slotta, J.D. (2000, October) WISE science. Educational Leadership, 29-32.

Linn, M.C. & Songer, N.B. (1991). Teaching thermodynamics to middle school students: What are appropriate cognitive demands? Journal of Research in Science Teaching, 28, 885-918.

Linn, M. C., Songer, N. B., & Eylon, B. (1996). Shifts and convergences in science learning and instruction. In Calfee, R. & Berliner, D. (Eds.), Handbook of Educational Psychology. Riverside, NJ, Macmillan**:** 438-490.

Mannes S.M., & Kintsch, W. (1987). Knowledge organization and text organization. *Cognition and Instruction, 4,* 91-115.

Marx, R. W., Freeman, J. G., Blumenfeld, P. C., & Krajcik, J. S. (in press). Learning in the new science classroom. In C. E. Weinstein & B. McCombs (Eds.), Strategic learning: Skill, will and self-regulation . Hillsdale, NJ: Erlbaum.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14,* 1-43.

Means, B. (1994). Technology and education reform: The reality behind the promise. San Francisco: Jossey-Bass.

Minstrell, J. (1992). Facets of students' knowledge and relevant instruction. In Duit, R., Goldberg, F. & Niedderer, H. (Eds.), Proceedings of the International Workshop on Research in Physics Education: Theoretical Issues and Empirical Studies (pp. 110-128). Kiel, Germany: Institut fur die Padagogik der Naturwissenshaften.

Mislevy, R.J. and Wilson, M.R. (1996). Marginal maximum likelihood estimation for a psychometric model of discontinuous development. *Psychometrika, 61,* 41-71.

Mislevy, R.J., & Gitomer, D.H. (1996). The role of probability-based inference in an intelligent tutoring system. User-Modeling and User-Adapted Interaction, 5, 253-282. [preprint] [pdf format]

Mislevy, R.J., Almond, R.G., Yan, D., & Steinberg, L.S. (1999). Bayes nets in educational assessment: Where do the numbers come from? In K.B. Laskey & H.Prade (Eds.), Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (437-446). San Francisco: Morgan Kaufmann. [preprint] [pdf format]

Mokros, J. R., & Tinker, R. F. (1987). The impact of micro-computer based labs on children's ability to interpret graphs. Journal of Research in Science Teaching, 24(4), 369-383.

Nakhleh, M. B., & Krajcik, J. S. (1994). The effect of level of information as presented by different technologies on students' understanding of acid, base and pH concepts. Journal of Research in Science Teaching, 31(10), 1077-1096.

National Academy of Education. Committee on Educational Research, Coleman, J. S., Suppes, P., & Cronbach, L. J. (1969). Research for tomorrow's schools: disciplined inquiry for education; report. New York: Macmillan.

National Council For Social Studies. (2001). Standards for Social Studies. http://www.ncss.org/standards/2.0.html

National Council of Teachers of Mathematics. (2001). Principles and Standards for School Mathematics. http://standards.nctm.org/

National Research Council (1996). National science education standards. Washington, DC.

National Research Countil (1999). High Stakes: Testing for Tracking, Promotion and Graduation. Washington, D.C.

National Science Teachers Association. (1996). NSTA scope, sequence, and coordination program. A framework for high school science education. NSTA Publication: Washington, D.C.

Neilson, J. (1999). Designing Web Usability: The Practice of Simplicity. Indianapolis: New Riders Publishing. Also see: http://www.useit.com/papers/

Nersessian, N. J. (1991). Conceptual change in science and in science education. History, philosophy, and science teaching. M. R. Matthews. [Ed.]. New York, Teacher's College Press**:** 133-148.

Novak, J. D., & Gowin, D. B. (1984). Learning how to learn. New York: Cambridge University Press.

Novak, J.D. & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. American Educational Research Journal, 28 (1), 117-153.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-monitoring and comprehension-fostering activities. Cognition and Instruction, 2, 117-175.

Palinscar, A. S. & Brown, A. L. (1984). "Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities." Cognition and Instruction, **1**: 117-175.

Papert, S. (1980). Mindstorms: Children, computers, and powerful ideas. New York: Basic.

PCAST. (1997). Report to the President on the use of technology to strengthen K-12 education in the United States . Washington, DC: Panel of Educational Technology.

Pea, R. D. & Gomez, L. (1993). Distributed multimedia learning environments: The Collaborative Visualization Project. Communications of the ACM, 36(5), 60-63.

Pea, R. D. (1999). New media communications for improving education research and practice. In E. C. Lagemann & L. S. Shulman (Eds.), Issues in Education Research : Jossey-Bass.

Pea, R. D., Tinker, R. F., Linn, M. C., Means, B., Bransford, J. D., Roschelle, J., Hsi, S., & Songer, N. (1999). Towards a Learning Technologies Knowledge Network. Educational Technology Research and Development, 47(2), 19-38.

Perkins, D. N. (1993, Fall). Teaching for understanding. American Educator, 28-35.

Peterson, P. (1998). Why do educational research? Rethinking our roles and identities, our texts and contexts. Educational Researcher, 27(3), 4-10.

Pfundt, H., & Duit, R. (1991). Students' alternative frameworks. (3rd ed.). Federal Republic of Germany: Institute for Science Education at the University of Kiel/Institut für die Pädagogik der Naturwissenschaften.

Piaget, J. (1971). Structuralism. New York: Harper and Row.

Pizzini, E. L., & Shepardson, D. P. (1991). Student questioning in the presence of the teacher during problem solving in science. School Science and Mathematics, 91(8), 348-352.

Poole, H., and Page, J. A. (1995). Beyond Tracking: Finding Success in Inclusive Schools. Bloomington, IN. Phi, Delta, Kappa Educational Foundation.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? Educational researcher, 29(1), 4-15.

Reese, W. (1999). What history teaches about the impact of educational research on practice. Review of Research in Education, 24, 1-20.

Reif, F. & Larkin, J. H. (1991). "Cognition in scientific and everyday domains: Comparison and learning implications." Journal of Research in Science Teaching (Special Issue: Students' models and epistemologies), **28**(9): 733-760.

Reiser, B.J., Kimberg, D.Y., Lovett, M.C., & Ranney, M. (1992). Knowledge representation and explanation in GIL, an intelligent tutor for programming. In J.H. Larkin & R.W. Chabay, Computer-assisted instruction and intelligent tutoring systems: Shared goals and complimentary approaches. Hillsdale: Erlbaum.

Resnick, L., & Hall, M. (1998). Learning organizations for sustainable education reform. Daedalus, 127, 89-118.

Resnick, L.B. & Resnick, D.P. (1989). Tests as standards of achievement in schools. In The uses of standardized tests in American education [Proceedings of the 1989 Educational Testing Center invitational conference] (pp. 63-80), Princeton: ETS.

Riel, M. (1998). Learning communities through computer networking. In Greeno & Goldman (Eds.), Thinking practices : Erlbaum.

Robinson, V. M. J. (1998). Methodology and the research-practice gap. Educational Researcher, 27(1), 17-27.

Rogoff, B. & Lave, J. (1984). Everyday cognition: Its development in social context. Cambridge: Harvard.

Roth, W. M. (1994). Experimenting in a constructivist high school physics laboratory. Journal of Research in Science Teaching, 31, 197-223.

SCANS (1991) What Work Requires of Schools. National Technical Information Service (NTIS), Technology Administration, U.S. Department of Commerce, Springfield, VA 22161. NTIS Order Number: PB92-146711INZ.

Scardamalia, M., & Bereiter, C. (1992). Text-based and knowledge-based questioning by children. Cognition and Instruction, 9, 177-199.

Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. Journal of the Learning Sciences, 1(1), 37-68.

Scardamalia, M., & Bereiter, C. (1993). Technologies for knowledge building discourse. Communications of the ACM, 36(5), 37-41.

Schauble, L., Leinhardt, G., & Martin, L. (1998). A framework for organizing a cumulative research agenda in informal learning contexts. Journal of Museum Education, 22(2&3), 3-8.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207-217.

Schmidt, W. H., Raisen, S. A., Britton, E. D., Bianchi, L. J., & Wolfe, R. G. (1997). Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Science. Dordrecht/Boston/London: Kluwer Academic Publishers.

Schneider, R., Krajcik, J., & Marx, R. W. (2000). The role of educative curriculum materials in reforming science education. Paper presented at the Annual meeting of the American Educational Research Association, New Orleans, LA.

Schoenfeld, A. (1994). Reflections on doing and teaching mathematics. In A. Schoenfeld, Mathematical thinking and problem solving (pp. 53-70). Hillsdale: Erlbaum.

Schoenfeld, A. H. (1986). On having and using geometric knowledge. In Hiebert, J. (Ed.), Conceptual and Procedural Knowledge: The Case of Mathematics. Hillsdale, NJ, Lawrence Erlbaum Associates.

Schwartz, D. L., Lin, X., Brophy, S., & Bransford, J. D. (1999). Toward the development of flexibly adaptive instructional designs. In C. Reigeluth (Ed.), Instructional Design Theories and Models, Volume II . Hillsdale, NJ: Lawrence Erlbaum Associates.

Shea, J.B., & Morgan, R.L. (1979) Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 179-187.

Shear, L. (1998). When Science Learners are Language Learners: Designing linguistically aware instruction to teach science "the knew way." Unpublished master's thesis, University of California, Berkeley, CA.

Shepard, R. N. (1990). *Mind sights: Original visual illusion, ambiguities, and other anomalies, with a commentary on the play of mind in perception* (p. 48). New York: W. H. Freeman.

Shrader, G., & Gomez, L. M. (1999). Design research for the Living Curriculum. In C. Hoadley (Ed.), Proceedings of Computer Supported Collaborative Learning 1999 (pp. 527-537). Palo Alto, CA.

Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), Handbook of research on teaching (pp. 3-36). New York: MacMillan.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher, 15*, 4-14.

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard *Educational Review, 57,* 1-22.

Shymanksy, J.A., Kyle, W.C., Jr., & Alport, J.M. (1983). The effects of new science curricula on student performance. Journal of Research in Science Teaching, 20, 387-404.

Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 907-912.

Simon, H. (1972). The sciences of the artificial. Cambridge, MA: MIT Press.

Slotta, J. & Linn, M.C. (2000). The Knowledge Integration Environment: Helping students use the Internet effectively. In M.J. Jacobson & R.Kozma (Eds.), Innovations in Science and Mathematics Education: Advanced Designs for Technologies of Learning (pp. 193-226). Mahwah: Erlbaum.

Slotta, J. Dodson, K. & Marantz (1998). Presentation at the annual conference of NARST.

Slotta, J. Dodson, K. & Marantz (1999). Presentation at the annual conference of AERA.

Slotta, J., & Linn, M. C. (2000). How do students make sense of Internet resources in the science classroom? In M. J. Jacobson & R. Kozma (Eds.), Learning the Sciences of the 21st Century . Hillsdale: Erlbaum.

Slotta, J.D. (2000). Toward a WISE Professional Development Model": Case studies from a whole-school study of teacher adoption. Paper presented at the National Association for Research in Science Teaching (NARST) annual meeting. April 28 - May 1,. New Orleans, LA.

Slotta, J.D., Chi, M.T.H. & Joram, E. (1995). Assessing the ontological nature of conceptual physics: A contrast of experts and novices. Cognition and Instruction.

Snyder, L., Aho, A. V., Linn, M. C., Packer, A., Tucker, A., Ullman, J., & Van Dam, A. (1999). Be FIT! Being fluent with information technology . Washington, D.C.: National Academy Press.

Songer, N. B (2000). Scaling Beyond mavericks: What do our experiments tell us? Proceedings of the NCSA Workshop on Visual Modeling of Scientific Phenomena. Washington, DC, Oct. 25-26.

Songer, N. B. (1996). Exploring learning opportunities in coordinated network-enhanced classrooms: A case of kids as global scientists. The Journal of the Learning Sciences, *5*, 297-327.

Ste-Marie, D. M., & Lemire, J. A. (1997). *Contextual Interference: A Strategy for Handwriting Skill Acquisition?* Unpublished paper.

Ste-Marie, D. M., Taylor, G. M., & Cumming, J. *The effects of contextual interference on handwriting skill acquisition*. Paper presented at the Canadian Society of Pscyhomotro Learning and Sport Psychology. Edmonton., Alberta, 1999

Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematics instruction: An overview of the TIMSS video study. Phi Delta Kappan(September), 14-21.

Stigler, J.W. TIMSS Video Study Coding and Analysis. Available on-line at: http://www.lessonlab.com/timss-r/coding.htm

Stigler, J.W. & Hiebert, J. (1999). The teaching gap: Best ideas from the world's teachers for improving education in the classroom. New York: Free Press.

Stratford, S. J., Krajcik, J., & Soloway, E. (1998). Secondary students' dynamic modeling processes: Analyzing reasoning about synthesizing, and testing models of stream ecosystems. Journal of Science Education and Technology, 7(3), 215-233.

Suppes, P. (Ed.). (1978). Impact of research on education: Some case studies. Washington, DC: National Academy of Education.

Taines, C., Patrick, H., & Middleton, M. (2000). Observations of urban middle school students engaged in technology-supported inquiry. Paper presented at the Annual meeting of the American Educational Research Association, New Orleans, LA.

Thorton, R. K., & Sokoloff, D. R. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. American Journal of Physics, 58(9), 858-867.

Tinker, R. (1999, Spring). Free and open to all: The potential for free and open source software for education. Concord Consortium Newsletter.

Tinker, R. (Ed.). (1996). Microcomputer-based labs: Educational research and standards. Berlin; New York: Springer.

Tschirgi, J. E. (1980). "Sensible reasoning: A hypothesis about hypotheses." Child Development, 51: 1-10.

Tsuchida, I., & Lewis, C. (1996). Responsibility and learning: Some preliminary hypotheses about Japanese elementary classrooms. In T. P. Rohlen & G. K. LeTendre (Eds.), Teaching and Learning in Japan (pp. 190-212). New York: Cambridge University Press.

Tufte, E. (1990). Envisioning Information. Cheshire, CT: Graphics Press.

Tufte, E. (1996). Visual Explanation. Cheshire, CT: Graphics Press.

Tyack, D., & Cuban, L. (1995). Tinkering toward utopia: A century of public school reform. Cambridge, MA: Harvard University Press.

Vanderbilt, Cognition and Technology Group at Vanderbilt (1990). Anchored instruction and its relationship to situated cognition. Educational Researcher, 19, 2-10.

Vanderbilt, Cognition and Technology Group. (1997). The Jasper Project: Lessons in curriculum, instruction, assessment, and professional development: Lawrence Erlbaum Associates.

Vye, N., Schwartz, D. L., Bransford, J. D., Barron, B. J., Zech, L., & CTGV. (1998). SMART environments that support monitoring, reflection, and revision. In Hacker, Dunlosky, & Graessar (Eds.), Metacognition in educational theory and practice . Hillsdale, NJ: Lawrence Erlbaum Associates.

Vygotsky, L. S. (1962). Thought and language (E. Hanfmann and G Vaker, Trans.). Cambridge MA: MIT Press.

Vygotsky, L. S. (1962). Thought and language. Cambridge, MA, MIT Press.

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

Walker, D.F. & Schaffarzik, J. (1974). Comparing curricula. Review of Educational Research, 44, 83-111.

Ware, M., Richardson, L., & Kim, J. (2000, March, 2000). What Matters in Urban School Reform. Study Monograph No. 1, Systemic Research, Inc.

Welch, W. W. (1979). Twenty years of science curriculum development: A look back. In D. C. Berliner (Ed.), Review of research in education (pp. 282-308). Washington, DC: American Educational Research Association.

Wellesley College Center for Research on Women (1992). How schools shortchange girls. Washington, D.C.: AAUW.

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

White, R.T. (1988). Learning science. Oxford: Basil Blackwell.

Wiesenmayer, R.L. & Meadows, G.R. (1997). Addressing science teacher's initial perceptions of the classroom uses of internet and world wide web-based resource materials. Journal of Science Education and Technology, 6 (4) 329-335.

Wilson, M., and Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13*(2), 181-208.

Wisnudel, M., Stratford, S., Krajcik, J., & Soloway, E. (1996). Using technology to support students' artifact construction in science. In M. C. Linn (Ed.), International Handbook of Science Education . Netherlands: Kluwer Publishers.

Wittgenstein, L. (1958). Preliminary studies for the "Philosophical investigations"; generally known as the Blue and Brown books. New York,: Harper.

Wolf, S. A., Carey, A. A., & Mieras, E. L. (1996). What is this literachurch stuff anyway? Preservice teachers' growth in understanding children's literary response. Reading Research Quarterly, 31(2), 130-157.

Wolf, S. A., Mieras, E. L., & Carey, A. A. (1996). What's after "what's that?": Preservice teachers learning to ask literary questions. Journal of literacy research, 28, 459-497.

[1]Ten Things We Know About Learning and Their Implications for Geoscience Education

Joel Michael[1,3] and Harold Modell[2,3]
[1]Rush Medical College, Chicago, IL
[2]National Resource for Computers in Life Science Education, Seattle, WA
[3]Physiology Educational Research Consortium, Seattle, WA

Learning and memory have been issues of central concern to psychology since the late 19th century. Over the past 40 years the development of cognitive science has greatly accelerated the pace of studies of these phenomena. More recently the learning sciences have extended such studies to the classroom.

One dominant school of thought about learning, one that is particularly prevalent in all discussions of science learning, is constructivism. There are two basic tenets of all forms of constructivism. (1) Knowledge can not be transmitted from one individual to another; rather, knowledge is built by the learner using internal cognitive processes acting on stimuli from the external environment. (2) The learner's process of building new knowledge starts with a foundation of everything that he or she already knows; the learner is never a blank slate (tabula rasa). Ten key ideas can be extracted from current attempts to understand learning. These are listed below and some of their educational implications are indicated.

Key Idea 1: All learning occurs on the foundation of already learned knowledge and skills.

Educational implications: There are several implications of this tenet about learning. First, it is essential to understand what students already know when they embark on any learning experience. Without an understanding of this input state of the students, it is impossible to set realistic goals for students and impossible to plan activities that will help them reach those goals.

Second, it is to understand that some of what the students bring to the classroom is flawed (incomplete or simply wrong). One goal for all teachers is presumably to help students correct existing mental models. Without knowing what misconceptions are prevalent, it is impossible to plan activities that will help students correct their mental models. Another goal must be to avoid inadvertently contributing to the students' building faulty mental models because of something that happens in your classroom.

Finally, since new knowledge will be learned in the context of old knowledge, it is important to understand the students' prior knowledge so that new material can be organized and presented in a way that can be most appropriately related to the old knowledge.

Key Idea 2: To the extent that the old knowledge is faulty, the learning of new knowledge will be compromised.

Educational Implications: Again, it is essential to understand that misconceptions exist, and to know what misconceptions students exhibit. That being said, it is important for you to understand that telling the learner that some piece of their knowledge is wrong, or simply attempting to provide them with, "give" them, the right knowledge, DOES NOT WORK.

Just as students come into the classroom with misconceptions, it is important to recognize that things that happen in the classroom may very well create new misconceptions or at least reinforce old ones, even though that is obviously not your intention. Teachers must be sensitive to the mechanisms that are known to contribute to misconception formation and seek to minimize their occurrence in the classroom. For example, while analogies can be powerful aids to learning, their misuse in the classroom can leads to serious misconceptions if the learner over-generalizes the analogy to situations in which it does not apply.

Key Idea 3: Declarative (what) and procedural (how) knowledge are different, and the processes of learning them are different.

Educational implications: It is important to understand the difference between declarative and procedural knowledge for several reasons. Decisions about the learning objectives for some content area in a course must reflect the type of knowledge being addressed. Learning resources must be made available to facilitate student mastery of both types of knowledge. In addition, assessments must be appropriate for the kinds of knowledge the students have been asked to learn. If students are expected to master both declarative and procedural knowledge, it is essential that assessment tools test for the mastery of both kinds of knowledge.

---

[1]This material is adapted from Michael and Modell, *Active Learning in the Science Classroom* (under contract to Lawrence Erlbaum Associates, Publishers). **DO NOT CITE!**

Key Idea 4: Learning declarative knowledge involves building mental models or representations.

Educational implications: It is essential that the learner be given opportunities to build many different models of the knowledge being learned, perhaps differing in the modality that is activated (vision or audition) or the source of the information that is used. Aids to relating the new knowledge being learned to old knowledge are also helpful. It is also important that learners be required to think about their models in an overt, and explicit way. In some instances it may be helpful for the learner to use specific model-building skills such as concept mapping.

Key Idea 5: Practice with timely and appropriate feedback is required for all procedural psychomotor learning.

Educational implications: Practice and appropriate feedback are the keys to procedural learning. Thus, the learning environment must provide opportunities for students to solve problems or practice some sensory-motor skill and receive appropriate, timely feedback. This can be difficult to arrange. Students learning to solve problems need to have access to many problems. Learning to prepare a thin section of a mineral sample for microscopy may require many samples and adequate access to the needed equipment. Moreover, many of the educational experiences commonly incorporated into science courses may not help students reach the desired level of mastery. Textbook problems with their minimalist solutions at the end of the chapter or book may not provide enough feedback to help the learner. Similarly, watching an instructor solve a problem at the board is all too often a passive experience for students, one that contributes little to procedural learning. Watching the instructor carry out a task in the laboratory will not lead to the level of student proficiency desired.

Key Idea 6: Retention and the ability to utilize knowledge (meaningful learning) is facilitated by building associations (links) between old knowledge structures and the new knowledge being learned.

Educational implications: In order to assist students to link their new knowledge to their old knowledge, it is essential to understand what is in the students' existing knowledge base. Knowing this, one can assist students to overtly and explicitly build links or associations with that old knowledge. Furthermore, having insight into the students' old knowledge base helps put the new material into a context that students are likely to find most relevant and, thus, more easily related to their old knowledge.

Key Idea 7: The ability to construct multiple representations of new knowledge is an important component of meaningful learning.

Educational implications: All learning environments must provide the learner with opportunities to engage new material in a context that facilitates establishing relationships between the old and the new. Further, the learning environment must make obvious the relationships that exist between what the learner already knows and what he or she is attempting to learn.

Success at meaningful learning will be facilitated if the learner is given opportunities to create multiple representations of the new knowledge being acquired. One way to do this is to provide new information using multiple modalities (vision, audition, touch). Another useful tactic is to provide different examples of the phenomenon being learned, thus facilitating links to many different, already stored mental models.

Key Idea 8: Some knowledge and skills, when acquired, are context-specific while other knowledge and skills may be more readily transferred to a new domain.

Educational implications: If certain skills can be generalized, and if certain scientific models can be applied to many situations, it is imperative that students be taught these skills and models. This focus on generality must start with the initial learning and must be reinforced as each new opportunity for application is encountered.

Key Idea 9: Collaborative or cooperative effort can yield more individual learning than individual effort alone.

Educational implications: The learning environments that teachers create must provide opportunities for students to work with one another in ways that will result in learning for all students. Students learning together share their knowledge, explain their position, argue and debate and in this way build more robust mental models (representations). Group work can occur in the lecture hall, the student laboratory or the discussion section.

Key Idea 10: Articulating explanations, whether to peers, teachers, or one's self, facilitates learning.

Educational implications: The benefit to students of having to articulate their knowledge is one explanation for the benefits of collaborative/cooperative learning. In all settings in which students work together to learn or solve a problem, one critical feature is the need for the learners to communicate with one another. Communication is essential if information is be shared and ideas exchanged. Arguing and debating issues, negotiating and reaching agreements in order to complete the assigned task all require that each individual be able to articulate to others what he or she believes, understands, and doesn't understand. When learners who are solving problems generate explanations of what they are doing, whether voiced or not, they seem to learn more. Clearly, then, it is essential to arrange learning environments in which students are encouraged to talk to students and to the instructor.

Where do we go from here? Incorporating these ideas about learning into our teaching

Constructivism has demonstrated great explanatory power about what happens in our classrooms. While it does not answer all of the questions about learning, and certainly does not provide a prescription for what we ought to do in the classroom, it offers solid evidence in support of many of things supporters of classroom reform are advocating. Incorporating what we have learned about learning into classroom practices will call for changing at least some of what you do, but if your goal is helping the learner to learn, the results of these changes will be gratifying.

References and a bibliography are available on request.

Conducting Discipline-Specific Educational Research: An Example

Joel Michael
Department of Physiology
Rush Medical College
Chicago, Illinois
and
Physiology Educational Research Consortium
Seattle, WA

Great strides are being made by cognitive psychology and the learning sciences to understand how learning occurs. At the same time the science education community is studying the learning and teaching of science in the classroom, while simultaneously attempting to bring about changes in the ways in which science is taught in our schools. Arguably these attempts at understanding and reform are most advanced in physics and chemistry, disciplines whose national societies (via their meetings and journals) have focused considerable attention on science education.

Physiology is a discipline that has more recently joined this endeavor. The American Physiological Society established a Section on the Teaching of Physiology in 1985 and established a journal, *Advances in Physiology Education*, in 1989. Shortly after this, a group of physiology teachers, educators, and educational researchers launched a program of research and faculty development. This group represents one example of how discipline-specific educational research might be implemented.

The Physiology Educational Research Consortium (PERC)

PERC is a group of 13 physiology teachers located at 12 post-secondary educational institutions across the country - from New York to California and Washington. These institutions range from community colleges to large research universities. We are all classroom physiology teachers with an interest in doing a better, more effective job of helping our students learn. We all believe that research on learning and teaching is necessary if we are to succeed in our job of helping the learner to learn. We all believe that changes in the ways we, and our colleagues, teach physiology require that we learn to change what we do in the classroom. Details about PERC and its members can be found on our web site at http://www.physiologyeducation.org.

PERC began as a group of consultants to an SBIR project developing physiology teaching software. Group meetings led to the recognition that we shared many interests and concerns about physiology education that went beyond the use of computers in the classroom. Grant proposals were written and our initial success in obtaining funding from NSF enabled us to successfully launch our research program. We are currently in our fourth year of funding from NSF.

How does PERC work?

Of PERC's current 13 members, three of them have functioned as investigators and 10 have served as collaborators, making available their classrooms for our research. *All of our research to date has occurred in the classroom as a routine part of the activities of the course.* In some instances student participation has been strictly voluntary, with participation, or lack of participation, having no effect on a student's grade. In other cases, instructors have offered "extra credit" for participation. Several experiments were conducted as routine parts of the activities of the course. Approval of institutional IRBs has been sought and obtained for all of our studies.

All of the research questions that have been pursued by PERC have arisen from the experiences that PERC members have had in the classroom. Issues of interest to the entire group are identified and discussed. The investigators then write a proposal for discussion by the group. PERC members volunteer their classrooms for use in the particular experiment being considered. When funded, the project proceeds with frequent interactions between the investigators and the classroom instructors. In many cases the investigator(s) will travel to the classroom where the study is being carried out. In other studies, particularly where the research involves the administration of assessment instruments, the instructor takes responsibility for the experiment and the collection of the data. Recently we have begun to use a web site for constructing and administering assessments, greatly facilitating the running of experiments.

When data has been collected, analysis is begun by the investigators with interaction with the classroom instructors. Members of PERC generally meet twice a year, once at the annual Spring meeting of Experimental Biology (the yearly meeting of the American Physiological Society) and once late in the summer in either Seattle or Chicago. On these occasions there is intense interaction between investigators and the classroom instructors. The data that has been collected is discussed, conclusions are drawn, publications are decided on, and plans for future research are initiated.

What is PERC doing?

Our research is based on a simple model of the educational process in which we focus on the input state of the students (what they know when they begin some educational experience - whether it be the course, a laboratory experiment, a lecture etc.), the desired output state of the students (what do they know and what can they do after the educational experience) and the educational "treatment" (the course, the lab exercise, the discussion section) which is designed to help students get from their input state to the desired output state.

Thus, we have studied the knowledge that physiology faculty believe is prerequisite to success in their course and whether students actually possess that defined knowledge. As part of determining the input state of students in our courses, we have surveyed large numbers of students to determine the misconceptions (alternative conceptions, preconceptions, conceptual difficulties) that are present in three different areas of physiology.

We have studied the use of conventional student laboratories to help students correct faulty mental models (misconceptions) and have found that simple changes to the laboratory protocol can have dramatic effects on the success of the lab experience.

We are presently studying the consequences of helping students to understand and apply certain general models of phenomena that occur in many different physiological systems. We are also looking at the effects of different problem solving experiences on the mental models that students develop.

In addition to our research, PERC members have been heavily involved in faculty development activities at national and international physiology meetings and on individual campuses. We have run brief, two hour workshops and three to four days workshops. Our goal has been to inform our colleagues of the growing body of knowledge about learning and to help them learn to change what they do in their classrooms.

Where is PERC going next?

We will be continuing both our educational research and our faculty development activities. Future research will explore the mental models that faculty believe are appropriate for students at different levels (introductory courses versus upper level courses), what difficulties students have in building those models, and how we can best help students to achieve the models defined by the instructors. We will be developing and testing innovative learning resources that will help students develop the models that are defined by their instructors as appropriate for their level of experience.

In the faculty development arena we will be attempting to organize faculty development programs in which week long summer “institutes” are followed by periodic activities throughout the following academic year. This kind of repeated reinforcement is known to result in more, and more sustained, change than brief exposures to new approaches to classroom.

Some recent publications from PERC include:

Michael, J. A. (1998). Students' misconceptions about perceived physiological responses. *American Journal of Physiology* , 274, (*Advances in Physiology Education*, 19), S90-S98. (Available as a pdf file at http://advan.physiology.org/cgi/reprint/274/6/S90.pdf)

Rovick, A. A., Michael, J. A. et al. (1999). How accurate are our assumptions about our students' background knowledge? *American Journal of Physiology*, 276 (A*dvances in Physiology Education* , 21), S93-S101. (See http://advan.physiology.org/cgi/reprint/276/6/S93.pdf)

Michael, J. A. et al. (1999). Undergraduate students' misconceptions about respiratory physiology. *American Journal of Physiology* , 277 (*Advances in Physiology Education*, 22), S127-S135. (See http://advan.physiology.org/cgi/reprint/277/6/S127.pdf)

Modell, H. I., Michael, J. A. et al. (2000). Helping undergraduate repair faulty mental models in the student laboratory. *Advances in Physiology Education*, 23, 82-90. (See http://advan.physiology.org/cgi/reprint/23/1/S82.pdf)

Michael, J. A. (2001). In pursuit of meaningful learning. *Advances in Physiology Education*, 25, 145-158. (See http://advan.physiology.org/cgi/reprint/25/3/145.pdf)

Michael, J. A. et al. (in press). Undergraduates' understanding of cardiovascular phenomena. *Advances in Physiology Education.*

Michael, J. A. (2002). Misconceptions - what students think they know. *Advances in Physiology Education*, 26 , 5-6. (See http://advan.physiology.org/cgi/reprint/26/1/5.pdf)

Research on Learning in the Geosciences: Contexts, Goals and Opportunities

David W. Mogk
Dept. of Earth Sciences
Montana State University

Prepared for a workshop on "Bringing Research on Learning to the Geosciences", sponsored by NSF and the Johnson Foundation, Racine WI, July 8-10, 2002

Some Contexts

What are the geosciences? Broadly defined, the geosciences encompass disciplines that study the solid earth (e.g. rocks and their structures), surficial deposits (e.g. soils, surface waters), the atmosphere, and the oceans; related disciplines are also included such as near-space physics, planetology, paleontology and physical geography. In detail, each sub-discipline in the geosciences has developed its own knowledge base, methodologies, philosophies and approaches to problem solving, and cultural attributes that inform the conduct of science in these varied domains, and by extension, educational practice in these subjects. At the same time, in the study of our planet we often apply "first principles" of science shared with sister disciplines in biology, chemistry, physics, engineering, and mathematics to understand the natural world around us—so our work is often inter- or multidisciplinary in its scope.

In recent years there has been a growing movement towards research and education using an Earth system approach (AGU 1997). Connections between different components of the Earth system are being emphasized, interesting research results are being realized at the interfaces between diverse sub-disciplines, and new hybrid disciplines are emerging, e.g. geomicrobiology. Earth system studies increasingly focus on the processes that connect the components of the Earth system: e.g. the transfer of mass and energy through complicated pathways and reservoirs, feedback mechanisms, the physical and biological evolution of systems. Time is an essential component of understanding the Earth system—the concepts of "deep geological time", rates, and fluxes. There is also an increasing awareness of the importance of understanding the linkages between the physical world and the life forms it nurtures (e.g. coupled biogechemical cycling) and with humanity (e.g. dynamics of coupled natural and human systems).

The Earth system is dynamic, heterogeneous, complex and often chaotic. And this presents a number of challenges to learning about a system that is naturally ambiguous, uncertain, and largely unpredictable. The Earth system operates on spatial scales ranging from atomic to planetary, and over time scales that may be considered instantaneous and catastrophic to inexorable over the eons. Earth system processes typically operate beyond every-day human experience, and we consequently rely on other ways of observing the Earth such as remote sensing (e.g. via satellite imaging) or by making inference based on indirect observations of things we can't see directly (e.g. seismic records and tomography). The geologic record is incomplete, and we are left with a detective's mystery trying to fill in the missing pieces. It is very difficult for geoscientists to conduct controlled experiments: we don't have a separate world that can be used as a control for comparison, and given the complexity of natural systems it is difficult to construct comprehensive physical or computational models. Nonetheless, we do find it useful to use simulations or visualizations to model, represent and interpret these complex systems.

The geosciences obviously encompass a wide variety of fields of study, and there are consequently many factors that influence what we teach, how we teach, and whether or not students can or will learn. Here are some observations for context:

- The geosciences have somewhat an identity problem. In some cases, departments self-identify as Geology (or Geological Science), Oceanography, Atmospheric Sciences along strict disciplinary lines. Other departmental names become more inclusive: Geosciences, Earth Sciences, Earth and Planetary Sciences, Geology and Geography…There's important meaning behind these names which reveals philosophy and structures for organizing information, inclusion or coverage of content and consequent curricular design, approaches to education, etc.
- There is not a curricular "canon" in the geosciences. Most departments continue to offer courses along disciplinary lines (geology, oceanography, meteorology…). Traditionally these topics have been taught as stand-alone courses with little reference to each other. Across the geoscience disciplines, or even within a given discipline, it is very difficult to find "core content" that would be common to similar courses offered by different instructors/departments. This appears to be the case for introductory courses as well as upper division courses. Even a topic as broad as plate tectonics leaves ample room for instructors to pick and choose among topics to be covered. When we convened the workshop that led to "Shaping the Future of Undergraduate Earth Science Education, Innovation and Change Using an Earth System Approach" (AGU, 1997), it quickly became apparent that participants were strongly resistant to the creation of a uniform, centralized (and prescriptive) curriculum. Consensus showed that most geoscience educators wanted to retain autonomy in their choices of *what to teach*, and insisted on the need to optimize opportunities afforded by their local geographic setting, institutional resources, and to meet the specific needs of their student clientele. This is in sharp contrast with our friends in physics and chemistry (personal communications) wherein a standard introductory course in these fields covers essentially the same material no matter where it is taught or by whom. The question of *what to teach* has been settled (at the introductory level) a century ago in these disciplines. The geoscience educator must make choices about what to teach, as well as how to teach it.
- At the Shaping the Future workshop we did encourage participants to look for commonalities among the geoscience disciplines. Since that report, the Earth system approach is increasingly being adopted within the context of traditional classes (i.e. some connections across the Earth system are being articulated), and dedicated Earth system classes are coming on-line.
- Fundamental concepts underlie our approach to understanding the world. For example, in geology the principles that "the present is the key to the past" and the "principle of uniformitarianism" are applicable throughout the Earth system and

geologic time (or are they?). Given the complex and incomplete geologic record, we often apply the tenets of "multiple working hypotheses" to help winnow out erroneous explanations as new evidence is acquired. This presages a very Popperian approach to our science in that we can rarely hope to "prove" our hypotheses, but rather, must often settle for interpretations that are internally consistent at best.

- Field work is at the heart of our science, although this is done in different contexts in the different disciplines: the classic field geologist with brunton compass and pick in hand hiking over hill and dale; a change in emphasis from mapping to sampling; shipboard activities on extended deep ocean cruises involving drilling, dredging, and submersible dives; airborne data collection using sophisticated instrumentation; and an increasing dependence on technology, e.g. global positioning systems, remote sensing instruments with data relayed by telemetry (i.e. you don't actually need to be in the field to collect field data), and robotics (in the deep sea, in volcanoes and on other planets).
- Geoscience education relies heavily on geospatial and temporal referencing—we need to know where we are in the Earth system, and what the physical, chemical, and biological contexts are;—when did an event occur, over what time scale, and at what rate? We must continually iterate between global principles and local examples. Concepts of relative and absolute time are essential, as well as the ability to represent "deep time" i.e. geological time over millions or billions of years. In virtually al of the geosciences it is important to understand Earth processes in three and four dimensions.
- We place a strong emphasis on use of data, real-time or archived, to interpret the world, although the range of information that is considered "data", the tools to manipulate and represent "data", and how we use data in the classroom can be quite varied (see the results of the recent "Using Data in the Classroom" workshop at dlesecommunity.Carleton.edu).
- At times a reductionist approach is required —selecting representative sub-systems to more precisely understand a given phenomenon at increasingly smaller scales of observation; at other times it is necessary to integrate or synthesize across many lines of evidence.
- To fully understand any part of the Earth system, we typically integrate a) over many scales of observation—processes that occur on the atomic scale, in aggregate, contribute to planetary scale processes; b) across disciplines—utilizing the principles, concepts, and methodologies from physics, chemistry, biology, mathematics, computer science, and engineering; c) using numerous, disparate types of data (e.g., physical and chemical measurements, imagery of many types, biological surveys, human census data…); and d) using a variety of approaches including observation (sometimes aided by technology), analysis, modeling, experiment, and theory.
- Representations of the Earth have become increasingly important to demonstrate natural processes. Models of all types, numerical, physical, visualizations, simulations, and projections are typically used to make complex systems more accessible to understanding.

Goals for Learning in the Geosciences

Goals for learning must be established in the context of the diverse attributes of the geosciences as outlined above, and in the knowledge that we must reach diverse audiences for different purposes. There are some broad domains where learning goals can be established, appropriate to each subject or audience. The challenge is to translate these broad goals into specific practice across the many interests in geoscience education.

- **An overarching goal from NSF** (NSF Geosciences Beyond 2000): *To benefit the nation by advancing the scientific understanding of the integrated Earth systems through supporting high quality research, improving geoscience education and strengthening scientific capacity.*
- **Mastery of content knowledge**: a minimum understanding of basic taxonomies, formulae, etc., is required for communication and understanding. (e.*g.* National Science Education Standards, NRC, 1996)
- **Mastery of fundamental "first principles" and concepts**: including the ability to apply a concept appropriately to a new situation.
- **Making connections**: "*Science is knowledge not of things, but of their relationships*" (Lucien Poincaré, Science and Hypothesis)
- **"Inculcating scientific habits of the mind"**: *Project 2061 Science for All Americans*, AAAS, 1989.
- **Skill development**: many skills are of a technical nature—the ability to operate an instrument or use a software package; other skills relate to personal growth—communication, quantitative, and interpersonal skills.
- **Attitudes about science, values, ethics**: in many cases, in our instructional practice it may be appropriate to demonstrate the personal and societal value of the scientific process and its products, and to address ethical issues that impact the conduct of science.
- **Recruitment, retention, diversity**: what practices can help to make the geosciences an attractive career option—for those who aspire to become a professional scientist, and those who need to use science in their daily activities (e.g. policy planners, journalists, …)? The geosciences have made important strides in recruiting women, but we are greatly in arrears with respect to recruitment from underrepresented groups. We need to identify barriers and incentives to make the geosciences more accessible and attractive to all people.
- **Training the "Workforce for the 21st Century"**: adapting curricula content and methods to meet the changing needs of the future job market.
- **A scientifically literate, scientifically capable public**: our personal and communal health, safety and economic well-being are directly impacted by natural hazards and resources; everyone should be able to read a weather map, be aware of local natural hazards, understand their connections to global systems.

An Invitation to Collaborate

The Earth is a wonderful natural laboratory, and geoscience classes provide great laboratories for research on learning. During the past decade the NSF has sponsored a series of workshops that have helped to build a vibrant community of geoscience educators. The professional societies (e.g. American Geophysical Union, Geological Society of America, National Association of Geoscience Teachers, American Meteorological Society, American Geological Institute) have contributed in many ways to support geoscience education through theme sessions, committee work, and publications. The Earth and Planetary Sciences are recognized as an important component of the K-12 National Science Education Standards (NRC, 1996), and geoscience courses are among the most popular on college campuses.

We recognize the significant advances that have been made in cognitive psychology (e.g. *How People Learn: Brain, Mind, Experience and School*, NRC, 1999) and in research on learning in our sister STEM disciplines. There are ample opportunities to adapt or adopt these lessons to the geosciences. At the same time, the geosciences have special interests and needs related to research on learning that are intrinsic to their subject matter, methodologies and audiences. Important outcomes of this workshop will be:

- Development of collaborations among all partners interested in pursuing future work on research on learning in the geosciences; establishing a common understanding of the contributions that can be made from other disciplines as applied to the special needs of the geosciences.
- Engaging the geoscience education community to contribute to research on learning projects. Collaborative projects are needed to design and implement research experiments that meet high scholarly standards in the fields of human cognition, education, and the geosciences.
- Providing a research environment in geoscience educational settings that will help to contribute to the larger arena of understanding human learning (e.g. with respect to 3 and 4-dimensional representations; optimizing learning in field settings; measuring the value of simulations and visualizations in instructional practice; understanding complex, dynamic, and chaotic systems).
- Translating the results of this research on learning into effective instructional practice in the geosciences—covering all geoscience disciplines, instructional settings, and for all audiences.

We anticipate that the proceedings of this workshop will be the first step towards long and productive collaborations with all interested contributors.

References:

American Association for the Advancement of Science, 1989, *Project 2061 Science for All Americans*, 269 pp.

American Geophysical Union, 1997, *Shaping the Future of Undergraduate Earth Science Education, Innovation and Change Using an Earth System Approach*. F. Ireton, C. Manduca, D. Mogk (eds).

National Research Council, 1996, *National Science Education Standards*, National Academy Press, Washington, DC.

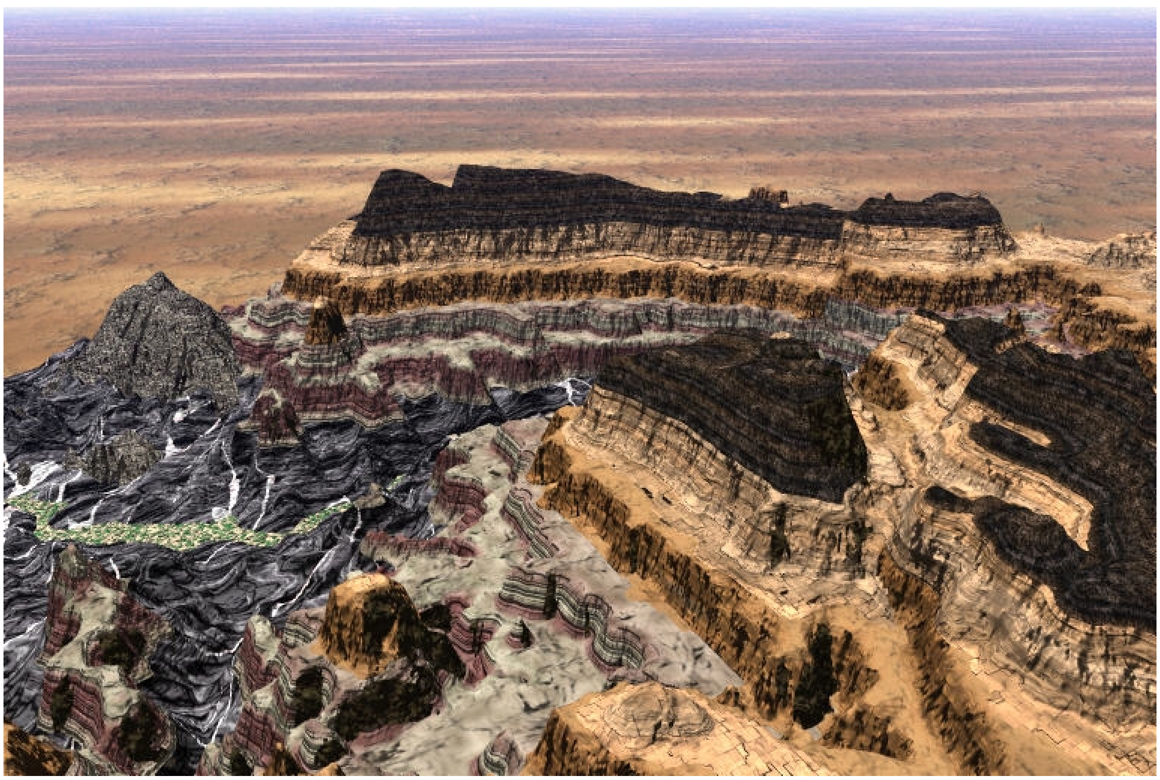
National Research Council, 1999, *How People Learn: Brain, Mind, Experience and School*, National Academy Press, Washington, DC.

National Science Foundation, 2000, *NSF Geosciences Beyond 2000 Understanding and Predicting Earth's Environment and Habitability*, NSF 00-27.

The Hidden Earth:
Visualization of Geologic Features and their Subsurface Geometry

Michael D. Piburn, Stephen J. Reynolds,
Debra E. Leedy, Carla M. McAuliffe,
James P. Birk, and Julia K. Johnson

ARIZONA STATE UNIVERSITY



Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA, April 7-10, 2002.

ABSTRACT

Geology is among the most visual of the sciences, with spatial reasoning taking place at various scales and in various contexts. Among the spatial skills required in introductory college geology courses are spatial rotation (rotating objects in one's mind), and visualization (transforming an object in one's mind). To assess the role of spatial ability in geology, we designed an experiment using (1) web-based versions of spatial visualization tests, (2) a geospatial test, and (3) multimedia instructional modules built around innovative QuickTime Virtual Reality (QTVR) movies.

Two introductory geology modules were created – visualizing topography and interactive 3D geologic blocks. The topography module was created with Authorware and encouraged students to visualize two-dimensional maps as three-dimensional landscapes. The geologic blocks module was created in FrontPage and covered layers, folds, faults, intrusions, and unconformities. Both modules had accompanying worksheets and handouts to encourage active participation by describing or drawing various features, and both modules concluded with applications that extended concepts learned during the program.

Computer-based versions of paper-based tests were created for this study. Delivering the tests by computer made it possible to remove the verbal cues inherent in the paper-based tests, present animated demonstrations as part of the instructions for the tests, and collect time-to-completion measures on individual items. A comparison of paper-based and computer-based tests revealed significant correlations among measures of spatial orientation, visualization and achievement.

Students in control and experimental sections were administered measures of spatial orientation and visualization, as well as a content-based geospatial examination. All subjects improved significantly in their scores on spatial visualization and the geospatial examination. There was no change in their scores on spatial orientation. Pre-test scores on the visualization and geospatial measures were significantly lower for the experimental than for the control group, while post-test scores were the same. A two-way analysis of variance revealed significant main effects and a significant interaction. The unexpected initial differences between the groups resulted from an uneven gender distribution, with females dominating the experimental group and males the control group. The initial scores of females were lower than those of males, whereas the final scores were the same. This demonstrates that spatial ability can be improved through instruction, that learning of geological content will improve as a result, and that differences in performance between the genders can be eliminated.

BACKGROUND

Visual-Spatial Ability

The exceptional role of spatial visualization in the work of scientists and mathematicians is well-known. The German chemist August Kekule described how atoms appeared to "dance before his eyes," and is said to have discovered the structure of the benzene ring by "gazing into a fire and seeing in the flames a ring of atoms looking like a snake eating its own tail (Rieber, 1995)." Roger Shepard (1988) discusses many examples of how spatial visualization was important to the creative imagination of scientists like Einstein, Faraday, Tesla, Watson and Feynmann.

The performance of scientists on standard tests of spatial ability is so high that Anne Roe (1961) had to create special measures for her studies of exceptionally creative scientists. Successful science students in high school and college have higher scores on traditional measures of spatial ability than is true of other students of their age and ability (Carter, LaRussa & Bodner, 1987; Pallrand & Seeber, 1984; Piburn, 1980).

Despite the obvious importance of spatial visualization to the geological sciences, there are few studies that explore this relationship. Muehlberger and Boyer (1961) found that students' scores on a standardized visualization test correlated positively with their grades in an undergraduate structural geology course, as well as grades in previously taken geology courses. In a more recent study, Kali and Orion (1996, 1997) reported that the "ability mentally to penetrate a structure," which they called visual penetration ability (VPA) is highly related to the ability to solve problems on their Geologic Spatial Ability Test (GeoSAT).

The exact nature of scientific abilities in the spatial realm is not clear. Spatial ability can be conceived of in a variety of ways, from recognizing rotated figures (Shepard & Metzler, 1971), to disembedding and "restructuring" information from visual arrays (Witkin, Moore, Goodenough & Cox, 1977) to "mental imagery" (Shepard, 1978).

It is possible to think of spatial abilities as a cluster of factorially distinct qualities. Studies of traditional measures show that they separate into at least two groups. Spatial orientation ("the ability to perceive spatial patterns or to maintain orientation with respect to objects in space") and visualization ("the ability to manipulate or transform the image of spatial patterns into other arrangements") are factorially distinct abilities (Ekstrom, French, Harman & Dermen, 1976). When considered in this way, the contribution of spatial ability to achievement in science is about the same as that of verbal ability (Piburn, 1992).

Another way to think about spatial ability has been provided through the work of Howard Gardner (1985). His theory of Multiple Intelligences proposes that spatial intelligence is one of several quite distinct intellectual abilities. These separate intelligences find their greatest expression in the specialized practices of society. He cites, for example, the case of a child in the South Pacific who has exceptional spatial abilities, and is specially trained for a career as a navigator. Presumably, there has been some kind of a similar tacit program in our culture that has resulted in those with similar abilities being identified and trained as scientists and mathematicians.

One of the discouraging results of much of this literature is that, although the importance of spatial abilities is clear, the correlations between the results of spatial measures and achievement in science class are low. One possible explanation for this comes from the study of expertise. Expert performance, it turns out, is very context specific. Chess players can remember more than 50,000 meaningful chess positions, but are no more able than others to remember the random positioning of chess pieces on a board. Expert map-makers have an incredible visual memory for maps, but no better memory than others for other kinds of displays (Ericsson & Smith, 1991). It is a reasonable hypothesis that the correlations would rise substantially if the measures of spatial ability were more closely aligned with the specific science content that was being tested.

Some recent proposals in cognitive science and education seem to reflect the idea that knowledge is contextual. These include Anchored Instruction (The Cognition and Technology Group at Vanderbilt, 1990), Problem-Based Learning (Albanese & Mitchell, 1993) and Situated Cognition (Brown, Collins & Duguid, 1989). These three psychological and educational models are similar insofar as they suggest that learning occurs best in situations that are complex, problem-based, realistic and reflective of the actual content of instruction. Very few of these have been attempted in science education, and even fewer in the earth sciences. However, Smith and Hoersch (1995) have reported on the application of problem-based learning in the college geology classroom, including tectonics, mineralogy and metamorphic petrology. They conclude that it "seems more effective than didactic learning at overturning incorrect preconceptions and encouraging interdisciplinary integration of content, independent learning, and active student participation."

Spatial Visualization in Geology

Practicing geologists engage in many kinds of spatial-visual activities. Much of classical geology is concerned with understanding the distribution, both on the surface and at depth, of geologic units, geologic structures, and natural resources. To help them visualize these distributions, geologists have developed various kinds of maps, diagrams, and other graphical representations of geologic data (Rudwick, 1976; Davis and Reynolds, 1996). Geologists use these types of illustrations to help them visualize landscapes, surficial and subsurface geology, and geologic changes over time.

Geologists use topographic maps to visualize the shape of the land surface from the contours. To do this, a geologist must mentally transform the abstract, two-dimensional map, with its squiggly contour lines, into a three-dimensional landscape. Geologists perform a similar spatial transformation by visualizing the landscape from a two-dimensional aerial photograph. In this case, geologists use visual clues from the aerial photo, such as shadows, the typical appearance of streams and other features, and a mental picture of what the landscape "ought to look like."

Geologists rely extensively on geologic maps, which show the types and ages of rock units exposed on the surface, as well as faults, folds, and other geologic structures. Most geologic maps incorporate a topographic base map so that geologic features can be referenced to their actual elevation and location on the map. A geologist examining a geologic map will alternately focus on the geology and the topography to gain a mental picture of how the two are related. This mental process is a type of disembedding, in which one aspect is mentally isolated from a multifaceted context.

From a geologic map, geologists may construct a geologic cross section, which is an interpretation of the subsurface geology from one point to another. A cross section is like cutting a big slice through the landscape, picking it up, and looking at it from the side, in the same way we look at layers inside a cake. Geologists use cross sections to visualize the subsurface geology and to explore for natural resources by determining the depth to a specific coal-bearing layer, copper deposit, or oil field.

Geologists also construct a sequence of diagrams to illustrate successive geologic changes in an area. Many geologic processes require so much time that humans are not around long enough to observe any changes in the landscape. To approach this problem, geologists have developed the technique of "trading location for time." By this it is meant that geologists look at several present-day areas and mentally arrange these into a sequence interpreted to represent an evolutionary sequence through time. A narrow deep canyon, for example, is interpreted to be a younger phase of landscape development than an area that has been eroded down into a series of low, subdued hills.

Spatial Visualization in Geology Courses

One of the main goals of a geology course is to teach students how to visualize geology in a way similar to practicing geologists. When the laboratory for Introductory Geology was redesigned, a decision was made to restrict the course to those aspects that are most important to real geologists. Students now learn how to:

- construct, read, and visualize topographic and geologic maps,
- visualize geology in the subsurface,
- visualize and reconstruct past environments from rocks and minerals,
- reconstruct geologic history from rocks, minerals, and maps, and
- understand the implications of geology for society.

To help students understand and visualize topographic maps, they construct a contour map by successively filling with water a plastic box containing a plastic mountain and drawing a map of the shoreline at each water level. After they have used such concrete manipulatives, the students interact on the computer with a module entitled Visualizing Topography. This experience is reinforced by having students use topographic maps throughout the semester to locate rock and mineral samples and decide which areas have safe slopes for situating a colony.

To help students understand and visualize geologic maps, they construct their own geologic map, on a topographic base, from three-dimensional perspectives of a computer-generated terrain called Painted Canyon (see front cover). To complete this map, students need to (1) recognize how geology and topography interact, (2) draw lines on the topographic map that correspond to boundaries between geologic units on the perspectives, and (3) reconstruct the order in which the rock units and geologic structures were formed. The students then use this geologic map to construct a cross section of the units in the subsurface and to determine the impacts of geology on a colony they must site.

Students then have a chance to apply these skills to several interesting real places. They use topographic maps, geologic maps, and rock samples from these places to reconstruct the geologic history. To help them better visualize how geologic structures appear on such maps and in the landscape, students interact with another computer-based module entitled Interactive 3D Geologic Blocks. This module and the one on visualizing topography are described in a later section of this paper.

The last several weeks of the laboratory are devoted to having students use geologic information to solve geologic problems, such as identifying the source of groundwater contamination. For these exercises, students again use contour maps, but this time of the elevation of the water table, to determine the direction of groundwater flow. Students also go on a field trip to make their own observations in the field and to use a topographic base map to construct a geologic map and cross section. They also go to the map library to use topographic and geologic maps to write a report on the geology of their hometown. The field and library assignments give the students an opportunity to apply what they have learned throughout the semester.

Developing Students' Spatial Ability

Although our schools specifically teach verbal and logical-mathematical skills, they rarely intervene in the spatial realm. But spatial ability can be taught, and the effects of such instruction have been shown to yield greater learning in science classes.

Practice with classification, pattern detection, ordering, rotation and mental manipulation of three-dimensional objects can improve spatial ability. Zavotka (1987) used computer animated graphics that "replicate mental images of rotation and dimensional transformation" with university students. The intervention was successful in improving scores on orthographic tests, but not those of mental rotation. In a computer-based intervention, McClurg (1992) created a series of puzzles for use with third- and fourth-grade students. Two, called Gertrude's Puzzles and The Pond, constituted the Spatial Patterning Group. Two others, called The Factory and Shifty Shapes, were referred to as the Spatial Rotation Group. Significant post-test differences between control and experimental groups were observed on both the Mental Rotation Test and the Figural Classification Test. In a review of visualization research in chemistry education, Tuckey and Selvaratnam (1993) present a number of techniques that have been proven effective in improving spatial skills. These involve interventions in which students observe diagrams showing successive steps in the rotation of molecules, as well as computer-based programs showing rotating molecules and their shadows.

Lord (1985, 1987) succeeded in improving the spatial ability of college students by having them try to visualize sections through three-dimensional objects, and then cut the objects to verify their predictions. His rationale for these experiments was that asking a subject to "picture in his mind the bisection of a three-dimensional form and to predict the two-dimensional shape of the cut surface" conformed with the demands predicted by the Shepard-Chipman theory of second order isomorphism. As individuals with poor spatial ability attempt to manipulate an image, they lose the one-to-one relationship between the mental image and the external object. Repeated practice appears to improve subjects' ability to maintain this correspondence between object and image.

Interventions constructed within the contexts of Piagetian theory have also been shown to improve spatial ability. Cohen (1983) conducted an experiment with elementary students studying the Science Curriculum Improvement Study (SCIS) curriculum. Students in a control group were told to leave the experimental apparatus stationary, while those in the control group were encouraged to seek a variety of alternative perspectives from which to view the experiment. Post-test scores in the experimental group showed significant improvement over those in the control group on three of eight measures of Piaget's projective groupings. Vasta, Knott and Gaze (1996) designed a "self-discovery training procedure" were able to show improvements in performance on Piaget's water-level task. They created an experiment in which they varied the shape of the bottle containing the water, thus manipulating the "field" bounding the task. This caused students to question their initial judgments and to reconsider the relationship of external boundaries of the contained water and the orientation of the water level.

Two studies (Eley, 1993; Schofield & Kirby, 1994) address the question of improving topographic map interpretation through intervention. Both show that improvement is possible, but use drastically different procedures to achieve that result. Schofield & Kirby rely heavily on Paivio's (1990) dual coding theory in the design of their experiment. They found that location of a position on a map involved both spatial and verbal strategies, as would be predicted by the theory, and that training in a verbal strategy could lead to improved performance. In contrast, the study by Eley involved training students to visualize a landscape from a topographic map and to state how the map would look to different observers. In this regard, the study was very similar to some portions of the present study. The results indicated the use of mental imagery was context specific, but that the choice of processing strategy was not, instead being more susceptible to the influence of training.

There is no doubt of a significant relationship between spatial ability and success in science. However, it is much more difficult to show that training programs leading to improved spatial ability have direct impact on school success. The review by Tuckey and Selvaratnam (1993) suggested that there was very little transfer from trained tasks to new settings. Similar results were found by Devan, et. al (1998), who found that modeling software in engineering graphics courses improves spatial skills, but that this improvement does not show any clear relationship to retention of students in engineering school.

This issue of transfer is a very important one. Proposals to create programs that improve students' visualization skills will only take on educational meaning if it can be shown that there is transfer from learning of these skills to other, more general problems, and especially those containing significant content from the sciences. The treatment provided by Pallrand and Seeber (1984) is perhaps the most detailed that has been attempted in the science education field. Students in an introductory college physics course were "asked to draw outside scenes" by viewing through a small square cut in a piece of cardboard. They were encouraged to draw the dominant lines of the scenery and to reduce the scene to its proper perspective. Subjects were also given a short course in geometry involving lines, angles, plane and solid figures, and geometric transformations. In addition, the "Relative Position and Motion" module from the Science Curriculum Improvement Study was used. Subjects located positions of objects relative to a fictitious observer, Mr. O. Individuals learned to reorient their perceptual framework with respect to observers with different orientations (pg. 510). These activities took place for 65 minutes weekly for 10 weeks. Students who went through the training showed improved visual skills, and achieved higher course grades than those who were enrolled in the same course but were not part of the experiment.

The effect of experience on spatial ability is an important question that requires further examination. Burnett and Lane (1980) were able to show that college students majoring in physical science and mathematics showed greater improvement in spatial ability than those in the humanities and social sciences. However, Baenninger and Newcombe (1989) conducted a meta-analysis that indicated that the effect of experience on spatial ability was relatively small. A newly emerging body of research may serve to offer profound new insights into this question. This research has shown a significant physiological relationship between neural structure and experience with spatial tasks. (Maguire, et al., 2000). It appears from this study that the posterior hippocampi of London taxi cab drivers are larger than those of other subjects, and that this enlargement shows a positive relationship with the amount of time spent as a taxi driver. The authors conclude that:

> "These data are in accordance with the idea that the posterior hippocampus stores a spatial representation of the environment and can expand regionally to accommodate elaboration of this representation in people with a high dependence on navigational skills. It seems that there is a capacity for local plastic change in the structure of the healthy adult human brain in response to environmental demands."

Such a result implies that prolonged experience with spatial tasks has the potential to significantly alter the physiology of the brain.

Relationships Among Gender, Spatial Ability, and Achievement

Many authors (McArthur & Wellner, 1996; Linn & Peterson, 1985; Voyer, Voyer and Bryden, 1995) trace our current awareness of the relationships among gender, spatial ability and achievement to the work of Eleanor Maccoby and Carol Jacklin (1974). In their pioneering book titled The Psychology of Sex Differences, Maccoby and Jacklin outlined the impact of gender on intellect, achievement and social behavior, and traced what was then known about the origins of psychological differences between the sexes.

In the category of "Sex Differences that are Fairly Well Established" the authors concluded that "girls have greater verbal ability than boys" and "boys excel in visual-spatial ability (pg. 351)." They also accepted the claim that boys are more analytic and excel in mathematical and scientific pursuits. They stated that "boys' superiority in math tends to be accompanied by better mastery of scientific subject matter and greater interest in science (page 89)." This led them to wonder about the link among variables discussed here, and in particular "whether male superiority in science is a derivative of greater math abilities or whether both are a function of a third factor (page 89)." It was not difficult for most people working in the field at that time to reach the tentative conclusion that spatial ability might be the link between gender and achievement in mathematics and science.

It is well known that differences in spatial ability are related to maturation. Gender differences are small in childhood, but develop in adolescence and adulthood. A number of theories have been proposed in order to explain this, each involving some combination of genetic and environmental factors. The most prominent among the former were those involving hormonal effects, maturation rates and neural development. Among the later were those that emphasized the differential socialization of boys and girls in our culture, and the resulting differences in attitude, behavior and experience that could be expected to create differences in performance on measures of mathematical, scientific or spatial ability. This discussion was developed at length by Maccoby in an article titled "Sex Differences in Intellectual Functioning (1966)." It did not seem possible at that time to resolve the conundrum. The best that Maccoby could say was that gender differences most probably resulted from "the interweaving of differential social demands with certain biological determinants that help to produce or augment differential cultural demands upon the two sexes (page 50)."

Despite a very large number of studies conducted and research reports published since the work of Maccoby and Jacklin, the issues remain unresolved now as they were then. Rather than attempting to review that massive literature, we will focus on three recent reviews that bring the reader more or less up to date on the status of the discussion. In the first, Marcia Linn and Anne Peterson (1985) question the basic assumptions of Maccoby and Jacklin. In particular, they ask about the magnitude of gender differences in spatial ability, when they first occur, and on exactly what aspects of spatial ability they are most pronounced. In the second, Daniel and Susan Voyer and M.P. Bryden (1995) re-examine these same questions. In the third, Julia McArthur and Karen Wellner (1996) perform a Piagetian analysis of spatial ability. We will follow these three questions in the fashion of Linn and Peterson.

Linn and Petersen reported a range of effect sizes for gender differences from 0.13 to 0.94 (Table 1, pg. 1486). Effect sizes greater than 0.30 (one-third of a standard deviation) are usually considered large enough to be meaningful. Those in the higher range seemed to contradict reports circulating at the time that as little as 5% of the variance in spatial ability was associated with gender, and the authors concluded that there were in fact important differences in some areas of spatial ability. The analysis conducted by Voyer, et al. confirmed this general result. They listed 172 studies (Tables 1-3, pp. 254-258), of which male performance was superior in 112, and females outperformed males in only three. There were no significant differences in the remainder. Effect sizes ranged from 0.02 to 0.66 (Table 4, page 258). Despite the fact that ten years separated these syntheses, the results remained approximately the same in their general form.

Both of these reviews also provide evidence supporting the contention that gender differences are quite small among younger children and increase with age. Linn and Petersen presented studies in which spatial ability was judged in children as young as four years old. At that age, girls were outperforming boys. But by 11 years male performance was superior, and remained so in all older samples. They showed a very rapid increase in effect sizes, from 0 to more than 1.0, in the ages between 10 and 20 years, with no further increases subsequently (Figure 4, page 1488). Voyer, et al. also documented differences with increasing age, concluding that "there is an increase in the magnitude of sex differences with age ($r=0.263$, $p<0.01$)" and that "participants below age 13 do not show significant sex differences in any of the categories of spatial tests, participants above age 18 always show sex differences, and those between ages 13 and 18 obtain significant sex differences in the spatial perception and mental rotations groupings (page 260)."

These three reviews also show how contingent the answer to the first two questions is on the nature of the task that is used to judge spatial ability. Each group of authors has created categories of spatial task for their purposes. However, they do not agree among themselves, nor are their categories the same as those which we are using in this study.

McArthur and Wellner (1996) devote their attention specifically to those tasks that were created by Piaget to describe the development of spatial reasoning. They follow his usage in categorizing tasks into three groupings: topologic, euclidian, and projective. In all of the comparisons they found in the literature, gender differences occurred only in 16% of the cases. Almost all of these were in the area of the euclidian grouping, and by far the most prominent occurred with respect to the water bottle task, in which subjects are asked to draw the water level in vessels tilted at a variety of angles.

Linn and Petersen and Voyer, et al. group spatial measures into three categories: spatial perception, mental rotation and spatial visualization. In the mental rotation category are those tests similar to the ones created by Shepard and his colleagues, in which people are asked to rotate three-dimensional figures in their mind and judge the outcome. The spatial perception category contains primarily the water-level task of Piaget and the Rod-and-Frame task of Witkin. Spatial visualization is defined primarily by various versions of the Embedded Figures Task. The paper Form Board test is the only instrument in the spatial visualization category similar to those used in this experiment. In this study, spatial visualization involves transformations of the sort that take place when paper is folded to create origami or boxes are created from flat pieces of cardboard.

Both Lynn and Petersen and Voyer et al. report very high effect sizes for measures of mental rotation. For all ages, the values given are 0.56 and 0.73. However, the results for spatial visualization are not as clear. The pooled results yield an effect size that is quite low (0.13 and 0. 19 respectively). This would lead one to conclude that the observed gender differences reside primarily in the area of mental rotation.

However, the remaining categories in both papers include measures of the cognitive style of field-dependence/field-independence within the category of spatial perception and visualization. These include several versions of the Rod-and-Frame, the Hidden Figures and the Embedded Figures tests. All involve an object that is embedded within a “field” that provides distracting stimuli. The solution to each involves overcoming field effects, an act often referred to as “restructuring” or “breaking set.” In many ways these are similar to the water bottle test described above. Voyer, et al. report an overall effect size of only 0.18 for pooled results from all versions of the embedded figures test. However, there are several forms of this instrument, of which the individually administered version is by far the most reliable. The same authors report an effect size of 0.42 for the individually administered version, a value that is not substantially different than that given for the rod-and-frame and the water bottle.

Because of the authors’ decisions to include results from the rod-and-frame and embedded figures tests, it is more difficult to judge the results of the analyses of spatial perception and visualization tests. Although the paper folding and surface development tests, in which judgments about spatially transformed figures are required, are mentioned in both studies neither group of authors reports the results of them separately in terms of effect sizes. However, Voyer, et al. report a weighted regression analysis of a variety of instruments against age of subject in which the paper folding test has the highest regression weight of any measure. The variance shared with age is almost 75%, and exceeds that of the next most powerful variables (mental rotations, card rotations and spatial relations) by a factor of three. Unfortunately, we are unable to confirm from the information given that this instrument would have had an equal superiority if its effect size had been reported separately.

From these studies, we conclude that sex differences in spatial ability are robust and that they have not changed much over time. They do appear to develop with age, and reach their peak in the late teens and early twenties. They are very situated in the task that is used to evaluate them. From the data given, the largest differences appear in the area of mental rotations, followed by those tasks that require disembedding or restructuring, and are smallest in the area of visualization. However, we believe that the final result, for the area of visualization, is untrustworthy and demands further study.

Erasing the Gender Gap

A number of the studies mentioned above (Chaim, et al., 1988; Cohen, 1983; McClurg, 1992) have shown no significant differences in the effects of training on spatial ability between females and males. If improvement has occurred, it has been approximately equivalent for the two genders, whether or not initial differences in spatial ability were observed. Others (Devon, et al., 1998; Lord, 1987; Vasta, et al., 1996) have shown that it is possible to use such interventions to improve the spatial ability of women differentially over that of men. These studies have typically involved cases where there were initial differences between males and females on pre-tests, but not on post-tests. We have reviewed no studies that have shown the spatial ability of males to improve more than that of females as the result of an intervention on spatial ability.

These results lead us to believe that observed gender differences in spatial ability and performance are probably more related to differences in experience than they are to any underlying differences in intellectual ability in the spatial arena. The fact that initial differences are either non-existent or favor males, and that they can be eliminated through relatively minor treatments, indicates that the interventions are providing important background information to females that males more often possess. This is almost certainly the result of differential experiences of men and women in our culture.

RATIONALE FOR THE STUDY

To go out into the field with a geologist is to witness a type of alchemy,
as stones are made to speak. Geologists imaginatively reclaim worlds
from the stones they're trapped in.
Frodeman (1996)

Geology is arguably the most visual of the sciences. Visualization by geologists takes place at a variety of scales, ranging from the outcrop to the region to the thin section. Many geologists have the ability to mentally transport themselves rapidly from one scale to another, using observations at one scale to constrain a problem that arose at another scale. Observations from the outcrop are used to construct a regional geologic framework, which in turn guides what features are looked for at the outcrop (Frodeman, 1996). Observations at two spatially separate outcrops may lead the geologist to visualize a major, regional anticline, along with its hidden subsurface geometry, its eroded-away projections into air, and perhaps even a causative ramp-flat thrust fault at depth.

From a rich trove of basic research in the cognitive sciences, as well as a more modest literature in science and geoscience education, it has been possible to isolate the processes of spatial orientation and visualization as crucial to the thought process of geologists. What we have constructed is a small demonstration project, carefully designed and executed, that substantiates the claim that this element of geological reasoning can be taught, and will transfer to improved performance in geology courses.

The specific objectives of the project are:

- to show that it is possible to train students to use spatial skills in real geological contexts;
- to demonstrate that such training improves performance on traditional measures of spatial ability;
- to eliminate gender differences in spatial ability;
- to show transfer from such training to extended context problems in novel settings; and
- to create innovative new computer-based materials that can be made available through the world wide web to instructors at colleges and universities.

MATERIALS DEVELOPMENT

Visualization Modules in Geology

In an effort to improve undergraduate geology education, two comprehensive modules were created. The purpose of these modules was to enhance students' spatial-visualization skills in the context of real problems presented to geologists in the field. The skills specifically targeted were spatial visualization and spatial orientation, and visual penetrative ability (Kali & Orion, 1996). The ability to reconstruct orders of events in a geologic time sequence are also crucial skills with which students have difficulty. Both modules were constructed using a learning cycle approach where students explore a concept, are introduced to the term or concept discovered during exploration, and then apply the concept in a new situation.

Multiple features of the modules, and the movies in them, such as maximum interactivity and open-ended discussions were designed to improve students' spatial visualization skills. Software packages were chosen that would accomplish this goal. Other considerations when choosing software included ease of navigation, clean screen layouts, ability to import multiple formats of images and movies, and the ability to provide feedback to students on conceptual questions. The topographic maps module was first designed in Macromedia's Authorware 5 (1998). A later version was also developed using html in FrontPage (2000) for web distribution. Once multiple screens were developed, Authorware had the advantage of easier modification when organizing screens. The editing capabilities were more user-friendly and required simple clicking and dragging to change the orders of screens. This same modification in FrontPage required changing either the page location of previous and next buttons or changing the script on each button when a page was inserted or deleted. FrontPage offered more flexibility in construction, ease of design, and distribution than did Authorware. Once the editing features were discovered in the first module, it was decided to design the blocks module using FrontPage for both cd and web distribution.

In both modules, movies were created in MetaCreations' Bryce4 (1999) and exported as QuickTime VR (virtual reality) files. Bryce4 is an animation program that can create the illusion of three-dimensional objects by using depth perception and varying lighting, shading, and color. Topographic maps of real geologic features were obtained and draped over digital topography using, MicroDEM, a program that displays and merges images from several databases. This method created the appearance of three-dimensional topography while simultaneously showing contour lines. MicroDEM is a downloadable program available on the internet.

Movies were created to rotate around various axes depending on the purpose of a module's section. The sections below on each module provide further explanation of how movies were made. All QuickTime Virtual Reality (VR) movies were created by designing image sequences in Bryce4 and importing them into VR Worx (2000). These can be viewed with Apple's QuickTime (2000) movie player. The gridlike layout of VR Worx is arranged such that each row consists of one feature (typically rotations), and columns allow elements such as shading, rotations (about another axis), transparency, deposition of layers, erosion, and faulting to change in combination with rotations.

Both modules were designed to be interactive, to achieve active learning and avoid screen-turning. Students can click buttons to choose sections from a menu or to move to different screens within a section. Active progression through the modules ensures that students will retain more information and understand more content from the movies. This encourages students to browse the sections in an order that makes the most sense to them. Since each topic progresses from simple to complex, suggestions were offered for an ideal sequence, but students were given freedom to navigate as they wished. This menu navigation also makes the modules ideal for whole classroom use. If a lesson ends mid-module, instructors can easily start the next lesson at the same point with only a few clicks of the mouse.

Another method to maximize interactivity with both modules was to create accompanying worksheets. These worksheets contain activities corresponding to random pages within the modules. The objectives of the worksheets were to ensure that students visited each section in the menu, to generate group discussions by posing open-ended questions, to encourage the interpretation and drawing of structures, and to have students describe images and movies seen on screen. The use of the worksheets also served to initiate whole class discussions at the conclusion of a module. These class discussion sessions helped students find their own areas of strengths and weaknesses as well as allowing lab instructors to determine what skills students gained from the modules.

Topographic Maps Module

The first module focused on topographic maps. Skills required for a thorough understanding of topographic maps and the use of contour lines are the identification of key geologic features on a topographic map, identification of elevation changes, and construction of topographic profiles. Students' difficulties arise from an inability to understand three-dimensional perspective depicted by two-dimensional representations. By being given topographic maps with four unique movie types, students are able to control the amount of shading in a black and white image, rotate colored landscapes from a top view to a side view, raise and lower water levels, and slice into terrains to understand how contour lines and intervals represent elevation changes. Figure M1 shows a simple hill landscape represented by each mode. This module was designed to cover three simple landscapes (hill, valley, and cliff) commonly encountered when reading and interpreting these maps. These three landscapes were presented with the four movie types mentioned above to encourage the visualization of simple features in three dimensions.

Movies were created to show the three-dimensionality of landscapes. The shading movies, both black and white and colored, were given the appearance of shadows by using the sun option in Bryce4 (see Figures M1b and M1d). Students could directly compare a flat, two-dimensional map with a three-dimensional map to draw a parallel between specific points and features on the two maps. The ability to see valleys and peaks in terms of shade and light allows to students to discover the relationship between shapes of contour lines and the geologic features they represent.

Upon entering the module, the terms topography and topographic maps are defined. Navigation suggestions are also provided. To notify users where they are within the module and to reduce the likelihood of getting lost, a title was added to the bottom of each page. The first four pages of the website serve to introduce users to the types of animations (user-controlled or instant playing and the four types of animation) they will see throughout the module. This module was constructed to be linear in order to group animations. By doing so, students adapted to each type of animation and were familiar with the changes that could be made to each landscape. This also allowed discussion questions to focus on the elements of an animation and enabled students to relate the landscapes to each other.

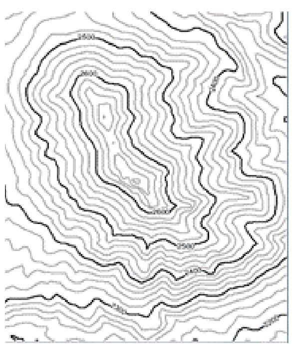


Figure 1a. Two-dimensional topographic map of a simple hill.

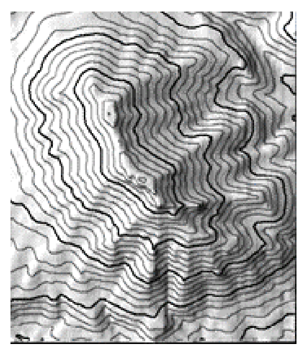


Figure 1b. Shading movie where users click and drag the mouse up and down to increase and decrease the amount of shade.

Most screens in the module are shown in a split-screen mode where the left half of the screen is a topographic map of the landscape being studied. On the right half, the various movies are presented. Directly above the movies, arrows are shown to direct users how dragging the mouse will alter the image. Figures 1a and 1b appear on one screen together. Both images on these split screens begin in identical orientations and scales so students can compare contour lines. As the user clicks and drags the mouse upward in the movie on the right, the amount of shading increases as the sun angle changes. Students immediately notice the appearance of hills, valleys, or cliffs, as well as high and low elevation points. The next screen shows colored topographic contours in which the movie rotates both vertically (to rotate up to a side view of the landscape as well as increase shading) and horizontally. Figures 1c and 1d

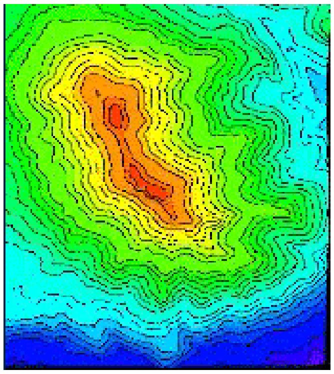


Figure 1c. Two-dimensional topographic map with color-coded elevations.

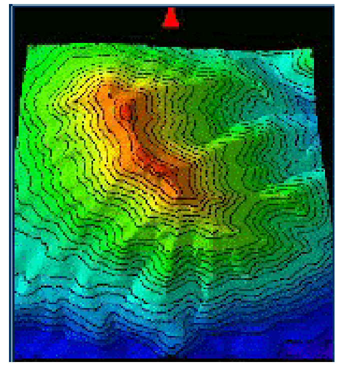


Figure 1d. Rotating and shading movie. Clicking and dragging the mouse up and down rotates vertically while changing shade. Landscape can be rotated horizontally by dragging sideways.

appear as a pair on screen. Students are then asked open-ended discussion questions that require observation and interpretation. The questions below represent types of questions asked about a still image of each landscape.

- Can you now envision what this terrain looks like, based on the map?
- What is the hill's overall shape?
- What are some of the finer details of its shape?
- Is it the same steepness on all sides?
- Is it aligned in some direction?

To check their responses, students are taken to another screen that shows a continuously playing movie that rotates both vertically (90º) and horizontally (360º). This allows students to discuss details in depth and modify any answers that were debated or unresolved. Students are then asked to write, on their worksheet, a clear verbal description for someone who has never seen each feature. They are given suggestions that may help students write their descriptions. More questions are then provided to help students clarify their descriptions. Finally, a sample description is provided by a field geologist.

The next mode of display for visualizing three dimensional features is the use of flooding water in a terrain (see Figure M1e). By clicking and dragging in the movie,

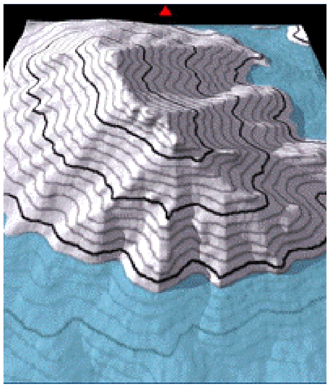


Figure 1e. Flooding movie. Users change the water level by clicking and dragging up and down.

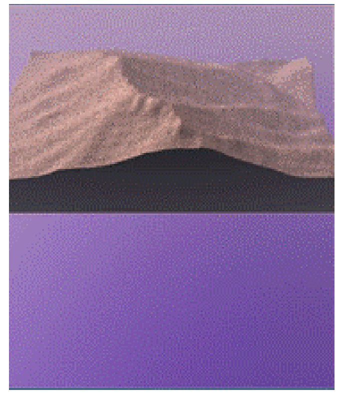


Figure 1f. Slicing terrains movie. Users change the depth of cut by clicking and dragging up and down.

users see how water rises to a level parallel to contour lines. The purpose of this mode is to clarify that contour lines represent a single elevation. Seeing the interaction of water and terrains helps students visualize basic features within an overall landscape. This interactive section allows students to set the water level at a contour line that might have previously been confusing for them. For example, not understanding how contour lines close together can represent a cliff often becomes clear when students altered the water level themselves. After students interact with each feature, they are again asked to clearly describe how the water flooded the area with the three questions below:

- Where does it flood first? Where does it flood last?
- What pattern does the water make when it is half way up the slopes?

After interacting with several flooding movies, groups are asked to verbally describe how the land would flood over time, and a sample description is given for the hill and valley but not the cliff. All of the screens up to this point represent the learning cycle exploration phase of the module. The last screen of this section defines contour lines and index contours. This represents the term introduction phase of the learning cycle.

The last mode of visualization consists of creating landscape profiles as slices are made in a terrain. Students actively change the profile by clicking and dragging up and down to slice into or build up, respectively, the terrain (see Figure 1f). The application phase of the learning cycle is then provided by showing a two-dimensional representation of a landscape with a red line drawn on it. Students are given the scenario that they want to hike along the line and shown an elevation profile that corresponds to that path. Then students are taken to several screens where they are asked to predict what the elevation profile for a different path in each of the three features would look like. Figure 2 shows several such screens. As they move to each new question, a different type of movie (increasing shading, rotating colored topographic maps, or slicing into terrains) is provided to help students determine the correct profile.

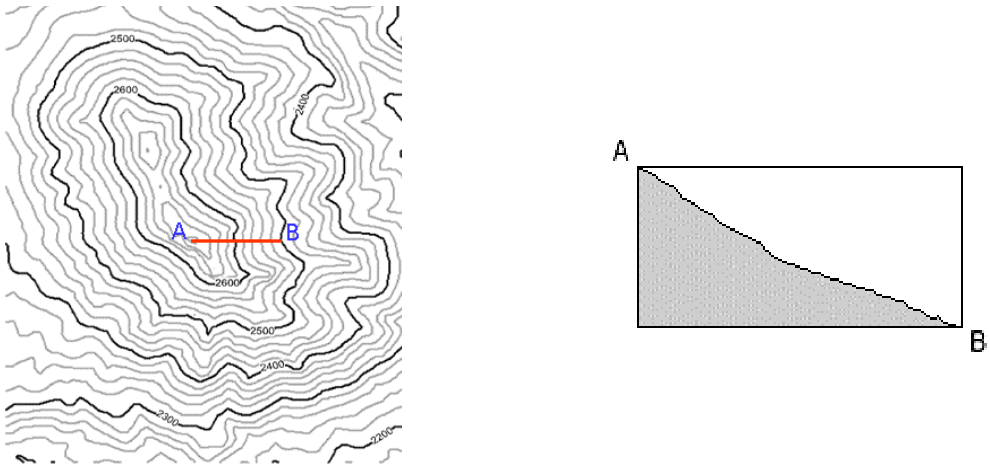


Figure 2. In the application phase of the learning cycle, students are asked to predict what profiles across the three featured landscapes would look like if they were to hike along indicated paths.

Block Diagrams Module

The interactive blocks module focused on developing students' visual penetrative ability. A crucial step in reconstructing geologic histories of an area requires the ability to sequence events from youngest to oldest. This is often done by interpreting the order in which events, such as layer deposition, folding, faulting, and intrusions, occurred. These features are often buried beneath the surface leaving only partial structures on which to base conclusions. The sections incorporated into the blocks module were designed to guide students in the visualization process of uncovering or disembedding underlying features. Techniques used to accomplish this included the rotation of blocks, making blocks partially transparent, slicing into blocks, offsetting faults, eroding the tops and sides of blocks, depositing layers, and revealing unconformities.

Since this module was created entirely in FrontPage for web distribution, the opening screen of the module contains links to instructional information. "List of Files" takes instructors to a list of individual movies used in the entire module. This allows them to access movies without entering the module. The second link, "Main Module Home", suggests students start here to receive introductory navigation and movie type information similar to the opening screens of the topographic maps module. Students are also informed in this section how the faces of blocks will be labeled (front, back, left, right, top, and bottom). The third link, "Main Module", takes students to the main menu of the module and lists the five features they can explore throughout the module. The five sections covered in the module are layers, folds, faults, intrusions, and unconformities. The fourth link takes instructors to Word and PDF files of the worksheets that accompany the module. A worksheet was developed for each of the five main sections in the module. See Figure 3 for an example of the worksheet from the layers section.

Once students reach the main menu, they can explore each feature in the order they choose. Students are informed that the topics are easiest to cover in the linear order presented in the menu, but if one section has already been covered or is too simplistic, it can easily be skipped. For the creation of this module, blocks were generated in Bryce4 as one row (rotations) or multirow (rotations in combination with other changes) movies.

Image sequences were loaded into VR Worx to generate QTVR movies. This format allows students to interact with movies to control the type and speed of changes that occur.

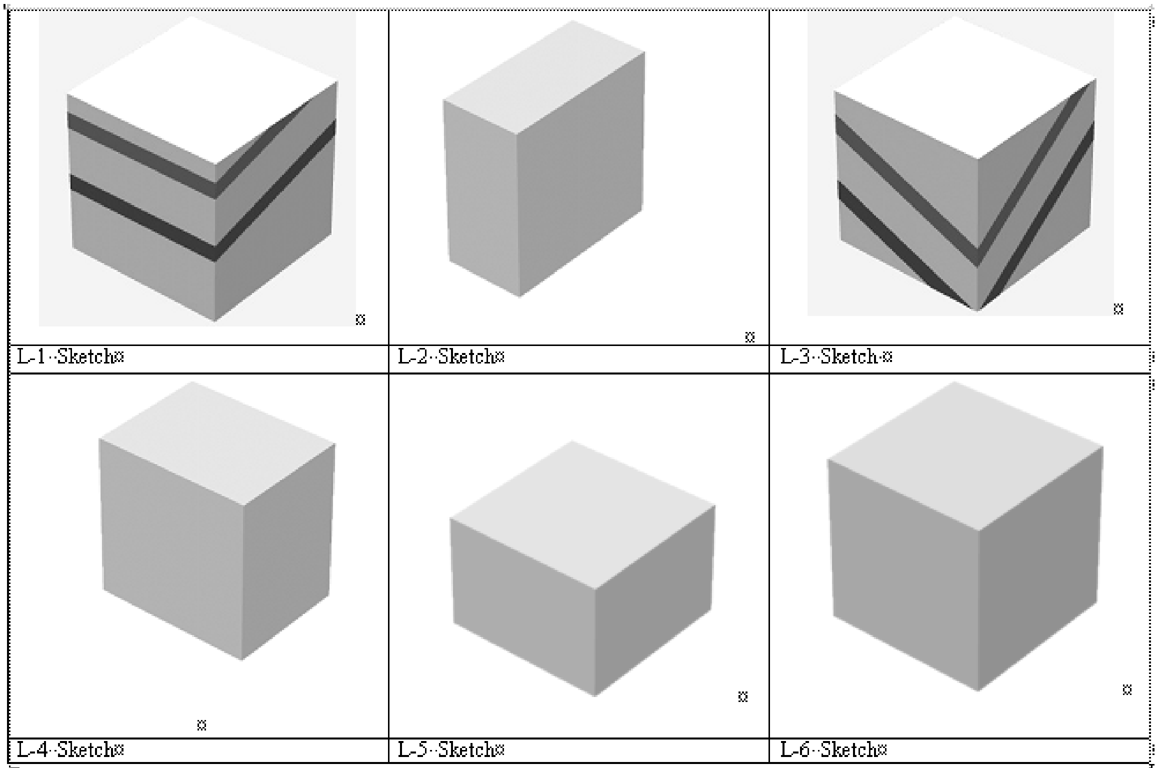
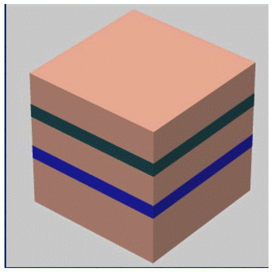


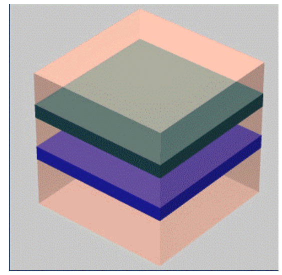
Figure 3. Layers worksheet used in the blocks module. Each block is shown on the worksheet exactly as students needed to draw it (e.g., cut in half or faces covered).

Each main menu topic contains its own submenu. For example, clicking on "Layers" takes students to a screen containing buttons to explore horizontal, gentle, moderate, steep, and vertical layers. Some sections begin with a prediction screen. Here, students are asked to predict how the layers continue from visible to hidden faces of the block. The sequence of screens after this include a rotating opaque block followed by a rotating/changing transparency block. The next screen in the section asks students to predict what the interior of the horizontal layer block looks like. Students are shown the block with a "cutting plane" intersecting it. The purpose of a cutting plane is to cut into a block and understand how subsurface features are oriented. In various movies, students can cut left to right, right to left, or top to bottom to fully understand orientations of features inside the blocks. Figure 4 shows examples of blocks from each of these screens for horizontal layers.

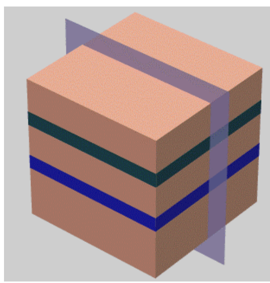
Quizzes were inserted at the end of each section so students could immediately test what they had learned. During the course of a single lab meeting, students completed one or two sections of the blocks modules. Testing after each section offered students feedback upon completion of a section and offered teaching assistants the chance to open the next lab discussion with a review of topics covered the previous day. Quizzes were designed to include a variety of questions, including multiple choice, sketches, and prediction, closely aligned to the types of questions asked throughout each section. Where possible, feedback was given for questions and movies were provided to have students verify their own answers. The last question in each quiz asks students to draw a block when given a series of geologic events.



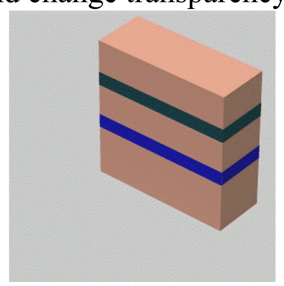
4a. Opaque block with horizontal layers students can rotate.



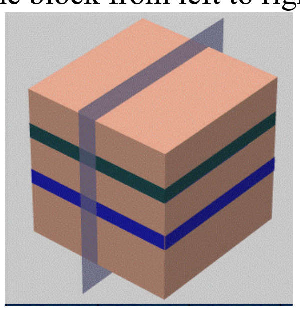
4b. Same block as 4a that students can rotate and change transparency.



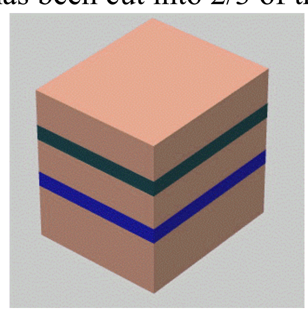
4c. Left cutting plane. Students are instructed to cut into the block from left to right.



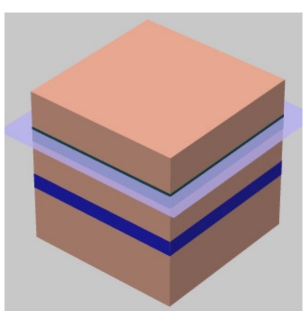
4d. Left cutting plane movie. The block has been cut into 2/3 of the way.



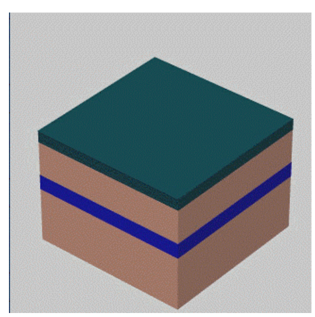
4e. Right cutting plane. Students can cut into the movie by clicking and dragging right to left.



4f. Right cutting plane movie. The block has been cut into 1/4 of the way.



4g. Top cutting plane. Students can cut into the block from top to bottom.



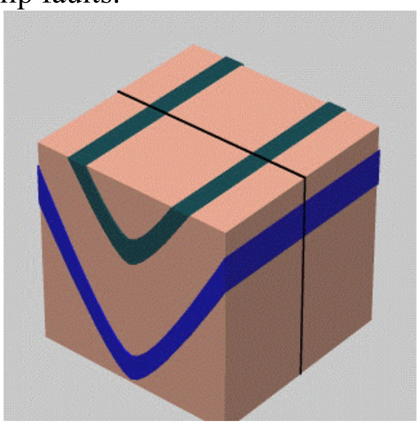
4h. Top cutting plane movie. The block has been cut into 1/3 of the way.

Figure 4. Block movies (transparency and cutting) used for the layers section. The same blocks and movies were also used throughout the folds section.

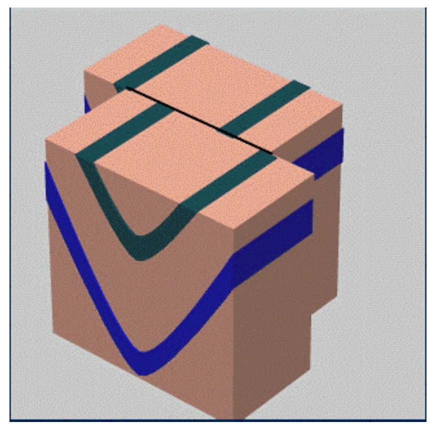
The folds section proceeds exactly as the layers section – with the same progression of screens and the same types of movies: rotations, transparency, cutting side to side and top to bottom. The five subsections of folds include horizontal anticline, horizontal syncline, plunging anticline, plunging syncline, and vertical.

The faults section contains several subsections: types of faults, layers in faults, and folds in faults. With these multiple subsections, clarity of navigation became an issue. In order to minimize student confusion when navigating, several versions of each section were developed. Each version would indicate with yellow text (rather than white) which section or screen was last visited. This helped students monitor their progress and keep track of which sections they completed.

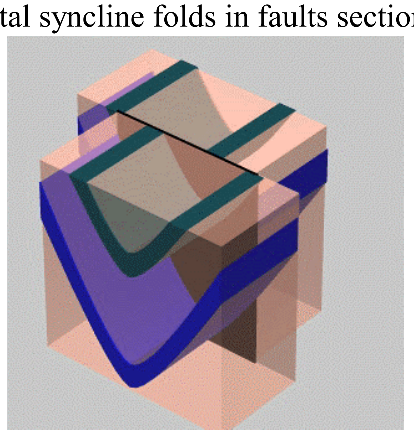
The first subsection of faults, types of faults, covered images and movies of dip-slip, strike-slip, and oblique-slip faults. Students were first given examples of the types of movies they would encounter in this section and then taken to a menu to choose what type of fault they wanted to explore. Movie types in the faults section include rotating, changing transparency, offsetting faults, and eroding surfaces in various combinations. Figure 5 shows movies before and after these changes for plunging syncline folds with strike-slip faults.



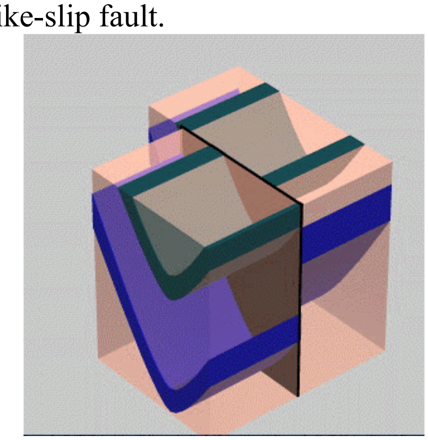
5a. Original image of opaque block with horizontal syncline folds in faults section.



5b. Same block as 5a now offset by a strike-slip fault.



5c. Same block as 5b now made partially transparent.

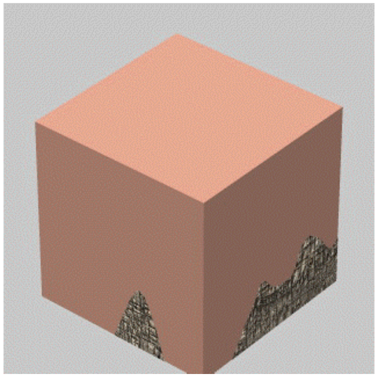


5d. Same block as 5c now eroded on the front side to make that face even.

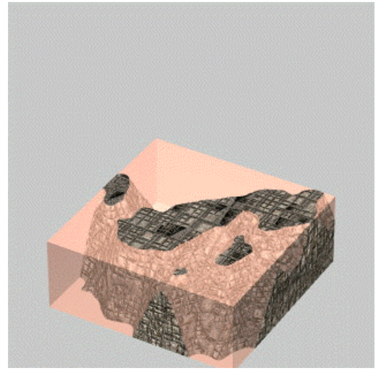
Figure 5. Four blocks showing the progressive types of movies covered in the faults section of the blocks module. These four blocks specifically show horizontal syncline folds offset by a strike-slip fault.

The next section of the module covers intrusions. The main types of movies seen here are rotations, changing transparency, and cutting from top to bottom in a block. This section begins with one intrusion type and adds another type to it. Throughout this section, only one block is shown on each screen. First, only a pluton is explored. Dikes are then added to the pluton to show students the relationship between the two. Sills are then added to the pluton and dike block. Figure 6 shows successive images of these movies. The first row shows only the pluton, the second row shows the pluton with a dike, and the third row shows a pluton, dike, and sill. The questions in this section's quiz were integrative and focused on having students reconstruct geologic histories from series of events. Students were shown rotating blocks and asked to list events in order they must have occurred. The difficulty in this task required students to identify whether faulting occurred before or after an intrusion based on the amount of offset visible on the surface.

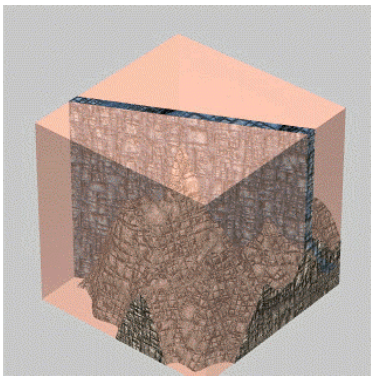
The last section covers unconformities. Students were presented with movies that revealed both horizontal and tilted unconformities. Other features from the module were included in combination with unconformities. For example, a block might contain faulted folds that were eroded and new layers deposited. Students could reveal the unconformities in this section by clicking and dragging the mouse up to examine the intersection of features between erosion and deposition. At the end of this section, and thus the end of the module, an integrated quiz was given. Questions in this quiz ask students to reconstruct a geologic history, predict what an unconformity looks like, sketch a block for a sequence of events, and interpret geologic events from an image taken in the field. Figure 7 shows a series of blocks presented in the integrated quiz section.



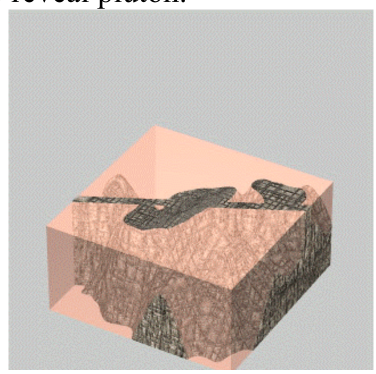
6a. Opaque block containing pluton.



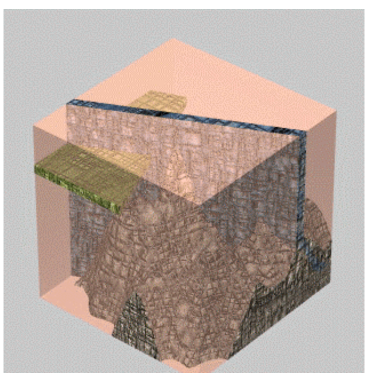
6b. Partially transparent block cut from top to reveal pluton.



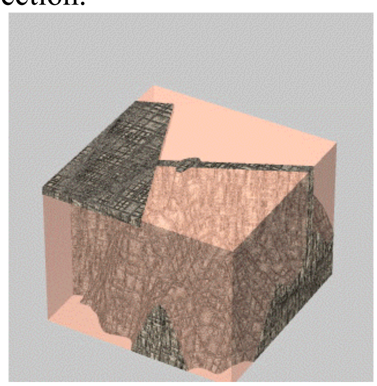
6c. Partially transparent block of pluton and dike.



6d. Partially transparent block of pluton and dike cut from top to reveal intersection.

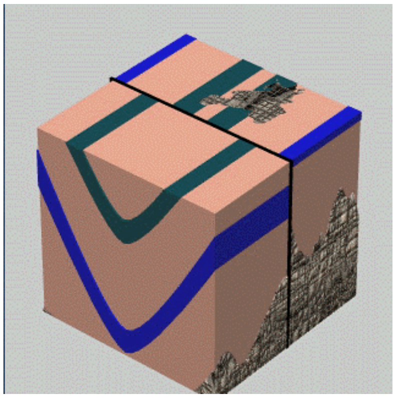


6e. Partially transparent block of pluton, dike, and sill.

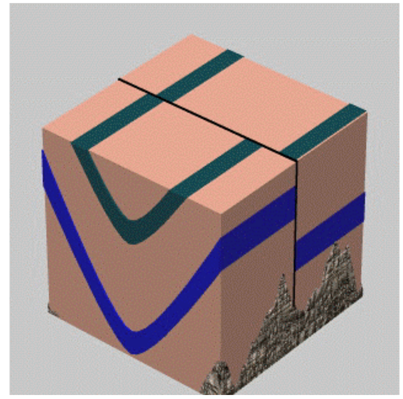


6f. Partially transparent block of pluton, dike, and sill cut from top to reveal intersection.

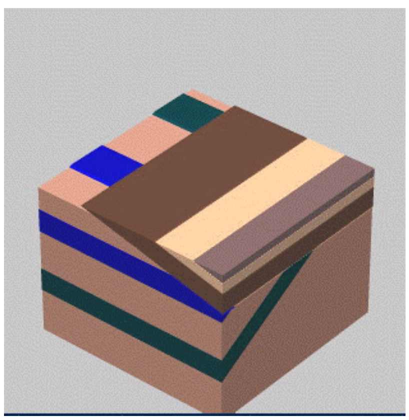
Figure 6. Blocks in intrusions section containing progressively more complex subsurface features.



7a. Integrative quiz question asking students to place the events (faulting versus intrusion) in the order they must have happened.



7b. Integrative quiz question asking students to place the events (faulting versus intrusion) in the order they must have happened.



7c. Integrative quiz question asking students to place events (tilting of layers, erosion, or unconformity) in the order they must have happened.

7d. Field-related question asking students to identify the key events that occurred to form this feature and the order in which they occurred.

Figure 7. Integrative quiz questions given at the end of the intrusions and unconformities sections.

Computer-Based Tests of Spatial Thinking

Visual-spatial thinking has been recognized as a facet of intelligence that is separate and distinct from verbal ability (Paivio, 1971, 1990; Ekstrom, French, Harmon, & Dermen, 1976; Gardner, 1983). Within the visual-spatial realm, psychometricians have identified a number of factors that contribute to spatial thinking. The Kit of Factor-Referenced Cognitive Tests (Ekstrom, et al., 1976) contains seven paper-based tests that each measure some aspect of spatial thinking. As part of a study investigating spatial thinking in college-level introductory geology class, computer-based versions were developed for two of these tests: the Surface Development Test and the Cubes

Comparisons Test (Ekstrom, et al., 1976). The Surface Development Test measures spatial visualization, the ability to manipulate a mental image while the Cubes Comparisons Test measures spatial orientation, the ability to perceive a spatial configuration from alternate perspectives.

Description of Paper Tests

In the Surface Development Test, subjects must imagine how a piece of paper can be folded into some kind of object. They are asked to compare numbered sides of the unfolded object with lettered sides of a folded object to determine which sides are the same. Figure 8 shows a sample item from the test. In Figure 1, the sides indicated by the numbers 2, 3, and 5 respectively correspond with the letters B, G, and H. The Surface Development Test contains six unfolded objects that each have five sides to be identified, resulting in a total of 30 items.

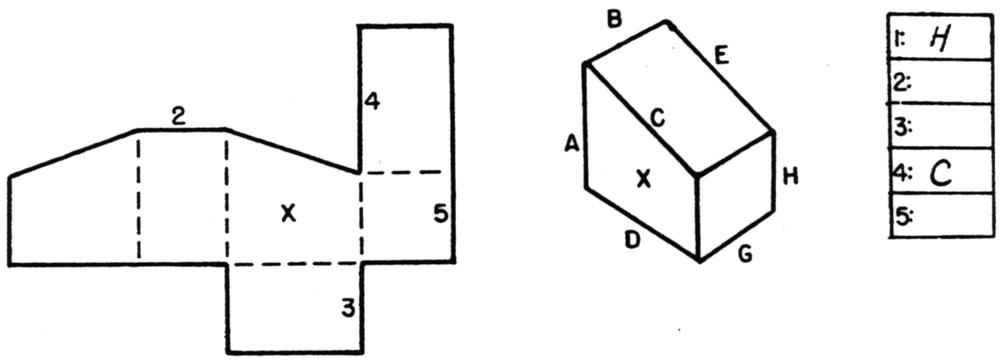


Figure 8. Sample item from the Surface Development Test from the Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976).

In the Cubes Comparisons Test subjects are given two cubes with a different letter, number, or symbol on each of the six faces. They must compare the orientation of the faces on each cube to determine if the two cubes are the same or different. Figure 9 shows a sample item from the test. The two cubes shown are not the same.

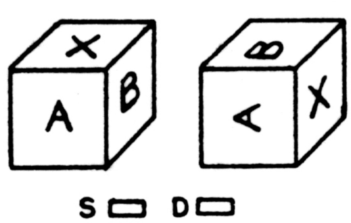


Figure 9. Sample item from the Cubes Comparisons Test from the Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976).

When the cube on the right is mentally rotated so that the face containing the "A" is in an upright position, then it can be readily seen that the face containing the "X" would now be at the bottom and would not be visible. Because no letter, number, or symbol may be repeated on any of the faces of a given cube, the "X" cannot be both on top and on the bottom of the cube. Therefore, these two cubes must be different. The Cubes Comparisons Test contains 21 pairs of cubes for a total of 21 items.

Design Considerations for Computer Tests

Creating computer-based versions of the spatial tests allowed the tests to be modified in ways that were not possible with the paper-based versions. These modifications included:

1) eliminating the verbal cues inherent in the paper tests,
2) providing animated demonstrations as part of the instructions for the tests, and
3) collecting time-to-completion measures on individual items.

In the paper version of the Surface Development Test, letters and numbers are used to identify the sides of the folded and unfolded objects, allowing subjects to indicate their response by recording letters next to numbers. In the computer version of the test, in order to eliminate these verbal cues, the sides of the unfolded objects were color-coded and both the folded and unfolded objects were hot-spot activated. One mouse click is used to select a side of the unfolded object and another is used to indicate its corresponding side. To visually show the response that has been chosen, a miniature folded object with the selected side highlighted, appears onscreen. In Figure 10, the blue side of the unfolded object has been chosen to correspond with the lower right edge of the folded object. This choice is displayed as a small diagram within the color-coded section of the answer box. In a similar manner, the brown side of the unfolded object has been chosen to correspond with the upper left edge of the folded object. As with the paper version, the computer-based version of the Surface Development Test has six unfolded objects that each have five sides to be identified, resulting in a total of 30 items.

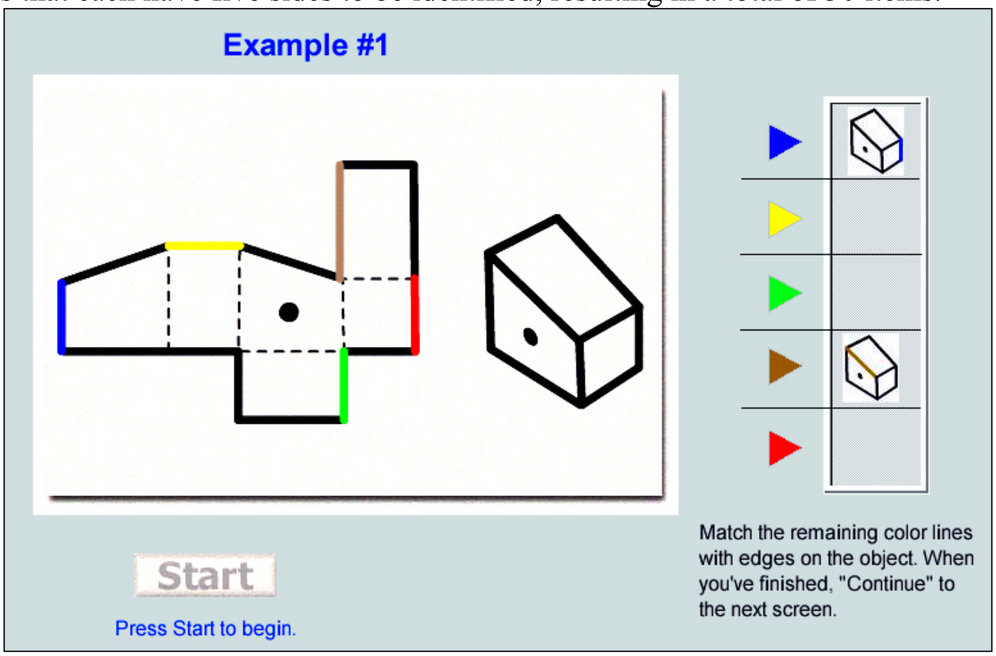


Figure 10. Sample item from a computer-based version of the Surface Development Test.

In the computer-based version of the Cubes Comparisons Test, the elimination of the verbal cueing was accomplished by replacing all the letters and numbers on the faces of the cubes with new symbols. In order to keep the computer-based items parallel to the original paper-based items, symbols were directly substituted on a one-to-one basis. For example in Figure 11, half-shaded circles in the cubes of the computer-based version replace the letter "A" on the faces of cubes of the paper version. Other substitutions include, a half-shaded triangle to replace the letter "F", a solid diamond to replace the letter "G", an open square for the letter "K", and a solid square inside an open circle for the letter "J". Whenever these letters occur on other cubes from the paper test, the same symbols are used to replace those letters. The computer-based version of the Cubes Comparisons Test contains 20 items.

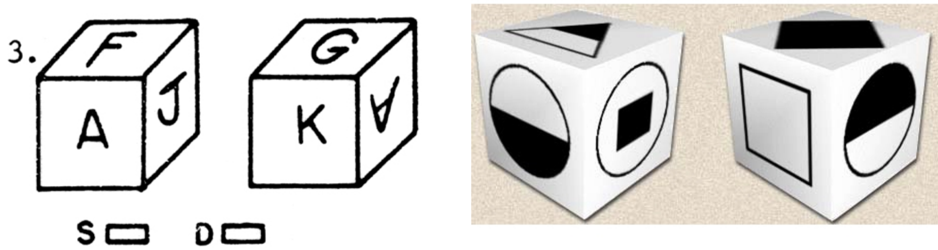


Figure 11. The paired cubes of item 3 from the paper-based (Ekstrom et al., 1976), and its equivalent computer-based, Cubes Comparisons Test.

At the beginning of both the paper and computer-based spatial tests, subjects are provided with instructions and given two sample items to make sure they understand the task that is required of them during the test. The mental operations that need to be performed in order to solve the problems on the test are verbally described to subjects. Animations that visually demonstrate these mental operations have been added to the computer-based versions. Thus, instead of simply describing in words that the right-hand cube in a pair of cubes could be rotated 90 degrees to the right, the subject sees the cube rotating 90 degrees via animation. In a similar manner, as an introductory example to the computer-based version of the Surface Development Test, an unfolded object folds up and then spins around to reveal the object from a 360 degree perspective. Thus, subjects taking the computer-based versions of the tests view animations that demonstrate the spatial tasks they need to perform during the tests.

The paper versions of the two spatial tests are administered with time limits. In many cases, subjects do not complete all the items on the test during the allotted time. In other cases, subjects complete all items before the time limit ends. How long it takes subjects to complete all items is difficult to measure. The ability to collect such time-to-completion data has been embedded within the computer-based versions of the tests. Whenever a subject is presented with an item on the test, he or she must click a start button to reveal the item. Clicking this start button activates a timer. When the subject leaves that screen, the timer stops. A total time-to-completion can be calculated by adding all the times for the individual items.

The decision to remove the time limit on the computer-based versions of the spatial tests was made in order to investigate basic patterns of performance. Time limits on spatial tests have implications for gender differences. On timed tests of mental rotation, male scores are consistently and significantly higher than that of females (Kimura, 1983; Linn and Peterson, 1985; Voyer, Voyer, & Bryden, 1995; Dabbs, Chang, Strong, Milun, 1998). However, there is some evidence to suggest that time, rather than ability per se, may be the differentiating factor in spatial tasks that involve mental rotations (Kail, Carter, & Pellegrino, 1979; Linn & Peterson, 1985).

Comparing Paper and Computer Tests

Scores on the paper and computer-based tests as well as time-to-completion data on the computer-based versions were analyzed to determine how their distributions relate to one another. Table 1 shows the correlation matrix for both versions of both tests.

Table 1: A correlation matrix for paper versions, computer versions, and time-to-completion on computer versions for two measures of spatial ability

Spatial Measure	SurDev computer	SurDev time	SurDev paper	Cubes computer	Cubes time	Cubes paper
Surface Development computer-based	1.00					
Surface Development time-to-completion	.471**	1.00				
Surface Development paper	.510**	-.148	1.00			
Cubes Comparisons computer-based	.536**	.048	.524**	1.00		
Cubes Comparisons time-to-completion	.121	.293*	-.098	.024	1.00	
Cubes Comparisons paper	.144	-.220	.364**	.245**	-.440**	1.00

Correlations that are significant at the .01 level are indicated by **
Correlations that are significant at the .05 level are indicated by *

The matrix shows that both spatial tests correlate with their computer-based versions. Moderate, but significant, correlations occur between the two computer-based versions and the two paper versions. The highest correlations exist with the computer-based version of the Surface Development Test. The scores on this test moderately correlate with the scores on the paper version, as well as with the scores on the computer-based Cubes Comparisons Test. A moderate negative correlation is found for the time-to-completion on the computer-based Cubes Comparisons Test and scores for the paper version of the Cubes Comparisons Test.

Cubes on Paper verses Cubes on Computer

An item analysis for the parallel versions of the Cubes Comparisons Test is shown in Table 2. For the paper version, the most difficult item appears to be number 21, with only 12% of the subjects responding. However, 73% of the students skipped this question, probably as a direct result of running out of time. For the paper test, item difficulty increases as the number of items skipped by subjects increases.

Table 2. Test statistics for the Paper and Computer versions of the Cubes Comparisons Test

Item Number	**Paper-based Cubes Test**					**Computer-based Cubes Test**						
	C	I	S	Mean (N = 146)	Reliability	C	I	S	Mean (N = 147)	Reliability	Average Time (sec)	Reliability
1	119	25	3	.8082	.7967	122	25	0	.8299	.5124	9.57	.8977
2	126	19	2	.8562	.7897	129	18	0	.8776	.5276	8.28	.8993
3	107	34	6	.7329	.7928	70	70	0	.4762	.5614	12.81	.8931
4	86	55	6	.5822	.7932	114	32	1	.7755	.5017	13.03	.8970
5	105	35	7	.7192	.7840	116	31	0	.7891	.4858	8.63	.8948
6	100	38	9	.6781	.7878	95	50	2	.6463	.5105	13.84	.8934
7	119	22	6	.8082	.7833	120	25	2	.8163	.4956	9.35	.8949
8	102	36	9	.6918	.7825	136	11	0	.9252	.5054	7.61	.8952
9	98	35	14	.6644	.7815	115	32	0	.7823	.5115	13.7	.8944
10	119	8	20	.8082	.7720	112	35	0	.7619	.5104	12.0	.8940
11	91	24	32	.6164	.7694	137	9	1	.9320	.5178	11.7	.8976
12	54	47	46	.3699	.7751	87	60	0	.5918	.5370	12.47	.8914
13	36	51	60	.2466	.7756	80	66	1	.5442	.5063	11.64	.8937
14	74	14	59	.5068	.7624	132	15	0	.8980	.5154	8.82	.8950
15	43	34	70	.2945	.7687	119	28	0	.8095	.4840	8.25	.8940
16	52	10	85	.3493	.7566	120	25	2	.8163	.4889	9.95	.8943
17	52	7	88	.3493	.7636	112	34	1	.7619	.5040	8.89	.8951
18	52	1	94	.3562	.7636	84	63	0	.5714	.6123	11.17	.8954
19	37	12	98	.2534	.7695	129	17	1	.8776	.5241	10.69	.8971
20	40	6	101	.2671	.7688	138	8	1	.9388	.5144	8.02	.8996
21	18	21	108	.1233	.7783							

C refers to the number of students selecting the correct response. I refers to the number of students selecting an incorrect response. S refers to the number of students skipping an item. The mean score reflects the difficulty level of an item.

The easiest item on the computer-based version was item 20. Whether the cubes are the same or different can be determined by using visual inspection, instead of rotation. Appendix A contains screen shots of each of the cube pairs created for the computer-based version. One of the least difficult items on both tests was item two, which also only requires visual inspection to solve. The most difficult item on the computer-based version was item 13. To solve this problem, one of the cubes must be rotated twice: 90 degrees on the x-axis and 90 degrees on the y-axis. Alternatively, a 180-degree flip along the z-axis also brings a cube into the necessary comparative position. Overall, the difficulty levels on items on both tests are very similar for the first half of the test. The difficulty levels diverge when subjects begin to run out of time to complete the paper-based version.

Surface Development on Paper verses Surface Development on Computer

The items on the Surface Development Test were not constructed in the same parallel fashion as with the Cubes Comparisons Test; therefore comparisons across items on the two tests cannot be made. Unlike with the paper-based Cubes Comparison Test, there is no general increase in difficulty as the paper-based Surface Development test progresses. In other words, difficult items are scattered throughout the test.

Table 3. Test statistics for the Paper-based Surface Development Test

Item Number	Paper-based Surface Development Test	
	Mean (N = 155)	Reliability
1	.6968	.9133
2	.8194	.9141
3	.5935	.9127
4	.8129	.9154
5	.6323	.9142
6	.7484	.9132
7	.5355	.9133
8	.5419	.9135
9	.7742	.9136
10	.5935	.9137
11	.8452	.9156
12	.8387	.9146
13	.4839	.9124
14	.3161	.9140
15	.7097	.9165
16	.6323	.9119
17	.5355	.9154
18	.4581	.9154
19	.2516	.9176
20	.5032	.9102
21	.1355	.9152
22	.3097	.9118
23	.3677	.9121
24	.4129	.9130
25	.3677	.9109
26	.3419	.9138
27	.4065	.9115
28	.3935	.9112
29	.4000	.9113
30	.4194	.9106

The mean score reflects the difficulty level of an item.

Table 4. Test statistics for the Computer-based Surface Development Test

Item Number	Computer-based Surface Development Test			
	Mean (N = 147)	Reliability	Average Time (sec)	Reliability
1	.7551	.9374	12.93	.9570
2	.7823	.9370	12.93	.9570
3	.7347	.9368	12.93	.9570
4	.7551	.9364	12.93	.9570
5	.6939	.9378	12.93	.9570
6	.8231	.9380	11.27	.9562
7	.6259	.9366	11.27	.9562
8	,6599	.9371	11.27	.9562
9	.8095	.9367	11.27	.9562
10	.7279	.9367	11.27	.9562
11	.8163	.9375	11.43	.9556
12	.8571	.9377	11.43	.9556
13	.5986	.9390	11.43	.9556
14	.0272	.9420	11.43	.9556
15	.0748	.9420	11.43	.9556
16	.7619	.9376	12.94	.9552
17	.5442	.9397	12.94	.9552
18	.5646	.9391	12.94	.9552
19	.2245	.9429	12.94	.9552
20	.5578	.9376	12.94	.9552
21	.4490	.9411	14.09	.9553
22	.2517	.9469	14.09	.9553
23	.0544	.9425	14.09	.9553
24	.7619	.9381	14.09	.9553
25	.6803	.9378	14.09	.9553
26	.0748	.9436	9.98	.9557
27	.7551	.9370	9.98	.9557
28	.7279	.9368	9.98	.9557
29	.6871	.9376	9.98	.9557
30	.7619	.9369	9.98	.9557

The mean score reflects the difficulty level of an item.

Reliability Data for all Four Tests

The reliabilities for each of the measures of spatial ability are listed in table 5 below.

Table 5: Overall Test Reliabilities for Spatial Measures

Spatial Measure	Number of Subjects	Number of Items on Test	Reliability
Surface Development computer-based	147	30	.9408
Surface Development time-to-completion	147	30	.9573
Surface Development paper	155	30	.9160
Cubes Comparisons computer-based	147	20	.5305
Cubes Comparisons time-to-completion	147	20	.9001
Cubes Comparisons paper	146	21	.7857

The Geospatial Test

The dependent variable in this experiment was a 30 item, multiple choice test that containing content from the geology laboratory that was judged to be spatial in nature. Although a paper-and-pencil test, the stems of all items included diagrams or pictures.

A pilot version of the instrument was administered to students who were just completing an introductory geology class at a local community college. Based upon the results of that administration, a final version of the test was prepared.

The data in Table 6 were obtained from the pretest administration of the instrument during the experiment.

Table 6. Item analysis and content of Geospatial Test

ITEM	Difficulty	Discrimination Index	Content
1	0.495	0.1	Finding point on map.
2	0.693	0.4	Finding point on map.
3	0.485	0.2	Finding point on map.
4	0.703	0.4	Finding point on map.
5	0.653	0.3	Identifying perspective.
6	0.525	0.4	Identifying perspective.
7	0.733	0.1	Identifying perspective.
8	0.733	0.6	Cross-section.
9	0.465	0.6	Cross-section.
10	0.594	0.6	Cross-section.
11	0.594	0.4	Sequence of events.
12	0.782	0.4	Sequence of events.
13	0.762	0.4	Sequence of events.
14	0.673	0.4	Sequence of events.
15	0.277	0.4	Sequence of events.
16	0.535	0.5	Sequence of events.
17	0.515	0.7	Sequence of events.
18	0.822	0.4	Block diagram.
19	0.713	0.5	Block diagram.
20	0.822	0.2	Block diagram.
21	0.584	0.7	Block diagram.
22	0.723	0.4	Block diagram.
23	0.653	0.6	Map problem.
24	0.733	0.6	Map problem.
25	0.594	0.3	Map problem.
26	0.723	0.6	Map problem.
27	0.347	0.3	Map problem.
28	0.307	0.2	Map problem.
29	0.762	0.3	Map problem.
30	0.762	0.6	Topographic profile

The K-R 20 reliability for the entire Geospatial Test was 0.75 on the pre-test and 0.78 on the post-test.

THE EXPERIMENT

Design of the Project

The project was to create and evaluate a group of computer-based modules for college-level instruction in geology. These were appropriate for use in introductory laboratories as well as upper division courses for geology majors. The materials focused on exposing "The Hidden Earth," presenting problems involving the surface expression of structural features and the shallow structure of the earth's interior. The modules were situated in complex, real-life problems and activities that are characteristic of the practice of geology, and its associated reasoning (Frodeman, 1995; Ault, 1998; Drummond, 1999).

Computer-based materials were built with the program Bryce3D. This program allows the creation of detailed and realistic, two-dimensional representations depicting three-dimensional perspectives of simple and complex geologic structures and landscapes. The 3D models can be rotated, sectioned, disassembled, or successively unburied. A series of images can be used to depict sequential geologic histories, such as deposition of successive layers, followed by erosion into realistic-looking landscapes. This approach is an analog of strategies that have been shown in previous research to be effective in the development of spatial reasoning.

This project sought to embed spatial learning in the context of real-life, complex problems that are authentic. They were taken from among actual problems that geologists deal with in everyday life. The expectation was that this would increase the development of spatial ability and improve the transfer to relevant problem solving. This hypothesis was to be tested in a quasi-experimental design in which control and experimental groups are administered a content assessment and two spatial/visual measures as pre- and post-tests.

The Context

The experiment was conducted during the first Arizona State University summer session, beginning on Tuesday, May 29 and ending on Friday, June 29, 2001. This consisted of 5 weeks of classes, meeting 1 ½ hours per day. Two sections met from approximately 7-9 a.m. and two from 11 a.m.-1 p.m.

Subjects were students in Geology 103, a one credit-hour introductory geology laboratory. Although Geology 103 is associated with the lecture course Geology 101, "Introduction to Geology," concurrent enrollment is not required, and the content of the lecture and the laboratory are not coordinated. The laboratory course enrolled approximately 100 students divided among four sections.

Four sections of Geology 103 were taught, each by a different graduate teaching assistant. Two sections each were assigned to either the control or experimental condition. To eliminate time-of-day effects, a control and experimental group were assigned to each starting time. Teaching assistants were fully briefed on the nature of the experiment, and members of the research team met with them weekly to discuss the nature of the experimental and control conditions. Members of the research team also observed both control and experimental classes on a regular basis to ensure that the experimental conditions are being met.

Both control and experimental classes studied from a laboratory manual written by Stephen J. Reynolds, Julia K. Johnson and Edmund Stump, titled "Observing and Interpreting Geology (2001)." This manual covers the traditional content of an introductory geology laboratory in an unconventional manner. The first seven chapters are anchored in a series of computer simulations created in a virtual environment called "Painted Canyon." In these chapters, students are introduced to topographic maps, minerals and rocks, geologic maps and geologic history, and environmental issues. Chapters 8 through 11 are devoted to the geology of selected regions of Arizona, and lead to a field trip at a location near the University. The final three chapters engage students in a study of the geology of their own home town, the exploration of a geological setting in a virtual environment, an evaluation of the economic potential of selected mineralized areas, and the fossils of the Colorado Plateau.

Two unique computer-based measures of spatial orientation and spatial visualization were created for this study. These were created by this research group as modifications of instruments contained within the "Kit of Factor-Referenced Cognitive Tests" by Ekstrom, et al. (1976). The dependent variable was a geospatial assessment based upon the content of the laboratory manual.

The geospatial assessment was administered as a paper-and-pencil test to all students in all sections on the first and the last days of the first summer session. They were told that their grade would depend in part on their performance on the second content assessment. The two computer-based spatial measures were administered to all students during the first and last weeks of the first summer session. Subjects were removed in groups of ten to an adjacent laboratory for computer-based testing. It required less than two days to complete this phase of the assessment.

The experiment was a quasi-experimental pre-test/post-test design with control and experimental groups. Analysis of Variance was used to test the hypotheses that there are no initial differences among experimental and control groups on any pre-test measure, and that the experimental groups perform at a significantly higher level than the control groups on all post-test measures. Step-wise multiple regression analysis was used to estimate the amount of variance in achievement that is shared with spatial measures.

RESULTS

Sample Distribution

The sample consisted of 103 subjects, of whom 48 were male and 55 were female. The groups were unequal in size, with 44 subjects in the control group and 59 in the experimental group. Although subjects self-selected into individual sections of the course, the distribution of by gender across the sections was not random (Table 7). Males exceeded females in the control group by a factor of 1.4/1 and females exceeded males in the experimental group by a factor of 1.7/1.

Table 7. Distribution of Subjects by Gender and Group

Group	Male	Female	Total
Control	26	18	44
Experimental	22	37	59
Total	48	55	103

This unusual sample bias is a classic example of the difficulties of quasi-experimental design with intact groups. The normal assumption of a quasi-experimental design of the sort used in this study is that the comparison groups will be equivalent. This has not turned out to be the case in this instance. As will be shown in the analyses that follow, initial mean scores on many variables were lower for females than for males.

This has led to a set of results in which initial mean scores of the experimental group tend to be significantly lower than those of the control group.

Attrition rates were relatively high. Only 89 students took both the pretest and the final examination for the course. In addition, many students failed to complete one or more of the spatial measures. The number of students completing each measure will be indicated in the analyses that follow.

The Geospatial Test

The effects of the experiment are analyzed through the application of a three-way Analysis of Variance. In this analysis, the dependent variable is performance on the Geospatial Test. SCORE reflects differences in performance from pretest to posttest, and is treated here as a repeated measure. CONDITION refers to control vs. experimental group, and GENDER to males vs. females. The results are shown in Table 8.

Table 8. Three-way Analysis of Variance (SCORExCONDITIONxGENDER) of Scores on Geospatial Test

	F	df	p
SCORE	161.266	1, 85	0.00*
SCORExCONDITION	3.844	1, 85	0.05*
SCORExGENDER	4.853	1, 85	0.03*
SCORExCONDITIONxGENDER	0.213	1, 85	0.65

There was a significant main effect for SCORE, with higher posttest than pretest scores for the entire sample. There were significant two-way interactions between SCORE and CONDITION, and between SCORE and GENDER. There was no significant three-way interaction.

In order to assess the magnitude of the experimental effect, normalized gain scores were computed for each student. Often referred to in the Physics Education literature as "Hake Scores," these reflect the increase from pretest to posttest score as a percentage of the total possible increase (normalized gain = posttest-pretest/total possible-pretest). The results are displayed as histograms in Figure 12.

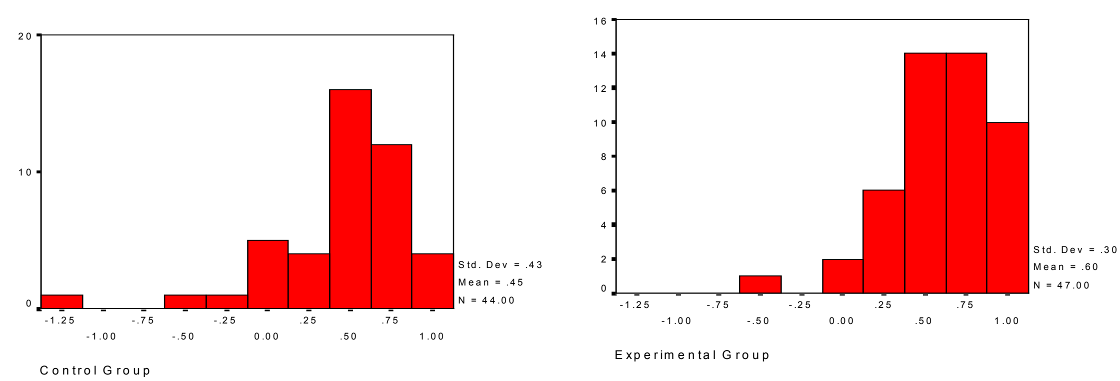


Figure 12. Normalized Gain Scores of Experimental and Control Groups on the Geospatial Test.

The mean control group gain scores were 0.45 (45%), and the distribution remained normally distributed. In contrast, mean experimental group gain scores were 0.60 (60%) and badly skewed as a result of a ceiling effect. A large number of students in the experimental group achieved gains in the upper ranges, 75% and above. If the Geospatial Test had been somewhat more difficult, it is likely that the distribution of experimental group scores would also have been normal, and the differences between the means even greater.

Pretest mean scores of the control group were lower than those of the experimental group, whereas posttest mean scores were approximately equal. This undoubtedly resulted from the unequal distributions of males and females in the control and experimental groups and differences in their performance on the Geospatial Test. The experimental treatment thus had the effect of equalizing previously unequal scores between the control and experimental groups, and demonstrating the effectiveness of the experimental materials. It also had the effect of equalizing initial differences in performance between males and females.

Normalized gain scores for the entire sample are displayed separately by gender in Figure 13. They are considerably larger for females (56%) than for males (48%). While there is a slight ceiling effect for females, it is not as dramatic as the earlier example.

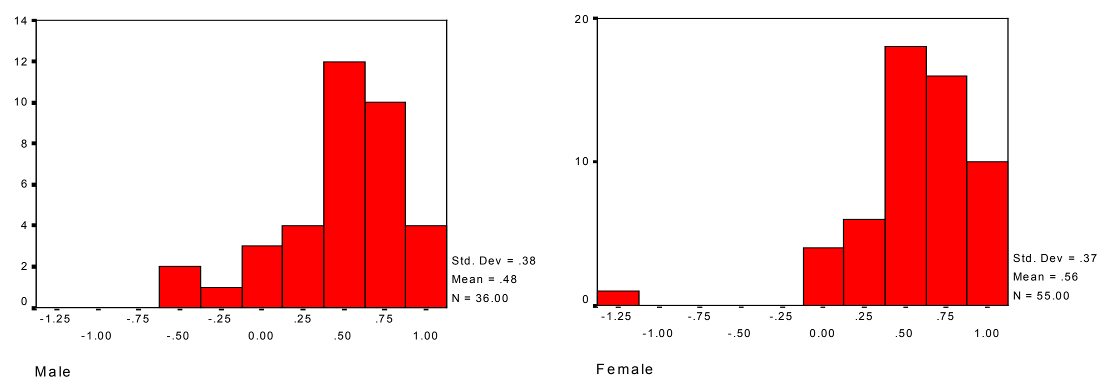


Figure 13. Normalized Gain Scores of Males and Females on the Geospatial Test.

Descriptive statistics for the sample of 89 students who took the Geospatial Test as both a pre-test and post-test are given in Table 9. These permit a more detailed comparison of male and female performances in the control and experimental groups.

Figure 14 demonstrates the importance of gender as a variable in performance on the Geospatial Test. Females in both the control and the experimental groups experienced greater growth in their Geospatial Test scores from pretest to posttest than did males. Although the effect was smaller, both males and females in the experimental group showed greater improvement than those in the control group. These results are exactly what were expected from the observation of a CONDITION x GENDER interaction.

Table 9. Descriptive Statistics for Sample Performance on Geospatial Test

PREGEOSPATIAL		MEAN	S.D.	n
Control	Male	21.67	4.50	24
	Female	18.61	5.01	18
	Total	20.36	4.91	42
Experimental	Male	20.67	3.45	12
	Female	16.63	4.73	35
	Total	17.66	4.75	47
Entire Sample	Male	21.33	4.15	36
	Female	17.30	4.87	53
	Total	18.93	4.99	89
POSTGEOSPATIAL				
Control	Male	25.79	3.05	24
	Female	24.33	5.04	18
	Total	25.17	4.04	42
Experimental	Male	26.17	3.43	12
	Female	24.57	3.88	35
	Total	24.98	3.80	47
Entire Sample	Male	25.92	3.14	36
	Female	24.49	4.26	53
	Total	25.07	3.89	89

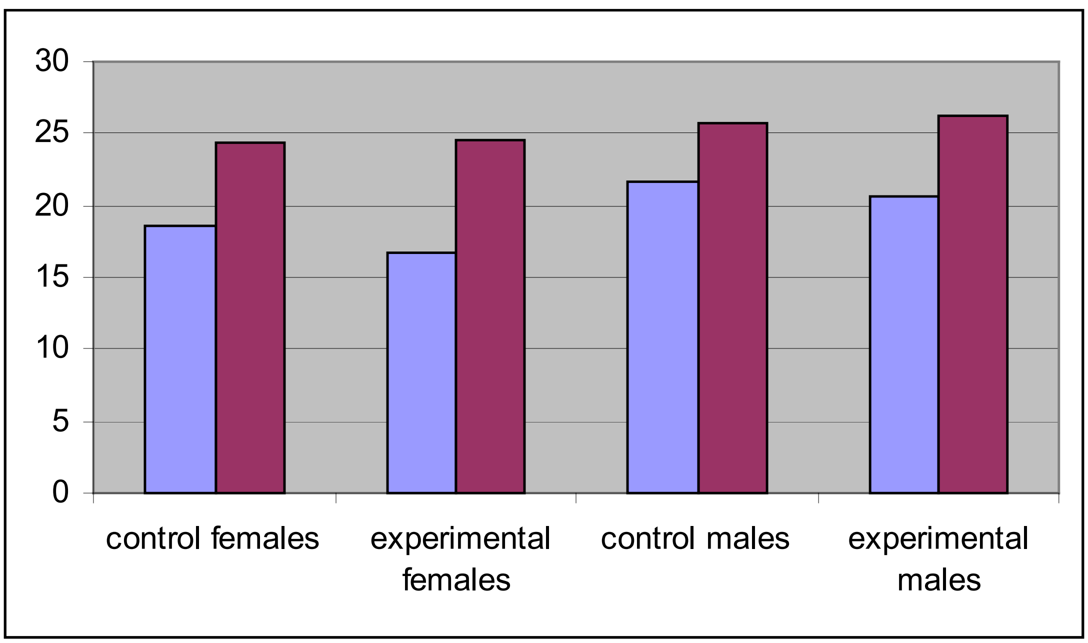


Figure 14. Pretest (left) and Posttest (right) Means of Males and Females in Experimental and Control Groups on the Geospatial Test.

The Spatial Measures

Measures of two types of spatial ability were given to all subjects as pretests and posttests. These were spatial orientation and visualization. Two values of each type of ability were generated for each instrument. The first was for the total score and the second for the time to completion.

A three-way Analysis of Variance revealed no significant main effect or interactions for the total score on the measure of spatial orientation. There was a significant main effect for time to completion (F = 16.956, df = 1, 82, p =0 .00), but there were no interactions with either CONDITION or GENDER. All subjects, both male and female in both the control and the experimental groups, showed improved time to completion on this measure.

The results for spatial visualization were somewhat different (Table 10). In this analysis, SCORE refers to the test of spatial visualization administered as a repeated measure, CONDITION refers to control versus experimental groups, and GENDER to males versus females. There was a significant main effect for SCORE, and a significant interaction between SCORE and CONDITION. There were no interactions between SCORE and GENDER, nor were there any three-way interactions.

Table 10. Three-way Analysis of Variance (SCORExCONDITIONxGENDER) on Total Score on Spatial Visualization Measure

	F	df	p
SCORE	4.533	1, 82	0.04*
SCORExCONDITION	6.830	1, 82	0.01*
SCORExGENDER	1.096	1, 82	0.30
SCORExCONDITIONxGENDER	0.618	1, 82	0.43

As demonstrated in Figure 15, the effect of the experiment was to equalize initial differences in spatial ability between the two groups. On the pretest, experimental group visualization scores were much lower than those for the control group, whereas on the posttest the scores of the two groups were quite similar. Because there was no significant interaction between SCORE and GENDER, it appears that the effect was about the same for females as for males.

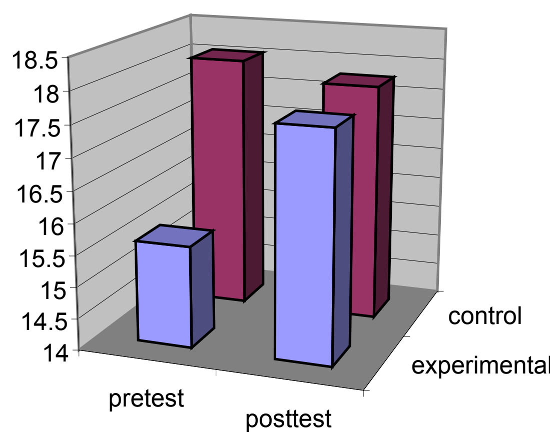


Figure 15. Pretest and Posttest Mean Total Scores of Experimental and Control Groups on the Spatial Visualization Measure.

This was not the case for time to completion on the test of spatial visualization Table 11). In this instance, there was a significant main effect for time to completion, with students completing the posttest more quickly than the pretest, and a significant interaction between SCORE and GENDER. There was no significant interaction between SCORE and CONDITION nor was there a significant three-way interaction.

Table 11. Three-way Analysis of Variance (SCORExCONDITIONxGENDER) for Time to Completion on Spatial Visualization Measure

	F	df	p
SCORE	75.899	1, 82	0.00*
SCORExCONDITION	2.199	1, 82	0.14
SCORExGENDER	5.683	1,82	0.02*
SCORExCONDITIONxGENDER	.115	1, 82	0.74

Figure 16 shows the effects of gender on time to completion. In this case, males began the experiment with somewhat longer times to completion than females, and the two groups were about the same at the end.

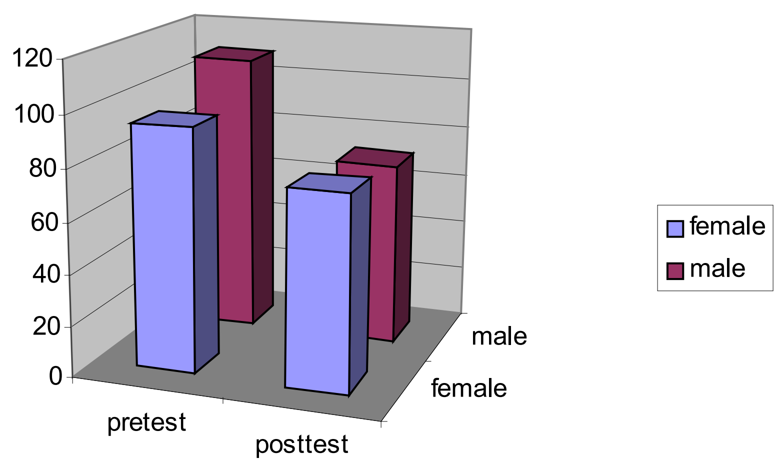


Figure 16. Pretest and Posttest Mean Times to Completion of Females and Males on Spatial Visualization Measure.

Spatial Ability and Achievement

The matrix of correlations between measures of spatial and geospatial ability is presented in Table 12. There are no significant correlations among measures of time to completion of spatial orientation or visualization. Because of this, time was eliminated as a variable in further analyses. However, the correlations between total scores on the spatial and geospatial measures are quite high, ranging from a low of 0.39 to a high of 0.57. This reflects shared variances (r-square) averaging 19% for spatial orientation and 29% for spatial visualization between the spatial and geospatial measures.

Table 12. Coefficients of Correlation Among all Variables

		1	2	3	4	5	6	7	8	9
1	PreOrientation -score	1.00								
2	PostOrientation -score	0.72*	1.00							
3	PreOrientation -time	-0.02	-0.06	1.00						
4	PostOrientation -time	-0.10	-0.01	0.81*	1.00					
5	PreVisualization -score	0.62*	0.59*	0.01	-0.05	1.00				
6	PostVisualization -score	0.59*	0.55*	0.11	0.09	0.84*	1.00			
7	PreVisualization -time	0.04	-0.01	0.45*	0.36*	-.20	0.31*	1.00		
8	PostVisualization -time	-0.10	-0.08	0.39*	0.47*	-0.03	0.15	0.64*	1.00	
9	PreGeospatial -score	0.46*	0.42*	-0.03	-0.16	0.57*	0.49*	0.00	-0.21	1.00
10	PostGeospatial -score	0.39*	0.48*	0.15	0.07	0.55*	0.55*	0.03	-0.06	0.57*

*p = 0.05

Because students entered the course with a good deal of prior geospatial knowledge, and because of the correlations between spatial and geospatial ability, it was necessary to estimate the amount of variance in posttest geospatial scores that was shared with spatial scores after the contribution of initial ability had been co-varied. In order to accomplish this, a Stepwise Multiple Regression Analysis, with pretest Geospatial scores entered as a covariate at the first step, was completed (Table 13). Prior knowledge, as

Table 13. Regression of Posttest Geospatial Scores Against Pretest Scores of Spatial Orientation and Visualization and of Geospatial Ability

	B	Std. Error	Beta	t	probability
(Constant)	16.009	1.992		8.306	0.000
Geospatial	.291	0.085	0.373	3.433	0.001
Orientation	4.4E-02	0.162	0.032	0.275	0.784
Visualization	0.173	0.074	0.296	2.339	0.022

measured by the Geospatial Test, and initial ability at spatial visualization achieved significant Betas in this analysis. The Beta for pretest scores on the spatial orientation measure did not reach the level of statistical significance.

The variance shared between the posttest geospatial ability and all pretest variables of spatial and geospatial ability was 38.4% (r=.620). The relative influence of the separate factors in the equation can be evaluated by comparing Beta weights, or standard partial regression coefficients, of the independent variables. Such a partial coefficient expresses the change in the dependent variable due to a change in one independent variable with the remaining variables held constant. In any regression, Beta weights are the same regardless of the order in which the variables are entered.

Both prior knowledge and visualization ability contributed significantly to the equation predicting posttest Geospatial Test scores. Although the Beta for prior knowledge was somewhat higher than the Beta for spatial visualization, the two are similar enough to state that as a first order approximation the two contribute equally to the regression equation.

Summary

Although all subjects profited from both the control and the experimental conditions, the effectiveness of the treatment experienced by the experimental group has been confirmed. Using both Analysis of Variance and a comparison of normalized gain scores, it has been demonstrated that students in the experimental group profited more than those in the control group.

Very powerful gender effects have also been demonstrated. The experiment had the result of equalizing the performance of males and females in a case where the performance of males was initially superior to that of females. Again, although females profited from both treatments, it appears that the experimental condition was slightly preferable.

There was little effect on the abilities of students in spatial orientation as a result of either condition, nor did this variable affect achievement. This was not, however, the case for spatial visualization. The experimental treatment was very effective at improving scores and lowering times to completion. In this instance, the performance of males appears to have been differentially improved over that of females. A regression of performance on the posttest Geospatial Ability measure against pretest variables showed that the normalized regression coefficients for prior knowledge and visualization ability were quite similar.

DISCUSSION

Reform of science education must be predicated on research on learning and teaching materials and practices that are developed from that research.

Geoscience Education Working Group (1997)

This project demonstrates that spatial ability can be improved through instruction, leading to improved learning, and that differences in performance between the genders can be eliminated with such an intervention. This result was reached through the creation and application of a set of innovative, computer based materials that can be widely used in introductory laboratory courses at colleges and universities. But more important, this study provides evidence from a naturalistic setting that demonstrates the effectiveness of those materials.

Spatial orientation and visualization are commonly understood as factorially distinct mental abilities. In this study, participants improved in visualization, but not in spatial orientation. In addition, visualization is a significant predictor of the amount learned, but spatial orientation is not. Even more important is the finding that visualization and prior knowledge have approximately equal predictive power in a regression equation against post-test knowledge scores. This may be the strongest demonstration yet of the potency of spatial ability in facilitating learning, and of the importance of being able to visually transform an image to the nature of that learning process.

Because of time limitations and difficulties with preparing computer-based materials, we limited our inquiry to the most obvious and well-known examples of spatial ability. Even then, questions remain about the nature of spatial orientation and visualization, and how these interact with student learning. The observation of significant correlations is interesting, but we must now move forward to an explanation of how students manipulate images and use that information to generate knowledge. We expect that this answer will not be reached through quasi-experimental studies such as this one. In fact, we hope to soon begin a series of studies of a more qualitative nature in which the question of how students use images to negotiate meaning is addressed.

At least two other important spatial factors remain unexamined in our study. The first is the process of "disembedding" or "restructuring," as defined by measures such as the Embedded Figures Test. We are confident that this is an important variable, and available tests are adequate for an appropriate study. However, we have not yet completely defined how a working geologist would apply this ability to field studies, nor have we been able to create computer-based activities that mimic this process. We intend to create an interactive, computer-based module that involves disembedding figure from ground in realistic geological contexts, and replicating the current study in the near future.

Although we did not examine the variable of visual penetrative ability (VPA) discussed by Kali and Orion (1996), we did observe student behaviors that suggested the operation of such a factor. This was especially true in problems involving block diagrams. When attempting to interpret a block penetrated by an inclined plane, many students seemed unable to see the projection of the plane through the block. When asked to complete a drawing of the intersection of a plane with the block faces, students often continued the line from the known face across the unknown one as though it were a linear rather than a planar element. The line seemed to be perceived as something found only on the outside of the block, that wrapped around the block in a continuous fashion. We also observed many solutions where the line was drawn at an angle someplace between this interpretation and the correct one, as though students had an insight but were drawn perceptually to the incorrect solution. We also observed that this problem generated spirited discussions within groups where the correct and incorrect interpretations were held by members.

This study also has important implications to the issue of factors that influence the success of women in science. Gender differences in both spatial ability and achievement have been found by almost all those who have studied the topic. As suggested in our review of the literature, the question of the origin of these differences has not been answered. In this study, a relatively brief intervention succeeded in eliminating gender differences in spatial ability and closing the performance gap between males and females. This replicates a recent finding, in a study of success in engineering, that "females improved more than males in spatial ability" (Hsi, et al., 1997). Both results speak very strongly in favor of the position that observed gender differences are the result of differences in experience, and not of innate mental abilities, and that they can be eliminated by relatively minor treatments.

Although this intervention was brief, and did not allow an extended qualitative examination of student behaviors, all of the members of the research team spent time in the experimental classroom watching students work and talking to them about what they were doing. One set of observations, to which all observers agree, deserves discussion in this context. It appeared that the dynamics of group interactions depended heavily on the gender mix of the groups. This was especially evident when all-male and all-female groups were compared.

In all-male groups, the interactions were extremely limited. Since only one person could control the computer terminal, that tended to be the individual who already knew the most about the topic and who directed the activities of the group. In fact, in all-male groups, those who were not running the computer were generally uninvolved, sitting quietly and inattentively until an answer was reached that they could record on their work-sheet. There was virtually no discussion among members of the group, except in cases where the dominant male explained the results and answer to others.

All-female groups tended to work in a much different fashion. The person managing the computer was more often directed by the group about what action to take. The origin of this dynamic is not clear. Perhaps it was because no single clear leader in terms of computer skills emerged in female groups, or perhaps it was because females prefer to work in a more collaborative manner. Whatever the reason, female working groups tended to negotiate the action to be taken, and then to discuss the results among themselves before moving on to another action. This applied also to decisions about what information and conclusions to record on their work-sheets. There was a great deal more discussion and negotiation of meaning in groups composed entirely of females.

Much of the research comparing technology-based instruction to other methods has proven to be inconclusive. In general, technology is expensive and difficult to use, and not clearly superior to more traditional methods of instruction. It is our opinion that the superiority of computer-based education only becomes evident in cases where it is not possible to deliver the instruction by any other means.

A case in point is the topographic mapping module in this study. The geology department at this university has been using the "volcano in a box" laboratory, which originated many years ago with the Earth Science Curriculum Project, for some years in its introductory laboratories. However, creation of other landforms for students to explore in the same way has proven difficult. We are able to render virtually any topographic feature in the world into a three-dimensional, manipulable image. In addition, we have been able to create many new ways for students to manipulate these images that are not possible with the physical model.

The same could be said for the geologic blocks module. A teaching laboratory typically has only one or two three-dimensional block diagrams for students to work with. We have been able to produce dozens, with an exceptionally wide variety of features. And we can allow students to do things, like making the blocks transparent, that are impossible to do with physical models.

We also present these modules as a proof-of-concept for the use of computer-based instructional materials in a constructivist context. We allow students to begin their work with a playful, exploratory investigation of a variety of images. They work in groups, interacting with the computer and using worksheets to record their emerging interpretations of what they are seeing. We ask them to create pictures in their mind long before we offer formalisms such as the definition of contour intervals or the names of particular kinds of folds or faults.

One of the characteristics of science curricula since the reform movement of the 1960's has been their attempt to accurately portray the nature of science. This was commonly expressed as a concern for the structure of the discipline (Bruner, 1960). Initially, this took form as something approximating what is usually described as the "scientific method," and curricula taught students to observe, infer and test hypotheses. More recently, science educators have recognized significant differences among scientists working under different paradigms, and come to see that there may be many structures of this discipline we call science.

We have been trying to emphasize what we believe is a structure of the discipline of geology that is especially important, and perhaps more so in this case than in other sciences. Geologists use time and space to construct theories about the earth.
While the more traditional processes of science remain important, they are to some extent subordinated to the temporal-spatial reasoning that we think is characteristic of geology.

We believe that instruction should be anchored in authentic contexts and faithful to the structure of the geological sciences. Unfortunately, introductory courses at the college and university level are often disconnected collections of topics with no apparent coherence, and the tasks given to students in the laboratory bear little resemblance to the work of practicing scientists. We have tried to create a single unifying structure in which we situate instruction. Painted Canyon, a computer-generated terrain, is the context within which students learn geology in the laboratory. We try to represent the thought process of the geologist through a series of tasks for students that are as similar to those being undertaken by practicing geologists as we can possibly make them.

This study challenges conventional methods of teaching science. Rather than working from dull and uninteresting workbooks, students need to be engaged actively in realistic settings that are like those experienced by geologists themselves. Rather than dealing entirely in verbal forms of learning, they should engage all of the mental faculties, including but not limited to spatial visualization.

Finally, engaging in situated activities helps students to develop a set of intellectual skills that are demonstrably important to the learning of science and to the practice of geology. And it gives them some sense of what it is like to be a geologist. That, it seems to us, is among the most important goals of any course in laboratory science.

Acknowledgements: This research has been supported by funds provided through NSF grant EAR-9907733.

For Information Contact:

- Michael Piburn – mike.piburn@asu.edu
- Stephen Reynolds – sreynolds@asu.edu

Most materials are available online at http://reynolds.asu.edu

REFERENCES

Albanese, M. & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcome and implementation issues. Academic Medicine, 68(1), 52-80.

Ault, C.R., Jr. (1998). Criteria of excellence for geological inquiry: the necessity Of ambiguity. Journal of Research in Science Teaching, 35, p. 189-212.

Authorware 5 [Computer Software]. (1998). San Francisco, California: Macromedia, Inc.

Brown, J., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(1), 32-41.

Bruner, J. (1960). The Process of Education. Cambridge, MA: Harvard University Press. Chapter 2. The Importance of Structure.

Bryce4 [Computer Software]. (1999). Carpinteria, California: MetaCreations Corporation.

Burnett, S. & Lane, D. (1980). Effects of academic instruction on spatial visualization. Intelligence, 4(3), 233-242.

Carter, C., LaRussa, M. & Bodner, G. (1987). A study of two measures of spatial ability as predictors of success in different levels of general chemistry. Journal of Research in Science Teaching, 24(7), 645-657.

The Cognition and Technology Group at Vanderbilt (1990). Anchored instruction and its relation to situated cognition. Educational Researcher, 19(6), 2-10.

Cohen, H. (1983). A comparison of the effect of two types of student behavior with manipulatives on the development of projective spatial structures. Journal of Research in Science Teaching, 20(9), 875-883.

Dabbs, J. M., Chang, E., Strong, R. & Milun, R. (1998). Spatial ability, navigation strategy, and geographic knowledge among men and women. Evolution and Human Behavior, 19, 89-98.

Devon, R., Engel, R. and Turner, G. (1998). The effects of spatial visualization skill training on gender and retention in engineering. Journal of Women and Minorities in Science and Engineering, 4, 371-380.

Drummond, C. (1999). Facilitating scientific thoughtfulness for non-science majors. Journal of Geoscience Education, 47, p. 45-55.

Ekstrom, R., French, J., Harman, H. & Dermen, D. (1976). Manual for Kit of Factor Referenced Cognitive Tests. Princeton, NJ: Educational Testing Service.

Eley, M. (1993). The differential susceptibility of topographic map interpretation to influence from training. Applied Cognitive Psychology, 7(23-42).
influence from training.

Ericsson, K. & Smith, J. (1991). Prospects and limits of the empirical study of expertise: An introduction. In Ericsson and Smith (eds.), Toward a General Theory of Expertise: Prospects and Limits. Cambridge: Cambridge University Press.

Frodeman, R.L. (1996). Envisioning the outcrop. Journal of Geoscience Education, 44, p. 417-427.

Frodeman, R.L. (1995). Geological reasoning: geology as an interpretive and Historical science. Geological Society of America Bulletin, 107, p. 960-968.

FrontPage 2000 [Computer Software]. (2000). Microsoft Corporation.

Gardner, H. (1985). Frames of Mind: The Theory of Multiple Intelligences. Basic Books.

Hsi, S., Linn, M. & Bell, J. (April, 1997). The role of spatial reasoning in Engineering and the design of spatial instruction. Journal of Engineering Education, 151-158.

Kail, R., Carter, P., & Pellegrino, J. (1979). The locus of sex differences in spatial ability. Perception and Psychoanalysis, 26, 182-186.

Kali, Y. & Orion, N. (1996). Spatial abilities of high-school students and the perception of geologic structures. Journal of Research in Science Teaching, 33(4), 369-391.

Kali, Y., & Orion, N. (1997). Software for assisting high-school students in the spatial perception of geological structures. Journal of Geoscience Education, 45, p. 10-21.

Kimura, D. (1992). Sex differences in the brain. Scientific American, 267, 118-125.

Linn, M. and Petersen, A. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child Development, 56, 1479-1498.

Lord, T. (1985). Enhancing the visuo-spatial aptitude of students. Journal of Research in Science Teaching, 22(5), 395-405.

Lord, T. (1987). A look at spatial abilities in undergraduate women science majors. Journal of Research in Science Teaching, 24(8), 757-767.

Maccoby, E. (1966). Sex differences in intellectual functioning. In Maccoby, E. (Ed.), The Development of Sex Differences. Stanford, CA: Stanford University Press. Pp. 25-55.

Maccoby, E. and Jacklin, C. (1974). The Psychology of Sex Differences. Stanford, CA: Stanford University Press.

Maguire, E., Gadian, D., Johnsrude, I., Good, C., Ashbruner, J., Frackowiak, R. & Frith, C. (2000) Navigation-related structural change in the hippocampi of taxi drivers. PNAS, 97(8), 4398-4403.

McArthur, J. and Wellner, K. (1996). Reexamining spatial ability within a Piagetian framework. Journal of Research in Science Teaching, 33(10), 1065-1082.

McClurg, P. (19). Investigating the development of spatial cognition in problem-solving microworlds. Journal of Computing in Childhood Education, 3(2), 111-126.

Muehlberger, W.L., & Boyer, R.E. (1961). Space relations test as a measure of visualization ability. Journal of Geological Education, 9, p. 62-69.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt Rinehart and Winston.

Paivio, A. (1990). Mental representations: A dual coding approach. New York: Oxford University Press.

Pallrand, G. & Seeber, F. (1984). Spatial ability and achievement in introductory physics. Journal of Research in Science Teaching, 21(5), 507-516.

Piburn, M. (1980). Spatial reasoning as a correlate of formal thought and science achievement for New Zealand students. Journal of Research in Science Teaching, 17(5), 443-448.

Piburn, M. (1992). Meta-analytic and multivariate procedures for the study of attitude and achievement in science. Dortmund, Germany: International Council of

Associations for Science Education (ICASE).
Rieber, L.P. (1995). A historical review of visualization in human cognition. Educational Technology Research and Development, 43(1), 45-56.
Roe, A. (1961). The psychology of the scientist. Science, 134, 456-59.
Rudwick, M.J.S. (1976). The emergence of a visual language for geological science 1760-1840. History of Science, 14, p. 149-195.
Schofield, J. & Kirby, J. (1994). Position location on topographical maps: Effects of task factors, training, and strategies. Cognition and Instruction, 12(1), 35-60.
Shepard, R. (February 1978). The mental image. American Psychologist, 125-137.
Shepard, R. (1988). The Imagination of the scientist. In K. Egan & D. Nadaner (Eds.), Imagination and education (pp. 153-185). New York: Teachers' College Press.
Shepler, R. & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171, 701-703.
Smith, D. & Hoersch, A. (1995). Problem-based learning in the undergraduate geology classroom. Journal of Geological Education, 43, 385-390.
The VR Worx 2.0 [Computer Software]. (2000). Pittsburgh, Pennsylvania: VR Toolbox, Inc.
Tuckey, H. & M. Selvaratnam (1993). Studies involving three-dimensional visualization skills in chemistry: A review. Studies in Science Education, 21, PP. 99-121.
Vasta, R., Knott, J. and Gaze, C. (1996). Can spatial training erase the gender differences on the water-level task? Psychology of Women Quarterly, 20(4), 549-567.
Voyer, D., Voyer, S. and Bryden, M. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. Psychological Bulletin, 117(2), k250-270.
Witkin, H., Moore, C., Goodenough, D. and Cox, P. (1977). Field dependent and field independent cognitive style and their educational implications. Review of Educational Research, 47, 1-64.
Zavotka, S. (1987). Three-dimensional computer animated graphics: A tool for spatial instruction. Educational Communications and Technology Journal, 35(3), 133-44.

The Implications of Cognitive Studies for Teaching Physics

Edward F. Redish

Department of Physics, University of Maryland, College Par,k MD 20742-4111

This article first appeared in the American Journal of Physics,62(6), 796-803 (1994) ; ©1994, American Association of Physics Teachers.

I. Introduction

Many of us who have taught introductory physics for many years recall with dismay a number of salient experiences: a reasonably successful student who can produce a graph but can't say what it means; a top student who can solve all the problems but not give an overview or simple derivation; many students of varying abilities who memorize without understanding despite our most carefully crafted and elegant lectures.

As physics teachers who care about physics, we have a tendency to concentrate on the physics content we're teaching. We often are most concerned for those students who are like we were -- that small fraction of our students who find physics interesting and exciting and who will be the next generation of professional physicists. But the changes in our society and in the role of technology for the general public mean that we must change the way we are teaching. It no longer suffices to reproduce ourselves. Society has a great need not only for a few technically trained people, but for a large group of individuals who understand science.

During the past decade, data has built up that demonstrates that as physics teachers we fail to make an impact on the way a majority of our students think about the world.[1,2,3,4,5] We have readjusted our testing so that the students can succeed and we have then either fooled ourselves that we are teaching them successfully or lowered our standards by eliminating understanding from our definition of successful learning. Alan van Heuvelen [6] has remarked that in his study of a typical introductory lecture class, 20% of the students entered the first semester of an introductory calculus-based physics class as Newtonian thinkers. The impact of the course was to increase that number to 25%. If we want to reach a substantial fraction of our students, we must pay much more attention to how students learn and how they respond to our teaching. We must treat the teaching of physics as a scientific problem.

A few physicists have begun to perform detailed experiments to determine what our students are thinking and what works in the teaching of physics. Some of their articles are of tremendous importance and I believe should be read by every physics teacher (see refs. 4-5 and references therein). But even among these few articles, only a small fraction of the authors attempt to place their results in a general theoretical framework -- to give us a way of thinking about and organizing the results.[7] Those of us in physics know well that advancement in science is a continual dance between the partners of theory and experiment, first one leading, then the other. It is not sufficient to collect data into a "wizard's book" of everything that happens. That's not science. Neither is it science to spout high-blown theories untainted by "reality checks". Science must build a coherent and clear picture of what is happening at the same time as it continually confirms and calibrates that picture against the real world.

The time has come for us to begin the development of a framework for understanding and talking about student learning. Some of the results of the past few decades in cognitive studies8 begin to provide such a framework.Cognitive studies focuses on how people understand and learn. It is still an amorphous field, and it is not yet really a single discipline. It overlaps many areas from anthropology to neurophysiology. It may not yet be "a science" as we in physics use the term, but developments in the past few decades have changed drastically what we know about how the mind works.

The issue of how to teach physics is a difficult one: the attempt of a naive student to build a good understanding of physics involves many intricate processes over a long period of time. These processes tend to be much more complex than those most cognitive scholars have addressed.[9] Nonetheless, some of the basic ideas of cognitive studies appear to be both firmly grounded and useful to the teacher of physics.

In this essay I briefly review some of the lessons I have learned from cognitive studies. This is not a review article, but a narrow selection from a small slice of a large field. For those interested in a more substantial introduction to cognitive studies I recommend Howard Gardner's historical overview [10], the collection of articles assembled by Gentner and Stevens [11], and some of the articles in the reprint volume collected by Collins and Smith.[12] These will give an entry point into the modern cognitive literature. The book by Inhelder and Piaget [13] has lots of discussion of experiments on how adolescents learn physics. Some articles by leading educational specialists also can help link to the existing literature.[14] Just for fun, I have to add Donald Norman's delightful book on how people interact with objects in their world.[15] For an introduction to the physics education research literature, Arnold Arons's book [1] and a few review articles [16] provide an appropriate entry point.

I have grouped what I have learned from the cognitivists into four broad principles with elaborative corollaries. One of the things students of cognitive processes have learned about thinking is that it is fuzzy. The sharp, crisp operations of formal logic or the digital computer are not appropriate models for the way most people think. Therefore, it's not correct to call these principles "theorems" or "laws of cognitive science". Nor is it correct to use them as such. They can't provide us with hard and fast rules for what to do. Using them incautiously without reference to experimental data can lead us to the wrong conclusions. But I have found that they help me to organize my thinking about my students and to refocus my attention. Instead of concentrating only on the organization of the physics content, I now also pay attention to what my students are doing when they interact with a physics course. This is not to suggest that an emphasis on content is somehow unimportant or should be neglected. What we are teaching is important, but it must be viewed in the context of what our students learn.

II. Building Patterns: The Construction Principle

The fundamental change that has led to the breakthroughs in cognitive studies in the past few decades has been a new willingness to model what is happening in the mind in terms of inferred structures. For the first half of this century10, studies of thinking were severely constrained by the "behaviorist" philosophy that one should formulate all one's theories only in terms of direct observables. This is like the S-matrix theory of elementary particles which insisted that we should only formulate our theories in terms of observable particles and their scattering amplitudes. Elementary particle physicists only made their breakthrough when they were willing to formulate their ideas in terms of quarks and gluons -- particles which could only be inferred and not be seen directly. Cognitive scholars started to make real progress when they began to be willing to formulate how people were thinking in terms of mental patterns or models that could not be directly observed or measured.

Principle 1: (Weak form) People tend to form mental patterns.

This is the fundamental hypothesis about how the mind works. On some levels, there is direct observation of this mental processing by patterning. For example, it has been demonstrated in detail that we process visual information on a variety of levels to form patterns beginning with the first layer of nerve cells attached to the retina, and the process continues through many stages deep into the brain.[17] I once attended a physics colloquium given by Jerome Lettvin on the subject of blind spots in the visual field. He passed out the standard blind-spot demonstration pages[18] that let us clearly find the blind spot in our eye. We then moved the end of a pencil into our blind spots and saw the end of the pencil disappear as if bitten off. Yet there was no "blank spot" in the visual field. The brain fills in the background -- here the simple white of a blank page. But it will fill in even a quite complex pattern. If the page had been covered with plaid, my brain would still have filled in my blind spot with the appropriate pattern. But note that the patterning was not sufficiently sophisticated to produce the "right" answer! My automatic filling led to my seeing paper in the blind spot, not the rest of the pencil.

The tendency of the human mind to form patterns is not just limited to the analysis of sensory data. This leads me to state the principle in a stronger (and more relevant) form.

Principle 1:(Strong form) People tend to organize their experiences and observations into patterns or mental models.

I use the term mental model for the collection of mental patterns people build to organize their experiences related to a particular topic. I use the term schema (pl. schemas or schemata) to describe the basic elements of these mental models. I think of a schema as a "chunk" or "object" (in the sense of object-oriented programming). It is a collection of closely linked data together with some processes and rules for use. Be careful of the use of the word "model". It tends to convey something clockwork -- a mechanism that has links and rules and operates in a well-defined way. These are not the characteristics of many mental models.

The characteristics of mental models and schemas are still vigorously debated.[19] Despite attempts to build a general representational system for mental models, none has yet been widely accepted. However, some results are clear.[20]

Properties of mental models

Mental models have the following properties:

1. They consist of propositions, images, rules of procedure, and statements as to when and how they are to be used.
2. They may contain contradictory elements.
3. They may be incomplete.
4. People may not know how to "run" the procedures present in their mental models.
5. Elements of a mental model don't have firm boundaries. Similar elements may get confused.
6. Mental models tend to minimize expenditure of mental energy. People will often do extra physical activities -- sometimes very time-consuming and difficult -- in order to avoid a little bit of serious thinking.

This inferred structuring of mental models is distinctly different from what we usually assume when teaching physics. We usually assume that our students either know something or they don't. The view of mental models we learn from cognitive scholars suggests otherwise. It suggests that students may hold contradictory elements in their minds without being aware that they contradict.

I had an interesting experience that illustrated for me vividly the surprising fact that one's mental model may contain contradictory elements. Ron Thornton visited the University of Maryland a few years ago to give a seminar on his now famous work on using the Sonic Ranger to teach the concept of velocity.[3] The Ranger detects the position of an object using sonar and can display the position or the velocity of the detected object on a computer screen in live time. Thornton set up the Ranger to display velocity and had the computer show a pre-set pattern (a square wave). He then called me up to the front of the room to serve as a guinea pig and try to walk so my velocity matched the pre-set pattern.

I had no hesitation in doing this. I had been teaching physics for nearly twenty years and felt perfectly comfortable with the concept of velocity. I did my first trial without thinking; I walked backward until my velocity reached the height of the pre-set square wave. Then I stopped and my velocity dropped to zero immediately! I asked for another chance, and this time, putting my brain in "velocity mode", I was able to reproduce the curve without difficulty.

What this experience said to me was that, for normal walking, I still maintained a naive (but appropriate!) position-dominated proposition in my mental model of motion. I also had a correct proposition for the concept of velocity, but I had to consciously apply a rule telling me to use it. I've also had personal experiences illustrating characteristic 6. I once spent an hour searching through my hard drive and piles of floppy disks to find a short paragraph (three sentences!) that I needed again. When I found it, I realized it would have taken me only five minutes to have rewritten it from scratch. A nice example of this characteristic is Donald Norman's study of how people use complex calculators (ref. 19).

An important aspect of Principle 1 is that "people ... organize their experiences into ... mental models" with the emphasis on the fact that people must build their own mental models. This is the cornerstone of the educational philosophy known as constructivism. For this reason, I refer to Principle 1 as "The Construction Principle."

An extreme[21] statement of constructivism is: You can't teach anybody anything. All you can do as a teacher is to make it easier for your students to learn. Of course, facilitation can be critical to the learning process. Constructivism shouldn't be seen as disparaging teaching, but as demanding that we get feedback and evaluations from our students to see what works and what doesn't. It asks us to focus less on what we are teaching, and more on what our students are learning.

Implications

A number of interesting corollaries, elaborations, and implications that are relevant for understanding physics teaching come from Principle 1. The first is the realization that what we want our students to get is not simply "the content" but to build their understanding of that content into an accurate and effective mental model.

Corollary 1.1:The goal of physics teaching is to have students build the proper mental models for doing physics.

This helps us identify the point of teaching physics. I really want my students to do three things:

1. develop the ability to reason qualitatively about physical processes;
2. structure that content into coherent and appropriately organized --and appropriately accessible -- mental models;
3. learn how to apply that model to "do" physics in an expert and creative way.

These goals suggest that we should broaden our evaluation procedures. We traditionally test only the content and part of the student's skill in doing physics, usually (at least at the introductory level), in pre-set limited contexts.

Some of the characteristics of mental models clarify what is happening when students make mistakes. Often in listening to my students explain what they think I used to become confused and sometimes irritated. How can they say x when they know the contradictory principle y? How come they can't get started on a problem when they certainly know the relevant principle? They just told it to me two minutes ago! How come they brought up that particular principle now? It doesn't apply here! The well-documented characteristics of mental models [22] listed above helps me understand that these sorts of errors are natural and to be expected.Corollary 1.2: It is not sufficient for students to "know" the relevant correct statements of physics. They also have to be able to gain access to them at the appropriate times; and they have to have methods of cross-checking and evaluating to be certain that the result they have called up is truly relevant.We must also test the underlying mental models that the students are developing. Traditional testing fails to do this, because many schemas can produce the correct solution to a problem. Even if the student goes through the same steps as we do, there's no guarantee that their schema for choosing the steps is the same as ours.[23] I once asked a student (who had done a problem correctly) to explain his solution. He replied: "Well, we've used all of the other formulas at the end of the chapter except this one, and the unknown starts with the same letter as is in that formula, so that must be the one to use."

Part of the way we have fooled ourselves with our current testing is that we are interested "in what students know". If they don't access to the right "information" in an exam, we give them clues and hints in the wording to trigger access. But since an essential component of a mental model are the processes for access to information, we are not testing the complete mental model. The student "has" the information, but it is inert and cannot be used or recalled except in very narrow almost pre-programmed situations.

To find out what our students really know we have to give them the opportunity to explain what they are thinking in words. We must also only give exam credit for reasoning, and not give partial credit when a student tries to hit the target with a blast of shotgun pellets and accidentally has a correct and relevant equation among a mass of irrelevancies. To know whether our students access the information in appropriate circumstances we have to give them more realistic problems -- ones that relate directly to their real world experience and do not provide too many "physics clues" that specify an access path for them.

My next corollary relates to the problem that students often seem to listen to us but not to hear what we think we are trying to say.

Corollary 1.3: The student is not a tabula rasa (blank slate). Each one comes to us having had experiences with the physical world and having organized these experiences into mental models.

As physics teachers, we must realize that students come to us with naive mental models for how the physical world works. These are often referred to in the literature on physics education as preconceptions (or misconceptions).[24] Even experienced teachers can be surprised by this. A few years ago, after reading the ground-breaking articles of Halloun and Hestenes [2] in this journal on students' preconceptions in mechanics, I excitedly related a brief description of the results to one of my colleagues -- someone whom I know to be a concerned teacher and a superb lecturer. He was skeptical at the idea that students had trouble learning Newton's first law because they had pre-existing mental models that were friction dominated. He insisted: "Just don't tell them about friction. They won't know about it if you don't tell them." This natural expectation of this experienced teacher is now strongly contradicted by an impressive body of data.

The presence of "false" preconceptions really isn't so surprising if we think about our students' previous experience. Why should we be surprised that students think that any moving object will eventually come to a stop? In their direct personal experience that is always the case. It's even the case in the demonstrations we show in class to demonstrate the opposite! When I slide a dry-ice levitated puck across the lecture table, I catch it and stop it at the end of the table. If I didn't, it would slide off the table, bounce, roll a short distance, and stop. Every student knows that. Yet I ask them to focus on a small piece of the demonstration -- the stretch of about four or five seconds when the puck is sliding along the table smoothly -- and extend that observation in their minds to infinity. The student and the teacher are focusing on different aspects of the same physical phenomena.[25]

Many teachers show surprise in response to the excellent educational physics studies that have graced the pages of this journal demonstrating that students regularly generalize their naive schemas incorrectly. Why should it be surprising that students think cutting off part of a lens will result in their seeing only part of an image?[26] Try it with a magnifying glass! (Yes, I know that's not a real image.) Where do students get the idea that electricity is something that flows out of the wall and is used up in the object?[27] Why don't they think circuits? Although we don't always think about it, most of our students have had extensive experience with electricity by the time they arrive in our classes. When I said the current had to come in one wire and go out the other, one of my students complained: "If all the electricity goes back into the wall, what are we paying for?"

Corollary 1.4: Mental models must be built. People learn better by doing than by watching something being done.

This is sometimes expressed in the phrase: active learning works better than passive learning.[28] In most cases, this means that reading textbooks and listening to lectures is a poor way of learning.[29] This shouldn't be taken as universally true! As physics teachers, most of us have had the experience of having a few "good" students in our lectures -- students for whom listening to a lecture is an active process -- a dialog between themselves and the teacher. Indeed, many of us have been that good student and remember lectures (at least some of them) as significant parts of our learning experience.[30] A similar statement can be made about texts. I remember with pleasure working through texts and lecture notes, reorganizing the material, filling in steps, and posing questions for myself to answer. Yet few of my students seem to know how to do this or even to know that this is what I expect them to do. This leads us to think about a fifth corollary.

Corollary 1.5: Many of our students do not have appropriate mental models for what it means to learn physics.

This is a "meta" issue. People build mental models not only for content, but also for how to learn and what actions are appropriate under what circumstances. Most of our students don't know what you and I mean by "doing" science or what we expect them to do. Unfortunately, the most common mental model for learning science in my classes seems to be:

* Write down every equation or law the teacher puts on the board that is also in the book. Memorize these, together with the list of formulas at the end of each chapter.
* Do enough homework and end-of-the-chapter problems to recognize which formula is to be applied to which problem.
* Pass the exam by selecting the correct formulas for the problems on the exam.
* Erase all information from your brain after the exam to make room for the next set of material.

I used to be flabbergasted to discover that when I stopped a lecture and said: "OK, here's a really important general principle for how to do things. It's not in the book but it comes from my years of experience as a physicist," my students would not write it down or consider it important, even if I wrote it on the board! (Well, after all, it wasn't going to be on the exam.)

I call the bulleted list above "the dead leaves model". It's as if physics were a collection of equations on fallen leaves. One might hold s= ? gt**2, another F = ma, and a third F = -kx. These are each considered as of equivalent weight, importance, and structure. The only thing one needs to do when solving a problem is to flip through one's collection of leaves until one finds the appropriate equation. I would much prefer to have my students see physics as a living tree!

I like the term mental ecology to describe the mental model that tells students what mental model to apply in what set of circumstances. It is a more important goal to reshape our students' mental ecologies so that they expand their idea of learning to make it more constructive, take it out of the classroom into their everyday lives, and understand what science is and how to apply it, than it is to teach them to parrot back equations or solutions to turn-the-crank problems. One final observation on the first principle is the following:

Constructing our own lectures and teaching materials can prove very useful in producing learning -- in the teacher!

Haven't we all remarked: I only really understood E&M (or classical mechanics, or thermodynamics, or whatever) when I finally taught it? This is really dangerous! For those of us who love learning, the experience of lecturing and teaching is such a powerful learning experience that we don't want to give it up, even when it proves less effective for our students than other methods.

III. Building on a Mental Model: The Assimilation Principle

The second and third principles have to do with the dynamics of modifying and extending one's mental models.

Principle 2: It is reasonably easy to learn something that matches or extends an existing mental model.

This principle states that mental models are not only the way that we organize our interactions with the world, but they also control how we incorporate new information and experiences.31 (The question of how they are first established in young children is interesting -- and controversial -- but it doesn't really concern us here.) I use the term "assimilate" to emphasize adding something smoothly to an existing set.32I pose three restatements and elaborations of this principle as corollaries to show what it means for teaching.

Corollary 2.1: It's hard to learn something we don't almost already know.

All students have things they know (some of which may be wrong!), things they are a bit familiar with, and things they have no knowledge about at all. In the last area my daughter would say they're "clueless".

I like to look at this as an archery target. What they know is the bull's-eye -- a compact black area; what they know a little about is a gray area surrounding the black; and the clueless region is a white "rest of the world". To teach them something, I do best to hit in the gray. A class full of students is a challenge because all of their gray areas are not the same. I want to hit as many of the grays as possible with each paint-tipped shaft of information to turn gray areas black. In communication studies, an important implication of this corollary is called the "given-new principle".33 It states that new information should always be presented in a context that is familiar to the reader and that the context should be established first. The analogous statement is very important in physics teaching, especially at the introductory level. As physicists with years of training and experience we have a great deal of "context" that our students don't possess. Often we are as fish in water, unaware of this context and that it is missing in our students.

There are a number of specifics that we can cite that are given-new problems. We often use terms that students are not familiar with -- or use in a different sense than we do. As a part of their study in the way speakers of English build their meaning of the term "force", Lakoff and Johnson [34] classified the characteristics of common metaphors using the term. Among their list of 11 characteristics, 8 involved the will or intent of an individual! But most of us are so familiar with the technical meaning of "force" that we surprised to learn that a significant fraction of our introductory students do not believe that a table exerts a force on a book it is supporting.[2] Why doesn't the book fall through? The table is just "in the way".

The problem caused by the interpretation of common speech words for technical ones is not a simple one. I know that the terms "heat", and "temperature" are not really distinguished in common speech and are used interchangeably for the technical terms "temperature" (average energy per degree of freedom), "internal energy", and "heat" (flow of internal energy from one object to another). In one class, I stated this problem up front and warned my students that I would use the terms technically in the lecture. Part way through I stopped, realizing that I had used the word "heat" twice in a sentence -- once in the technical sense, once in the common speech sense.[35] It's like using the same symbol to stand for two different meanings in a single equation. You can occasionally get away with it,[36] but it isn't really a good idea!Putting new material in context is only part of the story. Our students also have to see the new material as having a plausible structure in terms of structures they are familiar with. We can state this as another useful corollary.

Corollary 2.2: Much of our learning is done by analogy.

This strongly counters the image of the student as a tabula rasa. This and the previous corollary make what students know at each stage critical for what we can teach them. Students always construct their knowledge, but what they construct depends on how what we give them interacts with what they already have.

One implication of these results is that we should focus on building structures that are useful for our students' future learning. I state this as a third corollary.

Corollary 2.3: "Touchstone" problems and examples are very important.

By a touchstone problem, I mean one that the student will come back to over and over again in later training. Touchstone problems become the analogs on which they will build the more sophisticated elements of their mental models.

It becomes extremely important for students to develop a collection of a few critical things that they really understand well.[37] These become the "queen bees" for new swarms of understanding to be built around. I believe this is why some problems have immense persistence in the community. Inclined plane problems really aren't very interesting, yet the occasional suggestions that they be done away with are always resisted vigorously. I think the resisters are expressing the (usually unarticulated) feeling that these are the critical touchstone problems for building students' understanding of vector analysis in the plane. Corollary 2.3 is one reason why we spend so much time studying the mass on a spring. It's not really of much interest in itself, but it serves as a touchstone problem for all kinds of harmonic oscillation from electrical circuits up to quantum field theory.

Looking at a curriculum from the point of view of the mental models we want students to develop, their pre-existing mental models, and touchstone problems can help us analyze what is critical in the curriculum, which proposed modifications could be severely detrimental, and which might be of great benefit.

Combining this with the discussion of access and linking above leads us to focus on the presence of a framework or structure within a course. It suggests that building a course around a linked series of touchstone problems could be of considerable assistance in helping students understand the importance and relevance of each element. Such a structure is sometimes referred to as a story line.

IV. Changing an Existing Mental Model: The Accommodation Principle

Unfortunately, if students are not blank slates, sometimes what is written is -- if not wrong -- inappropriate for future learning in physics. Then it can seem as if we have run into a brick wall. This brings us to the next principle. I call this the "accommodation principle" to emphasize that changes have to be made in an existing structure. (Again, the term goes back to Piaget.)

Principle 3: It is very difficult to change an established mental model substantially.

Traditionally, we've relied on an oversimplified view of Principle 1, the patterning principle, to say: "Just let students do enough problems and they'll get the idea eventually." Unfortunately, the principle doesn't apply in this form if they already have a mental model about the subject. It has been demonstrated over and over again that simply telling somebody something doesn't easily change their deep ideas. Rather, what happens is that a poorly linked element is added with a rule for using it only in physics problems or for tests in one particular class. This and the fact that a mental model can contain contradictory elements is the reason why "giving more problems" can be ineffective. Once students learn how to do problems of a particular type, many will learn nothing more from doing more of them: new problems are done automatically without thinking. This also means that testing by varying homework problems slightly may be inadequate to probe the student's mental models of physics. More challenging tests involving a variety of modalities (problem solving, writing, interpreting, organizing) are required.

A few years ago I learned a lovely anecdote illustrating the barriers one encounters in trying to change a well-established mental model. A college physics teacher asked a class of beginning students whether heavy objects fall faster than light ones or whether they fall at the same rate. One student waved her hand saying "I know, I know". When called on to explain she said: "Heavy objects fall faster than light ones. We know this because Galileo dropped a penny and a feather from the top of the leaning tower of Pisa and the penny hit first." This is a touchstone example for me. It shows clearly that the student had been told -- and had listened to -- both the Galileo and the Newton stories. But she had transformed them both to agree with her existing mental model.[38]

Principle 3 can cause problems, both in getting students to change their mental models, and in getting ourselves to change the preconceptions we have about how students think! Fortunately, "difficult" does not mean "impossible". We have mechanisms that permit us to measure our mental models against the world and change them when we are wrong.[39] It appears as if the mechanism critically involves prediction and observation. The prediction must be made by the individual and the observation must be a clear and compelling contradiction to the existing mental model. A simple contradiction isn't sufficient. Posner et al.40 suggest that changing an existing mental model requires that the change have the following characteristics (which I state as a corollary).

Corollary 3.1: In order to change an existing mental model the proposed replacement must have the following characteristics:

* It must be understandable.
* It must be plausible.
* There must be a strong conflict with predictions based on the existing model.
* The new model must be seen as useful.

The clearer the prediction and the stronger the conflict, the better the effect. A nice example of how this process works concerns physics teachers rather than their students. It explains why the response to the Halloun-Hestenes test has been so great. Many teachers of introductory physics have a mental model that says the teacher is successful in teaching the concepts if the students can average 75% on a traditional exam. These teachers look at the H-H test and predict: "My students will be able to do very well on those problems. They're easy." Then their predictions fail miserably. On some critical questions, students average 20% or less and the conflict with the teacher's existing mental model is very strong. Many of the teachers who have gone through this process appear to be converted to a new way of looking at their students and what they know.[41]

Attempts are being made to combine the assimilation and the accommodation principles to yield new, more effective methods of teaching. John Clement[42] has proposed finding a series of bridges or interpolating steps that would help a student transform his or her mental model to match the accepted scientific one.

V. The Individuality Principle

One might be tempted to say: Fine. Let's figure out what the students know and provide them with a learning environment -- lectures, demonstrations, labs, and problems -- that takes them from where they are to where we want them to be. Since we all know that a few students get there from here using our current procedures, how come it doesn't work for all of them? We do in fact now know that the right environment can produce substantially better physics learning in most of the students taking introductory university physics.[43] But my final principle is a word of warning that suggests we should not be looking for a "magic bullet".

Principle 4: Since each individual constructs his or her own mental ecology, different students have different mental models for physical phenomena and different mental models for learning.

I like to call this the individuality or "line width" principle. This reminds us that many variables in human behavior have a large natural line width. The large standard deviation obtained in many educational experiments is not experimental error, it is part of the measured result! As scientists, we should be accustomed to such data. We just aren't used to its being so broad and having so many variables. An "average" approach will miss everyone because no student is average in all ways.[44]

Implications

One implication of this is that different students can have different reasons for giving the same answer. If we formulate our questions too narrowly we may misinterpret the feedback we are getting. This observation has influenced the style of educational physics research in a way that at first seems strange to a physical scientist.

When we try to take our "first look" at what students are doing, it is very important to consider them one at a time and to interview them in depth, giving them substantial opportunity for "thinking aloud" and not giving them any guidance at all.

This approach is characteristic of many important educational studies. Instead of hoping to "average out" the variation by doing large statistical experiments, one focuses on it and tries to learn the range of possible approaches that students are taking. Of course, at a later stage, one wants to be able to interrogate large numbers of students in order to obtain the frequency with which various modes of thinking are occurring in the population at large. But many valuable studies in educational physics (and in cognitive studies in general) are done with a sample that seems very small to a physicist.

An excellent example is the work of McDermott and Goldberg (ref. 25). They start with extensive interviews of a fairly small number of students, then develop short answer exams based on those observations that can be applied to large groups, and finally test that those exams are giving the same results as the interviews. The various Hestenes tests were developed in a similar fashion.[45] In addition to the fact that students have different experiences and have drawn different conclusions from them, their methods of approach may differ significantly. I state this as a corollary.

Corollary 4.1: People have different styles of learning.

There is by now a vast literature on how people approach learning differently. Many variables have been identified on which distributions have been measured. These include authoritarian/independent, abstract/concrete, and algebraic/geometric to name a few.[46] The first means that some students want to be told, others to figure things out for themselves. The second means that some students like to work from the general to the specific, some the other way round. The third means that some students prefer to manipulate algebraic expressions while others prefer to see pictures. Many of us who have introduced the computer in physics teaching have noted that some students want to be guided step by step, others try everything. These are only a few of the variables.

Once we begin to observe and use these differences in our students, we have to be exceedingly careful about how we use them. A preference does not mean a total lack of capability. Students who prefer examples with concrete numbers to abstract mathematical expressions may be responding to a lack of familiarity with algebra rather than a lack of innate ability. Many of our students preferences come from years of being rewarded for some activities (such as being good memorizers) and chastised for others (such as asking questions the teacher couldn't answer). Expanding our students' horizons and teaching them how to think sometimes requires us to overcome years of negative training and what they themselves have come to believe are their own preferences and limitations!

An interesting observation that has been made by a number of observers,[47] is that physics as a discipline requires learners to employ a variety of modalities (methods of understanding) and to translate from one to the other -- words, tables of numbers, graphs, equations, diagrams, maps. Physics requires the ability to use algebra and geometry and to go from the specific to the general and back. This makes learning physics particularly difficult for many students. One of our goals should be to have our students understand this, be able to identify their own strengths and weaknesses, and while building on the former, strengthen the latter.

An important implication is the following:

Corollary 4.2: There is no unique answer to the question: What is the best way to teach a particular subject?

Different students will respond positively to different approaches. If we want to adopt the view that we want to teach all our students (or at least as many as possible), then we must use a mix of approaches and be prepared that some of them will not work for some students. An important set of studies that are just beginning to be done ask the question: What is the distribution function of learning characteristics that our students have in particular classes?

Another implication that is very difficult to keep in mind is:

Corollary 4.3: Our own personal experiences may be a very poor guide for telling us what to do for our students.

Physics teachers are an atypical group. We selected ourselves at an early stage in our careers because we liked physics for one reason or another. This already selects a fairly small subclass of learning styles from the overall panoply of possibilities. We are then trained for approximately a dozen years before we start teaching our own classes. This training stretches us even further from the style of approach of the "typical" student. Is it any wonder why we don't understand most of our beginning students and they don't understand us? I will never forget one day a few years ago when a student in my algebra-based introductory physics class came in to ask about some motion problems. I said: "All right, let's get down to absolute basics. Let's draw a graph." The student's face fell, and I realized suddenly that a graph was not going to help him at all. I also realized that it was going to be hard for me to think without a graph and to understand what was going through the student's mind. I never minded doing without derivatives -- motion after all is the study of experimental calculus and you have to explain the concept (maybe without using the word) even in a non-calculus based class; but I have found it difficult to empathize with students who come to physics and can't read a graph or reason proportionately.[5] It takes a special effort for me to figure out the right approach.

This is very natural given the earlier principles. Our own personal mental models for how to learn come from our own experiences. However, to reach more of our students than the ones who resemble ourselves, we will have to do our best to get beyond this. It makes the following principle essential.

Corollary 4.4: The information about the state of our students knowledge is contained within them. If we want to know what they know, we not only have to ask them, we have to listen to them!

VI. Conclusion

The typical university course is a complex structure. It involves physics content, a teacher, perhaps graders or teaching assistants, a classroom, a laboratory, and, for each class, a particular set of students. Above all, it involves expectations and contexts for both the teacher and the students. If we are to make serious progress in reaching a larger fraction of our students, we will have to shift our emphasis from the physics content we enjoy and love so well to the students themselves and their learning. We must ask not only what do we want them to learn, but what do they know when they come in and how do they interact with and respond to the learning environment and content we provide.

The principles we are learning from cognitive studies can provide a framework for how we think about the complex issues of teaching and learning. The four principles that I have presented can help us begin to construct such a framework.

Acknowledgments

This article arises from the valuable discussions I have had with many friends and colleagues, especially Arnold Arons, Seth Chaiklin, Dewey Dykstra, John Guthrie, David Hestenes, Priscilla Laws, Lillian McDermott, Ed Taylor, Ron Thornton, Jack Wilson, and Dean Zollman. I would like to thank John Clement, Priscilla Laws, and Len Jossem for commenting on the manuscript. The suggestions of two very conscientious AJP reviewers were helpful in guiding me towards revisions that improved the paper significantly. Finally, I would like to thank my wife, Dr. Janice C. Redish, for numerous discussions on the subject of cognitive studies. Her expertise has played a major role in the development of my views on these issues. I would also like to thank her for her detailed editorial comments on the presentation. She is not to be held responsible for my persistent grammatical idiosyncrasies.

Endnotes

[1] Arnold Arons, A Guide to Introductory Physics Teaching (Wiley, 1990).

[2] David E. Trowbridge and Lillian C. McDermott, "Investigation of student understanding of the concept of velocity in one dimension", Am. J. Phys., 48 (1980) 1020-1028; "Investigation of student understanding of the concept of acceleration in one dimension", Am. J. Phys. 49 (1981) 242-253.

[3] A. Halloun and D. Hestenes, "The initial knowledge state of college physics students", Am. J. Phys., 53, 1043-1055 (1985); "Modeling instruction in mechanics", Am. J. Phys. 55, 455-462 (1987).

[4] R. K. Thornton and D. R. Sokoloff, "Learning motion concepts using real-time microcomputer- based laboratory tools", Am. J. Phys. 58 (1990) 858-867.

[5] Lillian C. McDermott, "Millikan Lecture 1990: What we teach and what is learned -- Closing the gap", Am. J. Phys. 59, 301-315 (1991); "Guest Comment: How we teach and how students learn-- A mismatch?" Am. J. Phys. 61, 295-298 (1993); and refs. therein.

[6] Alan van Heuvelen, Bulletin of the American Physical Society, 36, 1359 (April, 1991).

[7] One that has been very useful for me is the article by David Hestenes, "Toward a modeling theory of physics instruction", Am. J. Phys. 55, 440-454 (1987).

[8] I use the term "cognitive studies" as opposed to the term "cognitive science", which is more prevalent in the literature, advisedly. The field is extremely broad and cuts across many disciplines. Very little in cognitive science satisfies the scientific criteria we are used to in physics of being precisely stated and well-tested experimentally, as well as useful. So as not to raise the expectations of my scientist readers as to the nature of the information pr*cised here, I use the weaker term (and use it as a singular noun like "physics").

[9] There are a few notable exceptions. Piaget has made extensive studies of how children build their concepts of the physical world, and a number of trained physicists such as Fred Reif, John Clement, and Jose Mestre, among others, could be counted as cognitive scholars.

[10] Howard Gardner, The Mind's New Science: A History of the Cognitive Revolution (Basic Books, 1987).

[11] Derdre Gen tner and Albert L. Stevens, Eds. Mental Models (Lawrence Erlbaum Associates, 1983).

[12] Allan Collins and Edward E. Smith, Readings in Cognitive Science (Morgan Kauffmann, 1988).

[13] BSrbel Inhelder and Jean Piaget, The Growth of Logical Thinking from Childhood to Adolescence (Basic Books, 1958).

[14] L. B. Resnick, "Mathematics and science learning: A new conception", Science 220, 447-478 (1983); S. Carey, "Cognitive science and science education", Am. Psych. 41, 1123-1130 (1986); R, Driver, "Students' conceptions of the learning of science", Int. J. of Sci. Ed. 11, 481-490 (1989).

[15] Donald Norman, The Design of Everyday Things (Basic Books, 1990). (This book was originally published under the title: The Psychology of Everyday Things.)

[16] Lillian McDermott, "Research on conceptual understanding in mechanics", Phys. Today 37(7), 24-32 (1984); Jose Mestre and Jerrold Touger, "Cognitive research -- what's in it for physics teachers?", The Physics Teacher 27, 447-456 (September, 1989); Robert Fuller, "Solving physics problems -- how do we do it?", Physics Today 35:9, 43-47 (1982) .

[17] David H. Hubel, "The Visual Cortex of the Brain", Scientific American 225:11, 54 (1963); "The Brain", Scientific American 241:3, 44(1979).

[18] The page is blank with two marks: a black dot and a cross about six inches apart. One closes one eye and focuses on the black dot with the other. The page is oriented so the cross is off to one side. By moving one's head back and forth, one finds a distance at which the cross appears to vanish.

[19] One must take particular care with these terms in the cognitive literature as they are used by different authors in different senses.

[20] This list of properties is based in part on Donald Norman, "Some observations on mental models", in Gentner and Stevens, op. cit., pp. 7-14 and refs. therein.

[21] Actually, a very wide variety of educators and cognitive sp ecialists consider themselves constructivists. The statement here is not really extreme in the current spectrum of views. Radical constructivists reject the idea that knowledge has any base outside of construction in the human mind, a view which leads to results that could be classified as solipsism.

[22] Andy diSessa has documented and attempted to classify some of the levels of principles that people create and their modes of access to them. Although his papers are rather technical, he uses many examples from physics and includes references to the large literature on the subject. See for example, "Phenomenology and the evolution of intuition", in Genter and Stevens (Eds.), op. cit., 15-33; or "Knowledge in Pieces", Cognition and Instruction 10, issues 2 and 3.

[23] The difficulty is that the mapping from underlying schema to problem solving steps is not one- to-one. As nice example of this is given in a recent article in this journal: J. Bowden, et al., "Displacement, velocity, and frames of reference: phenomenographic studies of students' understanding and some implications for teaching and assessment", Am. J. of Phys. 60, 262-269 (1992) .

[24] I prefer the term preconceptions or naive schemas, with "naive" used in the non-pejorative sense of "untutored". The typical student's naive schema about motion and mechanics, for example, is not wrong. It correctly correlates their observations at low velocities in friction- dominated systems. It just does not provide them with the power and range of the schema we are trying to substitute for it. See refs. 1-5 and the references referred to therein for details and documentation.

[25] This argument is made in a slightly different context in T. S. Kuhn, The Structure of Scientific Revolutions, 2nd ed. (Chicago University Press, 1970).

[26] F. C. Goldberg and L. C. McDermott, "An investigation of student understanding of the real image formed by a converging lens or concave mirror", Am. J. Phys. 55, 108-119 (1987); "An investigation of student understanding of the images formed by plane mirrors", The Physics Teacher 34, 472-480 (1986).

[27] R. Cohen, B. Eylon, and U. Ganiel, "Potential difference and current in simple electric circuits: A study of students' concepts", Am. J. Phys. 51, 407-412(1983).

[28] We have to b e careful what we mean by "active learning". This shouldn't necessarily be taken to mean that "solving more problems is the answer". See the discussion after principle 3.

[29] J. S. Brown, A. Collins, P. Duguid, "Situated Cognition and the Culture of Lea rning", Educational Research, 32-42 (Jan-Feb, 1989), and refs. therein.

[30] In many research groups, a seminar more resembles a panel discussion than a lecture. These are very active learning experiences, both for the speaker and the listener.

[31] J. D. Bransford and M. K. Johnson, "Contextutal prerequisites for understanding: Some investigations of comprehension and recall", J. of Verbal Learning and Verbal Behavior 11, 717-726 (1972); "Considerations of some problems of comprehension", in W. Chase (Ed.), Visual Information Processing (Academic Press, NY, 1973).

[32] The term "assimilation" and the matching term "accommodation" used in the next section were introduced by J. Piaget.

[33] H. Clark and S. Haviland, "Comprehension and the given-new contract", in Discourse Production and Comprehension, R. Freedle, ed. (Lawrence Erlbaum, 1975).

[34] G. Lakoff and M. Johnson, Metaphors We Live By (U. of Chicago Press, Chicago, 1980), p. 70.

[35] "If there is no heat flow permitted to the object we can still heat it up by doing work on it."

[36] The energy levels of hydrogen in a magnetic field could be written with an m in the denominator for the electron mass and one in the numerator for the z-component of the angular momentum. Most physicists can correctly interpret this abomination without difficulty.

[37] In addition to giving them centers on which to build future learning, kn owing a few things well gives the student a model of what it means to understand something in physics. This valuable point that has been frequently stressed by Arnold Arons. It is an essential element in the mental ecology of a scientist.

[38] Of course the students' mental model in this case is in fact correct. Lighter objects do fall more slowly than heavy ones if they fall in air, and few of us have much direct experience with objects falling in a vaccuum. But that observation does not yield a useful idealization. The observation that objects of very different mass fall in very nearly the same way does.

[39] Unfortunately, these mechanisms appear to atrophy if unused.

[40] G. J. Posner, K. A. Strike, P. W. Hewson, and W. A. Gertzog, "Accommodation of a scientific conception: toward a theory of conceptual change", Science Education 66:2, 211-227 (1982).

[41] A nice description o f this process is given in Eric Mazur's brief but insightful article on teaching, "Qualitative vs. quantitative thinking: Are we teaching the right thing?", Optics & Phonics News 38 (February, 1992).

[42] John Clement, "Overcoming Students' Misconception s in Physics: The Role of Anchoring Intuitions and Analogical Validity", Proc. of Second Int. Seminar: Misconceptions and Educ. Strategies in Sci. and Math. III, J. Novak (ed.) (Cornell U., Ithica, NY, 1987); "Not all preconceptions are misconceptions: Finding 'anchoring conceptions for grounding instruction on students' intuitions", Int. J. Sci. Ed. 11 (special issue), 554-565 (1989).

[43] Priscilla Laws, "Calculus-based physics without lectures", Phys. Today 44:12, 24- 31(December, 1991); Alan Van Heuvelen, "Overview, Case Study Physics", Am. J. Phys. 59, 898-907(1991); Richard Hake, "Socratic Pedagogy in the Introductory Physics Laboratory",The Physics Teacher, 33 (1992); L. C. McDermott and P. S. Shaffer, "Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding", Am. J. Phys. 60, 994-1003 (1992); erratum, ibid. 61, 81 (1993); P. Shaffer, and L. C. McDermott, "Research as a guide for curriculum development: An example from introductory electricity. Part II: Design of an instructional strategy", Am. J. Phys. 60, 1003-1013 (1992).

[44] This is analogous to the story of the three statisticians who went hunting deer with bows and arrows. They cam across a large stag and the first statistician shot -- and missed to the left. The second statistician shot and missed to the right. The third statistician jumped up and down shouting "We got him!"

[45] Halloun and Hestenes, Op. Cit.; David Hestenes, M. Wells, and G. Swackhammer, "Force Concept Inventory", The Physics Teacher, 30:3, 141-158(1992); D. Hestenes and M. Wells, "A Mechanics Baseline Test", The Physics Teacher, 30:3, 159-166(1992).

[46] David A. Kolb, Experiential learning: experience as a source of learning and development (Prentice Hall, 1984); Noel Entwistle, Styles of integrated learning and teaching: an integrated outline of educational psychology for students, teachers, and lecturers (Wiley, 1981); Howard Gardner, Frames of Mind (Basic Books, 1983).

[47] Priscilla Laws, David Hestenes, private communication.

[Image]

Maintained by University of Maryland PERG
Comments and questions may be directed to E. F. Redish

Last modified 5 February 2002

Chapter 7 There's more than content to a physics course: The hidden curriculum [1]

Education is what survives
when what has been learned
has been forgotten.
B. F. Skinner

(New Scientist , 21 May 1964)

Remembrance and reflection, how allied.
What thin partitions, sense from thought divide!

Alexander Pope

A consideration of the cognitive model of student thinking allows us to create lessons that can help them readjust their schemas rather than totally recreate them and that can lead to their substantially improving their understanding of basic concepts. As we begin to be more aware of the complexity and the strong context dependence of student thinking even when they are giving simple answers, we begin to identify the *executive components* of their reasoning as important — those cognitive functions that control their access to declarative and reasoning elements in their schema.

It is not only ideas about how the physical world works that a student brings into the physics classroom. We are often frustrated by the tendency many students have to seek "efficiency" — to achieve a satisfactory grade with the least possible effort — often with a severe undetected penalty on how much they learn. They have a sense of what is appropriate for them to do in order to succeed in our class that may or may not be correct. They may spend a large amount of time memorizing long lists of uninterpreted facts or performing algorithmic solutions to large numbers of problems without giving them any thought or trying to make sense of them. Although some students consider this efficient, it is only efficient in the short term. The knowledge thus gained is superficial, situation dependent, and quickly forgotten.

Each student, based on his or her own experiences, brings to the physics class a set of attitudes, beliefs, and assumptions about what sorts of things they will learn, what skills will be required, what they will be expected to do, and what kind of arguments and reasoning they are allowed to use in the various environments found in a physics class. In addition, their view of the nature of scientific information affects how they interpret what they hear. I use the phrase *expectations* to cover this rich set of understandings that are particular to a given class. Students' views of the nature of knowledge and how they learn are often referred to in the education literature as their *epistemologies*.[2]

These attitudes, expectations, and epistemologies affect what they listen to and what they ignore in the firehose of information provided during a typical course by professor, teaching assistant, laboratory, and text. It affects which activities students select in constructing their own knowledge base and in building their own understanding of the course material. The impact can be particularly strong when there is a large gap between what the students expect to do and what the instructor expects them to do.

Although we don't often articulate them, most physics instructors have expectation-related goals for their students. In our college and university physics courses for engineers, biologists, and other scientists, we try to get students to make connections, understand the limitations and conditions on the applicability of equations, build their physical intuition, bring their personal experience to bear on their problem solving, and see connections between classroom physics and the real world. Above all, we expect students to be *making sense* of what they are learning. I refer to this kind of learning goal — a goal not listed in the course's syllabus or the textbook's table of contents — as part of the course's *hidden curriculum.*

Studies of Learning Attitudes

There are a number of studies of student expectations in science in the pre-college classroom that show that student attitudes towards their classroom activities and their beliefs about the nature of science and knowledge affect their learning. Many studies (see, for example, [Carey 1989] and [Linn 1991]) have demonstrated that pre-college students often have misconceptions both about the nature of scientific knowledge and about what they should be doing in a science class. Other studies indicate some of

---

[1] This chapter is based in part on the paper by Redish, Saul, and Steinberg. [Redish 1998]

[2] This word is borrowed from philosophy where it develops all kinds of arcane and delicate meanings. (For example, see [von Glaserfeld].) I use it here in a very limited way – "How do we know what we know?"

the critical items that make up the relevant elements of a student's system of expectations and beliefs. For example, Songer and Linn studied students in middle schools and found that they could already categorize students as having beliefs about science that were either *dynamic* (science is understandable, interpretive, and integrated) or *static* (science knowledge is memorization-intensive, fixed, and not relevant to their everyday lives).[Songer 1991] Alan Schoenfeld studied the assumptions high schools students make about learning mathematics. He concludes, "Student's beliefs shape their behavior in ways that have extraordinarily powerful (and often negative) consequences." [Schoenfeld 1992]

Two important large-scale studies that concern the general cognitive expectations of adult learners are those of Perry and Belenky et al. [Perry] [Belenky] Perry tracked the attitudes of Harvard and Radcliffe students throughout their college career. Belenky et al. tracked the views of women in a variety of social and economic circumstances. Both studies found evolution in the expectations of their subjects, especially in their attitudes about knowledge.[3] Both studies frequently found their young adult subjects starting in a *binary* or *received knowledge* stage in which they expected everything to be true or false, good or evil, etc., and in which they expected to learn "the truth" from authorities. Both studies observed their subjects moving through a *relativist* or *subjective* stage (nothing is true or good, every view has equal value) to a *consciously constructivist* stage. In this last, most sophisticated stage, the subjects accepted that nothing can be perfectly known, and accepted their own personal role in deciding what views were most likely to be productive and useful for them.

Although these studies both focused on areas other than science,[4] Sagredo and I both recognize a binary stage, in which students just want to be told the "right" answers, and a constructivist stage in which students take charge of building their own understanding.[5] Consciously constructivist students carry out their own evaluation of an approach, equation, or result, and understand both the conditions of validity and the relation to fundamental physical principles. Students who want to become creative scientists will have to move from the binary to the constructivist stage at some point in their education.

		Everyday domain	Scientific Domain
Domain Goals			
Main goals			
	Central goal	Leading a good life	Optimal predication and explanation
	Subgoal	Adequate prediction and explanation	
	Requirements	Adequate generality, parsimony, precision, consistency	Maximum generality, parsimony, precision, consistency
Working goals			
	Understanding	Few inferences, various acceptable premises	Many inferences, well-specified premises
	Assessing validity	Moderate importance, various acceptable premises, plausible inference rules	Central importance, observation-based premises, well-specified inference rules
Domain Cognition			
Knowledge structure			
	Concept specification	Implicit and schema-based	Explicit and rule-based
	Knowledge organization	Locally coherent, associative organization	Globally coherent, logical organization
Methods			
	Problem solving	Short inferences based on rich compiled knowledge	Long inferences based on parsimonious knowledge
	Types of methods	Non-formal	Complementary formal and non-formal
Quality concerns			
	Quality control	Non-formal	Strict and explicit
	Efficiency	Naturally efficient for everyday tasks	Designed for efficiency in complex tasks

Table 1: Comparison between everyday and scientific knowledge domains. [Reif 1991]

3 This brief summary is an oversimplification of a complex and sophisticated set of stages proposed in each study.

4 Perry specifically excludes science as "the place where they *do* have answers."

5 In my experience true relativism is rare, but not unheard of, among physics students.

An excellent introduction to the cognitive issues involved is given by Reif and Larkin who compare the intellectual domains of spontaneous cognitive activities that occur naturally in everyday life with those required for learning science.[Reif 1991] They pinpoint differences between these domains and show how application of everyday cognitive expectations in a science class causes difficulties. The extensive differences they identify are summarized in table 1.

Although there is no space to go into each of these entries in detail, even these brief descriptions are enough for an experienced instructor to recognize that students often applies everyday-domain cognition when we want them to apply scientific-domain cognition.

Another excellent introduction to the cognitive literature on the difference between everyday and in-school cognitive expectations is the paper by Brown, Collins, and Duguid, discussed in chapter 2 in the section on situated cognition and cognitive apprenticeships.[Brown]

The Structure of Student Expectations: The Hammer Variables

In order to get a handle on these complex issues, we need to begin defining specific characteristics so that we can talk about them and begin to think about ways to further them with instruction. In a series of interesting papers, David Hammer has begun this task.[Hammer 1996] [Hammer 1996a] [Hammer 1997] In these papers, he identifies a number of parameters that arise from the expectations and epistemologies that a student brings into the physics class. Hammer's three variables are listed in table 2.

I refer to these attitudes as *favorable* or *unfavorable*, since to make reasonable progress towards becoming a scientist or engineer, a student will find unfavorable attitudes limiting and will have to make a transition to the attitudes listed in the favorable column.

Sagredo complains, "I certainly expect my students to have the attitudes that you call favorable when they enter my class. If they didn't learn these attitudes in school, what can I do about it?" One of the problems, Sagredo, is that we often actually encourage unfavorable attitudes without really being aware of it. While working on his dissertation, Hammer did a case study with two students in algebra-based physics at Berkeley who were carefully matched as to grade point average, SAT scores, etc., but who had decidedly different approaches to learning physics.[Hammer 1989] The first student tried to make sense of the material and integrate it with her intuitions. She didn't like what she called "theory" by which she meant

> *"...it means formulas...let's use this formula because it has the right variable, instead of saying, OK, we want to know how fast the ball goes in this direction... I'd rather know why for real."*

	Favorable	***Unfavorable***
independence	takes responsibility for constructing own understanding	takes what is given by authorities (teacher, text) without evaluation
coherence	believes physics needs to be considered as a connected, consistent framework	believes physics can be treated as unrelated facts or independent "pieces"
concepts	stresses understanding of the underlying ideas and concepts	focuses on memorizing and using formulas without interpretation or "sense-making"

Table 2: The "Hammer variables" describing students expectations. . [Hammer 1996]

The second student was not interested in making sense of what she was learning. For her, the physics was just the set of formulas and facts based on the authority of the instructor and text. Consistency or sense-making had little relevance.

> *"I look at all those formulas, say I have velocity, time, and acceleration, and I need to find distance, so maybe I would use a formula that would have those four things.*

"Student A was able to make sense of the material for the first few weeks. Soon, however, she became frustrated, finding it difficult to reconcile different parts of the formalism with each other and with her intuition. Eventually she compromised her standards in order to succeed. Student B's failure to seek consistency or understanding did not hurt her in the course.

This small example indicates that we may inadvertently wind up encouraging students in holding unfavorable attitudes. After learning about these issues, I tried to change the way I taught in order to change this situation. How one might do this is discussed in chapter 8 on homework and testing and in chapter 9 on surveys and assessing our instruction. I used the MPEX survey we developed to test student expectations (described in chapter 9 and given in the Appendix). Although at first I didn't get improvement, I learned that at least my grades were somewhat correlated with the results on my survey whereas those of my colleagues were not. This can be taken in two ways! Either my survey is not measuring something we want students to learn, or our classes are not rewarding those behaviors we want to encourage.

As we begin to develop a more complex view of what is going on in a physics class, what we want the students to get out of it, and what we value, we begin to realize that sometimes "the right answer" is not the only thing we should be looking for. A dramatic demonstration of student variability on attitudinal issues and how these issues play out in a classroom setting is given by Hammer's analysis of a discussion among a teacher and a group of high school students trying to decide whether a ball rolling on a level plane would keep moving at a constant speed. [Hammer 1996] The students had been told the arguments made by Galileo that under ideal conditions it would do so.[6] I've numbered the lines in the discussion so we can refer to them later.

1. Prior to this moment, the debate had mostly focused on the question of whether it is friction, gravity, or both that causes the ball to slow down. The students also debated whether it is appropriate to neglect friction or gravity, or both, and whether it is possible to neglect one without neglecting the other.
2. About 20 minutes into the debate, Ning argued that Galileo's ideal conditions would mean no forces on the ball, including no friction and no gravity; and, she claimed, "*if you don't put any force on it, it's going to stay still or go at constant speed*." Bruce elaborated on Ning's statement, adding that there must be a force to make the ball move:
3. *Bruce: If there is no gravity and no friction, and there is a force that's making it move, it's just going to go in a straight line at a constant speed. . . . What's making the ball move?*
4. *Amelia [over several other voices]: The forces behind it.*
5. *Susan: He [Galileo] said there was no force.*
6. *Bruce: If there's no force pulling it down, and no force slowing it down, it would just stay straight.*
7. *Harry: The ball wouldn't move.*
8. *Jack: There's no force that's making it go.*
9. *Steve: The force that's pushing it.*
10. *Bruce: The force that's pushing it will make it go*
11. *Jack: Where'd that force come from, because you don't have any force.*
12. *Steve: No there is force, the force that's pushing it, but no other force that's slowing it down.*
13. Many voices at once, unintelligible. Sean says he has an example.
14. *Teacher: Sean, go ahead with your example.*
15. *Sean: If you're in outer space and you push something, it's not going to stop unless you stop it.*
16. *Teacher: If you're in outer space and you give something a push, so there's a place with no gravity -*
17. *Sean: No gravity, no friction.*
18. *Teacher: - it's not going to stop until you stop it. So Penny what do you think about that?*
19. *Penny: But we talked about the ball on [a surface], but when we talk about space, it's nothing like space. So I was just saying that gravity will make it stop.*
20. Amelia objected to Sean's example for another reason, saying that something moving in space will still stop.
21. *Amelia: No. Maybe there's no gravity and no air there, but there are other kinds of gases that will stop it.*
22. *Teacher: But those are other, those are outside things.*
23. *Amelia: The outside friction should stop it.*
24. *Bruce: That's not, that makes it an un-ideal state.*
25. *Scott: Space is a vacuum. Like a vacuum there's no -*
26. *Amelia: There are other kinds of gases.*
27. [Several voices, unintelligible.]
28. *Harry: We're talking about ideal space. (students laugh)*
29. I intervened at this point to steer the discussion away from the question of whether there are gases in space and toward the question of whether there is a "force that's moving" the ball.
30. *Teacher: . . . So how can one side say there are no forces on it, and the other side say there is a force that's moving it.*
31. *Bruce: Well there was an initial force.*
32. *Susan: There's an initial force that makes it start, giving it the energy to move.*

[6] Student names are pseudonyms.

In analyzing this discussion, Hammer identifies half a dozen perspectives that could be used to evaluate the students' responses. I want to focus on four.

- *Content answer:* Does the student have the correct answer?
- *Reasoning:* Does the student display a common naïve conception? Is it related to a reasoning primitive?
- *Coherence:* Does the student understand that scientific laws are developed to unify a wide variety of circumstances and that science should be consistent?
- *Understanding idealizations:* Can the student see the relevance of idealized or limiting conditions?

In the dialog, Ning gave the correct answer (line 2) but did not participate in defending it. The discussion revealed that many of the students had the common naïve conception represented by the facet "motion is caused by force" (lines 3, 8/11, 12). Almost all of the discussion was by claim and counter-claim without citing reasoning or evidence. The discussion in lines 15-19 shows a distinction between Sean, who is trying to make a link between two rather different physical situations and Penny, who wants to keep them separate. This can be interpreted as a difference in their understanding of the need for coherence in science. Sean's claim in line 15 tried to take the analysis to an idealized situation, without gravity or friction. Amelia (lines 23 and 26) did not appear to be comfortable in thinking about the simplified example.

In other examples cited by Hammer, students gave the correct answer to a problem, but argued its validity by citing the text or teacher and being unwilling to think about the issue for themselves.

These examples illustrate the complexity of our hidden curriculum and show how we can begin to think both about what the student is bringing in to our classes and what the student can gain from our classes in a more sophisticated way than just "are they giving the right or wrong answers."

Reflection: Thinking about thinking

The transcript from David Hammer's high school class in our discussion above shows that different students access different kinds of reasoning in their discussion of a physics problem. This variety arises from students having different expectations about the nature of science and what it means to learn science. Unfortunately, many of these expectations are inappropriate for learning science. They may be learned in school, from movies and TV, or from reading science fiction books.[7] When students have the wrong expectations about what they are supposed to do in a class, those expectations can serve as a filter, causing them to ignore even explicit instructions given by the instructor.

Most of my students expect that all they have to do to learn physics is read their textbooks and listen to lectures. Although some students who believe this don't actually carry out this minimal activity, even those who do often fail to make sense of physics in the way I want them to. This leads me to believe that reading textbooks and listening to lectures is a poor way of learning for most students. Sagredo objects, "This is clearly not universally true!" Remembering principle 4, I concur. As physics teachers, most of us have had the experience of having a few "good" students in our lectures — students for whom listening to a lecture is an active process — a mental dialog between themselves and the teacher. Indeed, many of us have been that good student and remember lectures (at least some of them) as significant parts of our learning experience.[8]

A similar statement can be made about texts. I remember with pleasure working through texts and lecture notes, reorganizing the material, filling in steps, and posing questions for myself to answer. Yet few of my students seem to know how to do this or even to know that this is what I expect them to do. This leads us to think about an additional observation.

> *Many of our students do not have appropriate mental models for what it means to learn physics.*

This is a "meta" issue. People build schemas not only for content, but also for how to learn and what actions are appropriate under what circumstances. Most of our students don't know what you and I mean by "doing" science or what we expect them to do. Unfortunately, the most common mental model for learning science in my classes seems to be:

- Write down every equation or law the teacher puts on the board that is also in the book.

---

7 Some science fiction books, especially those written by scientists (such as David Brin, Gregory Benford, or John Kramer) have excellent descriptions of the way science develops its knowledge.

8 In many research groups, a seminar more resembles a discussion than a lecture. These can be very active learning experiences, both for the speaker and the listener. However, see the discussion of the traditional lecture in chapter 11.

- Memorize these, together with the list of formulas at the end of each chapter.
- Do enough homework and end-of-the-chapter problems to recognize which formula is to be applied to which problem.
- Pass the exam by selecting the correct formulas for the problems on the exam.
- Erase all information from your brain after the exam to make room for the next set of material.

I call the bulleted list above *the dead leaves model*. It's as if physics were a collection of equations on fallen leaves. One might hold $s = \frac{1}{2}gt^2$, another $\vec{F} = m\vec{a}$, and a third $F = -kx$. These are each considered as of equivalent weight, importance, and structure. The only thing one needs to do when solving a problem is to flip through one's collection of leaves until one finds the appropriate equation. I would much prefer to have my students see physics as a living tree!

In part, these approaches to learning physics arise from a misunderstanding of the nature of scientific knowledge and how one has to learn it. As pointed out so clearly by diSessa and discussed in chapter 2, for most ordinary people (even for some of our best students[9] knowledge of the world comes in "pieces" about how particular situations work.[diSessa 1993] [diSessa 1988] As pointed out by Reif and Larkin, [Reif 1991] building a consistent and economical set of principles — at the cost of in many cases requiring long and indirect explanations of many phenomena — is not the way most people create their models of the physical world in their everyday lives. It seems that quick and direct explanations are what people tend to look for. (See table 1.) The complex consistent and parsimonious net of links built by science is not a natural type of mental construction for most people. It has to be learned.

The key element in the mental model I want my students to use in learning physics appears to me to be *reflection* — thinking about their own thinking. This includes a variety of activities including evaluating their ideas, checking them against experience, thinking about consistency, deciding what's fundamental that they need to keep and what is peripheral and easily reconstructed, considering what other ideas might be possible, and so on.

[9] Recall that in [diSessa 1993] the subjects studied were MIT freshman.

My experience with students in introductory classes — even advanced students[10] — is that they rarely expect to think about their knowledge in these ways. Students often come to my office hours for help with problems. I always ask them to show me what they have tried so far and proceed to offer help via questions. They frequently have an error close to the start of their analysis — in a principle or equation that they bring up from their memory. As I lead them to implausible and unlikely results through my questioning they become troubled, but they are much more likely to try to justify a ridiculous result by difficult and inconvenient contorted reasoning than by asking if one of their assumptions might be wrong!

From our cognitive model we understand that to create new, coherent, and well-structured mental models, students need to go through a number of well-designed activities addressing the issue to be learned, to repeat them, and to reflect on them. Similar principles hold for *metacognition* — thinking that acts on the thinking process itself. I add an additional learning goal to the list developed in chapter 3.

Goal 4: Metalearning — Our students should develop a good understanding of what it means to learn science and what they need to do to learn it. In particular, they need to learn to evaluate and structure their knowledge.

This is not a trivial goal and it does not happen automatically for most students as they work to learn physics content.

> *In order for most students to learn how to learn and think about physics, they have to be provided with explicit instruction that allows them to explore and develop more sophisticated schemas for learning.*

"Hold on!" Sagredo complains. "I never have time enough to teach all the content I'm supposed to teach. How can I find time to give them lessons in how to learn?" I sympathize, Sagredo. But in fact, the problem is not as bad as it looks. If we are teaching them to learn, we have to be teaching them to learn *something*. That something can easily be the appropriate physics content. Some introductory discussion, lessons designed to encourage particular activities, and reflections analyzing what they've done

[10] Many of the students in my algebra-based physics classes are upper division students who have previously taken many science classes in chemistry and biology.

should help substantially. Specific instructional techniques focused on learning to learn are discussed at the end of this chapter.

Connecting to the Real World

Although physicists believe that they are learning about the real world when they study physics, the context dependence of cognitive responses (see chapter 3) opens another possible gap between faculty and students. Students may believe that physics is related to the real world in principle, but they may also believe that what they are learning in a physics class has little or no relevance to their personal experience. This can cause problems that are both serious and surprising.

Even if our students develop strong concepts related to real-world meanings, the strong context dependence of the cognitive response makes it particularly easy for students to restrict their learning in physics classes to the context of a physics class. This seems unnatural to Sagredo. "Practically every problem I assign for homework or do on the board involves some real world physical context." True, Sagredo. But that doesn't mean that students will easily or naturally make the connections that you do.

When an instructor produces a demonstration that has been "cleaned" of distracting elements such as friction and air resistance, the instructor may see it as displaying a general physical law that is present in the everyday world but that lies "hidden" beneath distracting factors. The student, on the other hand, may believe that the complex apparatus is *required* to produce the phenomenon, and that it does not occur naturally in the everyday world, or is irrelevant to it. A failure to make a link to experience can lead to problems not just because physics instructors want students to make strong connections between their real-life experiences and what they learn in the classroom, but because learning tends to be more effective and robust when linked to real and personal experiences.

Even worse, students' failure to connect their personal experience to what is happening in their physics class can put up barriers to understanding that grow increasingly impenetrable. As discussed in chapter 5, multiple representations are used in physics in order to code knowledge in a variety of interlocking ways. A critical element in all of them is the map to the physical system. An essential part of solving a problem is understanding what the real world version of the problem is, what's important in that situation, and how it maps onto physical principles and equations. If students don't understand that part of the process, they can have great difficulty in seeing the physics as a way to make sense of the physical world.[11]

A classic word problem that illustrates this difficulty is shown in figure 1.

A shepherd has 125 sheep and 5 dogs. How old is the shepherd?

Fig. 1: A word problem for middle-school math students.

Although this problem is patently absurd and cannot be answered, some middle-school students will struggle to find an answer (Expectation: "The teacher wouldn't give me a problem that has no solution.") and will come up with an answer of 25. ("There are only two numbers to work with: 5 and 125. Adding, multiplying, and subtracting them doesn't give something that could be an age. Only dividing gives a plausible number.")

Another example comes from the mathematics exam given by the National Assessment of Educational Progress (NAEP). A national sample of 45,000 13-year-olds was given the problem shown in figure 2.[Carpenter]

An army bus holds 36 soldiers. If 1,128 soldiers are being bused to their training site, how many buses are needed?

Fig. 2: A problem for the NAEP math exam for middle-school students.

Although 70% of the students who worked the problem carried out the long division correctly, only 23% gave the correct answer — 32. The answer "31 remainder 12" was given by 29% and the answer 31 was given by another 18% of those doing the problem. Thus, nearly half of the students who were able to carry out the formal manipulations correctly, failed to perform the last simple step required by the problem: to think about what the answer meant in terms of a real world situation. (Expectation: "The mathematical manipulation is what's important and what is being tested.")

In these two examples, students are making somewhat different errors. In the shepherd problem they <u>are</u> using some real world information — what ages are plausible as answers; but they are not asking how the numbers they are given could relate to the answer. They are not making sense of the problem. In the soldiers and buses problem, students are <u>not</u> using their real-world knowledge that you cannot rent a fraction of a bus. In both

---

[11] The Physics Education Group at the University of Massachusetts - Amherst has done interesting research using problem posing as a technique to help students develop these skills.[Mestre] See also the variety of problems discussed in chapter 8.

cases, students who make these errors focus on the mathematical manipulations and fail to "make sense" of the problem in real-world terms.

The same problems occur frequently in introductory physics. In my experience with introductory college physics, more than half of the students do not spontaneously connect what they learn in their physics class to their everyday experiences — either by bringing their everyday experiences into their physics classes or by seeing the physics they are learning in the outside world. Two anecdotal examples of this show how this plays out in a college physics class.

A student in my algebra-based physics class missed a mid-semester exam due to an illness and I agreed to give her a makeup exam. One of the problems on the exam was the following. "A high jumper jumps so his center of gravity rises 4 feet before he falls back to the ground. With what speed did he leave the ground?" This is a typical projectile problem. My student knew the formula and punched the numbers into her calculator. When she handed in her test and I looked over her answers she had come up with the answer 7,840 feet/second. (Can you guess what she had done wrong on her calculator?) I asked her whether her answer to that problem had bothered her. She shrugged and said, "That's what the formula gave me." She saw absolutely no need to check her answer against her experience — and incidentally, it had never entered her mind that she might have mis-remembered the formula, incorrectly recalled the value of a parameter, or made an error in pressing the calculator keys. This overconfidence in their memory and processing is a symptom I have seen in very many students. They assume anything they remember must be correct.

A second example occurred in my engineering (calculus-based) physics class. For many years now, I have been requiring estimation (Fermi-type) problems in my classes.[12] Almost every homework assignment has one and every exam is guaranteed to have one. One of my students came into my office hours and complained that this wasn't fair. "I don't know how big these things are," she scowled. "Well," I said. "How about a foot? Do you know how big a foot is?" "I have no idea," she replied. Assuming that she was overstating her case to make her point, I said, "How about making a guess? Show me how far up from the floor a foot would be." She placed her hand at about waist level. "And how tall are you?" I asked. She thought for a second, said "Oh" and lowered her hand somewhat. She thought again and lowered her hand again — to about the right height above the ground. She looked at her hand — and at her foot a few inches away and remarked with great (and what appeared to be genuine) surprise, "Oh! Does it have anything to do with a person's foot?"

Since these real-world connections turn out to be critically important in developing an understanding of how physics helps us to make-sense of our everyday experiences,[13] I specify a fifth learning goal.

<u>Goal 5:</u> *Reality Link — Our students should connect the physics they are learning with their experiences in the physical world.*

To what extent does a traditional course satisfy this goal? There are a number of ways of probing these issues (see chapter 4). The simplest is to ask them.[14] In our study of student expectations in a calculus-based physics class for engineers [Redish 1998], using the MPEX survey (see chapter 9 for a detailed discussion) we found that student expectations of the connection between physics and the real world typically tended to deteriorate as a result of the first semester of instruction. The four items of the MPEX reality cluster are shown in table 3. The ask whether the student expects to / has needed to[15] make the link to their outside experiences for the class and whether the student expects to / has found that what they learn in physics can be seen in their real world experiences. Both issues are addressed in two statements, one positive and one negative. The student's response is considered to be *favorable* if she sees the need for a connection and *unfavorable* if she does not. The polarity of the favorable result is indicated after the item by a (+) when the favorable result is *agree* and by a (-) when the favorable result is *disagree*. The students are asked to report on a 5 point scale (strongly agree, agree, neutral, disagree, strongly disagree) but for a favorable / unfavorable analysis, we ignore whether or not there is a "strongly". The responses come from pre and post surveys given in my first semester of an engineering physics class. The class was calculus-based and covered mostly Newtonian mechanics. The results are shown for N = 111 students (matched, i.e., who completed both pre and post surveys).[16]

---

12 For examples of these types of problems, see the discussion in chapter 8 and the sample problems in the Appendix.

13 This is especially true for our service students in engineering and biology.

14 This method is not very accurate since students often do not reflect and do not necessarily know how they think. A better approach is to watch them solving problems alone or in a group using think-aloud protocols. (See chapter 4.)

15 The alternate forms are for the pre and post class surveys.

16 A total of 158 students completed the class.

The results are discouraging, especially on the last two items. I tried to help them make the connection by giving some estimation problems, but that was clearly insufficient. Similar results have been found with other faculty teaching this class at Maryland and at many other colleges and universities. [Redish 1998]

MPEX Item	***Fav. Pre***	***Unfav. Pre***	***Fav. Post***	***Unfav. Post***
Physical laws have little relation to what I experience in the real world. (-)	84%	5%	87%	2%
To understand physics, I sometimes think about my personal experiences and relate them to the topic being analyzed. (+)	59%	11%	54%	22%
Physics is related to the real world and it sometimes helps to think about the connection, but it is rarely essential for what I have to do in this course. (-)	73%	9%	61%	19%
Learning physics helps me understand situations in my everyday life. (+)	72%	10%	51%	18%

Table 3: Results on the MPEX Reality Link cluster items in a calculus-based first-semester physics class for engineers. (N=111, matched data) The polarity of the favorable answer is indicated in parentheses: (+) implies "agree" is favorable, (-) implies "disagree" is favorable.

There has been little published work on how to help students achieve the goal of this section. In my experience, regular essay questions asking the students to relate the physics to their experience and regular estimation questions (being sure to include both on every exam so that students take them seriously) only help a little bit. Even in lessons where physicists see real-world implications immediately, students rarely make the connections spontaneously if not led to them. I expect this goal will only be achieved by a thorough interweaving of the physics with explicit connections to the students' experience.[17] Further research and development on this issue would be most welcome.

Affect: Motivation, Emotion, and Self-evaluation

It is patently clear to most university physics instructors that motivation, how students feel about the class, and how the students feel about themselves play a significant role in how students respond to instruction and how well they learn. The issues of feeling, emotion, and mood are summarized by the term *affect* or *affection* in psychology. These issues have been discussed extensively in the educational literature, [Graham] [Stipek] but I do not attempt to review this literature here as it is my sense that it does not yet meet my "triangulation" conditions of a convergence being achieved between researchers in neuroscience, cognitive psychology, and education. In addition, the interaction between affect and cognition is extremely complex and it is difficult to provide intellectual guidance. This is not to say these issues are not of great importance. I therefore make a few comments, but refer the reader to the literature cited above for more details.

Motivation

Motivation can be a major factor in distinguishing students who will make the effort to learn and those who will not. We encounter a variety of motivations.

- *Internally motivated* — Some students who come to our classes are self-motivated by an interest in physics and a desire for learning.
- *Externally motivated* — Some students have no internal interest in physics but are strongly motivated to get a good grade because our class is hoop that must be jumped through in order to get into a program for which they are motivated.
- *Weakly motivated* — These students are taking physics because it is a requirement but they only are concerning about passing, not getting a good grade.
- *Negatively motivated* — Some students are motivated to fail — for example, in order to demonstrate to a controlling parent or mentor that they are not suited to be an engineer or a doctor.

17 Preliminary results with a more synergistic approach appear quite favorable. [Redish 2001]

Those in the first group are a physics instructor's delight. Whatever you give them they make the most of. We can work with those in the second group by controlling the learning environments we set up and making clear what will be evaluated on exams. (See examples in chapter 8.) I can rarely do much with the last group. Their goals in the class are distinctly different from mine.

Finding ways to motivate your students to want to learn physics can be an extremely effective lever to improve the success of your teaching. Unfortunately, this is easier said than done and is where much of the "art" in teaching comes in. It is easy to mistake student happiness for student motivation. Making your lecture "entertaining" does not necessarily increase students' motivation for learning. Indeed, it can set up the expectation in their mind that matches your lecture with a TV program where they don't have to think.

Providing connections to their chosen career might help. I evolve my estimation problems into design problems in my engineering physics class and create problems with a medical and biological context for my algebra-based students. I hope this helps them see the relevance of physics towards a profession towards which they should, in principle, be motivated. (Interviews with a small number of volunteers — usually the better students — suggests that at least this group is making the connection. [Lippmann])

Motivation is perhaps the primary place where the teacher in fact makes a significant difference. A teacher with the empathy and charisma to motivate the students can create substantially more intellectual engagement than one who reads from the book and does not take the time to interact with the students. Perhaps the most critical element in creating motivation is showing your students that you are interested in them, you want them to succeed, and you believe that they can do it.

Self-Image

Sagredo is a bit skeptical about the issue of students' self-image. He feels that the education community pushes "helping students feel good about themselves," sometimes to the detriment of serious critical self-analysis and learning, at least if the letters to the editor published in newspapers are to be believed. In my experience with university level physics students this issue cuts two ways. Some students are supremely overconfident while others think that they cannot possibly understand physics. Both groups are difficult to deal with.

In our small group learning sessions we use the Tutorial materials developed at the University of Washington. These lessons are research based group-learning worksheets (see chapter 12 for a detailed description) and use a cognitive conflict model. As a result, students who are used to being right often feel the Tutorials are trivial and therefore useless — even when they are consistently getting the wrong answers. When I am facilitating in one of these sessions I see this as a terrific learning opportunity. I circulate through the class, asking what they got on the tricky questions. When I find a group that has been overwhelmed by an overconfident student with a wrong answer I say, "Now remember: Physics is not a democracy and physics is not determined by charisma. You can't tell who's right by who says it the most forcefully or by what most people think. It has to make sense and it has to be consistent. Perhaps you want to go back and think that question out again." The result is almost always that someone else in the group who had previously been intimidated into silence can bring everyone to the correct result. This sends a really useful message — both for the overconfident student and for the other members of the group.

On the other side I have had experience with students who were absolutely convinced that they were incapable of learning physics. In one case, I had a student in algebra-based physics who was convinced "she couldn't do this stuff" and told me so repeatedly. On the other hand, I often watched her vigorously argue difficult issues in Tutorial with another student who was sublimely confident of her ability and answers. My underconfident student was almost always right and my overconfident student almost always wrong.

I was not successful convincing the student in the above story about her ability and she did poorly on exams. In other cases, I was able to help students who were good in other classes, but who, perhaps because of bad experiences in high school, were convinced that they "couldn't do physics." All these cases are best treated carefully and individually, in my opinion, using all the empathy and understanding you can bring to bear. Unfortunately, in many college and university situations, the pressure of time and numbers makes it difficult if not impossible to allow one to offer the individualized responses needed.

There has been some research on the topic of math anxiety or "math phobia". See for example [Tobias 1995]. I do not know of comparable work on "science phobia."

Affect

"I'm a physicist, not a song-and-dance-man!" Sagredo complains, echoing Star Trek's Dr. McCoy. Perhaps, Sagredo, but making your students feel good about your class can have an influence on their learning. For one, if they hate your lectures and don't come to class they won't be able to learn anything from them.[18] On the other hand, if you fill your lecture with jokes, films, and cartoons, they are unlikely to take them seriously.

The best thing you can do to make students "feel good" about your class is to make it worthwhile, at an appropriate level, and fair. Students like to feel that they are learning something valuable and that they can get a "good" grade (this may have different meanings for different students) without having to work so hard that their other classes (and their social life) suffers. Getting students to learn a lot from our classes is a process of negotiation. As a teacher, I want them to work hard, but as a student, they don't want to work hard without a clear payoff. In physics, learning can be frustrating and non-linear. Often you have to work for a long time without felling that your making much progress. Then, suddenly, everything falls into place and it all makes sense. But until the "click," you couldn't be sure how much time you would need to "get it" and it's difficult to plan. Students have to first learn what understanding the physics feels like and be slowly drawn into working hard enough to learn harder and harder topics.

But entertainment and "song-and-dance" don't' have to be shunned, Sagredo. In our context it can mean little physics jokes, personalized stories, and dramatic demonstrations. All of these can be effective — or not. Jokes should be relevant, not off-color, and not derogatory to groups or individuals. Personalized stories should be relevant to the physics involved and have some point that will make sense to a novice. Demonstrations can be the best but are also dangerous. As explained in chapter 3, demonstrations can be entertaining but misleading. Students often don't see what you think they are seeing. A careful and involving class discussion, both before and after the demonstration are usually needed.

The most entertaining and dramatic demonstration I use in my classes is the electromagnetic can crusher. In the lecture demonstration incarnation at the University of Maryland a 400 microfarad capacitor is charged to 3000 volts (storing 1.8 kilojoules) and is discharged through a three-turn coil into which an aluminum soft drink can has been positioned, as shown in photo in figure 3. With the circular windows open, as in the photograph at the left, the two pieces of the can are blasted over thirty feet to the sides of the large lecture hall with a very loud noise. Charging the capacitor to less voltage results in a can with a "waist," as seen in the photograph in figure 3 at the right.

Students always remember this one, even long after the class. The trick is to try to tie some real learning to the demonstration. In my experience, if I explain why the can is crushed (The collapsing magnetic field produces an EMF that induces a large current circulating around the can. That current then feels the magnetic force and is pushed inward.) students still want to know why the can is thrown outward. This leads to a discussion of the fringing fields whose directions and resulting forces can be worked out in detail — a very entertaining and satisfying exercise.

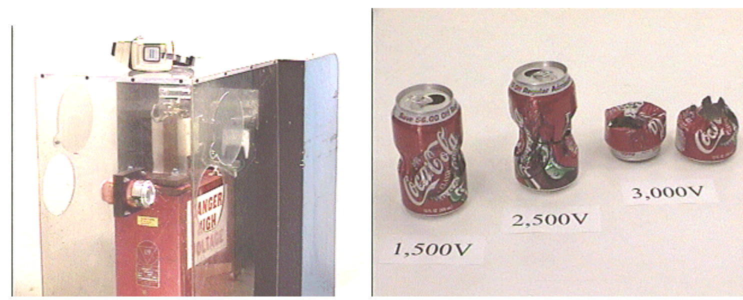


Fig. 3: An entertaining lecture demonstration: crushing a can with an EM field.

Instructional Methods That Can Help

In Recitation or Small Class: Group Problem Solving

Alan Schoenfeld, in a problem-solving college math class, developed a group-problem solving method that focused on helping students strengthen their judgment and control. The class was small enough (on the order or less than 25 students) that he could use a guided cooperative-group problem-solving approach.[19]

---

[18] Students tend to learn little from lectures anyway unless special tools are used. See chapter 11.

[19] See chapter 12 for a discussion of a method of this type employed in physics to help develop students' conceptual development and problem-solving skills,

In his observations of the class's behavior, Schoenfeld found that they often wasted a lot of time in following unproductive approaches through a lack of metacognitive activity. They quickly jumped on the first idea that came to their minds and then proceeded to "churn" through extensive manipulations, frequently losing track of what they were doing and rarely evaluating whether their approach was productive or not.

What (exactly) are you doing? (Can you describe it precisely?)
Why are you doing it? (How does it fit into the solution?)
How does it help you? (What will you do with the outcome when you get it?)

Fig. 4: Schoenfeld's questions for helping students learn to focus on metacognitive issues.

Schoenfeld developed an instructional method to help students become more metacognitively aware. The key was the mantra of metacognitive questions posted on the wall shown in figure 4. His comments on how this worked are worth repeating.

> *"Students' decision-making processes are usually covert and receive little attention. When students fail to solve a problem, it may be hard to convince them that their failures may be due to bad decision-making rather than to a lack of knowledge. The instructor had the right to stop students at any time while they were working on the problems and to ask them to answer the three questions on [figure 4]. At the beginning of the course the students were unable to answer the questions, and they were embarrassed by that fact. They began to discuss the questions in order to protect themselves against further embarrassment. By the middle of the term, asking the questions of themselves (not formally, of course) had become habitual behavior for some of the students…"*

He not only implemented a focus on metacognition and control in the group activity, but he modeled it in his approach to modeling solutions for the class as a whole. His description outlines the process in detail.

> *When the class convened as a whole to work problems (40-50% of class time), I served as orchestrator of the students' suggestions. My role was not to lead the students to a predetermined solution,…my task was to role model competent control behavior – to raise the questions and model the decision-making processes that would help them to make the most of what they know. Discussions started with 'What do you think we should do?' to which some student usually suggested 'Let's do X.' Of ten the suggestion came too rapidly, indicating that the student had not adequately thought through what the problem called for or how the suggestion might be useful. The class was then asked, 'Are you all sure you understand the problem, before we proceed with X?' A negative response from some students would result in our taking a closer look at the problem. After doing so, we returned to X as a possible solution approach. Did X still seem reasonable? Not infrequently the answer was 'no.' When it was, this provided the opportunity to remind students about the importance of making sure that one has understood a problem before jumping into its solution…After a few minutes of working on the problem – whether or not we were on a track that would lead to a solution – the process would be halted for an assessment of how things were going. The class was asked 'We've been doing this for 5 minutes. Is it useful, or should we switch to something else? (and why?)' Depending on the evaluation, we might or might not decide to continue in that direction: we might decide to give it a few more minutes before trying something else. Once we had arrived at a solution, I did a post-mortem on the solution. The purpose of that discussion was to summarize what the class had done and to point out where it could have done something more efficiently, or perhaps to show how an idea that the class had given up on could have been exploited to solve the problems …The same problem was often solved three or four different ways before we were done with it."*

[Schoenfeld 1985 p. 221-222]

Topics in Geoscience Research

William Slattery

Conference on "Bringing Research in Learning to the Geosciences"

I've read the questions to be addressed at the conference, and I'm sure they will lead to very interesting and informative discussions. I would ask that we consider how we can bring research on learning to the preparation of pre-service teachers and professional development to in-service classroom educators. Since K-12 teachers play such an important role in developing a scientifically literate population it's critical that we consider how best to bring research on learning to present the perspectives of Earth science and the Earth system to them.

I would also suggest that we consider two questions specifically dealing with on-line K-12 teacher professional development. These two questions are specifically focused on K-12 teachers, but the first is also apropos to on-line instruction in general:

- What is the most effective way to structure on-line professional development courses for K-12 teachers?

- How do K-12 teacher understandings of the Earth as a system translate into effective learning for K-12 students, integrating aspects of Physical, Life, and Earth/Space science?

Developmental Science 3:3 (2000), pp 247–286

ARTICLE WITH PEER COMMENTARIES AND RESPONSE

Seeing the big picture: map use and the development of spatial cognition

David H. Uttal

Northwestern University

Abstract

This paper considers the relation between the development of spatial cognition and children's use of maps and models. A new theoretical perspective is presented that takes into account the influences of maps on the development of spatial cognition. Maps provide a perspective on spatial information that differs in important ways from the perspective gained from direct experience navigating in the world. Using and thinking about maps may help children to acquire abstract concepts of space and the ability to think systematically about spatial relations that they have not experienced directly. In addition, exposure to maps may help children to think about multiple spatial relations among multiple locations. The results of previous studies that have demonstrated developmental differences in children's cognition of large-scale environments are examined from this theoretical perspective. This review suggests that the development of spatial cognition consists partly of the acquisition of models of large-scale space, and that maps influence the development of the modern Western model.

While making one of his first orbits of the earth, John Glenn said, 'I can see the whole state of Florida laid out just like on a map' (Wilford, 1981). Glenn was referring to the unique perspective that one can gain by viewing the earth from space. Looking down at the earth allows us to see spatial information, such as the layout of an entire state, that can be almost impossible to perceive directly while navigating.

Although Glenn's observations were about traveling in space, his comments are also a testimony to the importance of maps. Like space travel, maps provide a perspective that can be difficult to acquire from direct experience navigating in the world. Moreover, maps can alter how we think about the represented information. For example, we may realize from looking at a map that our hometown is much closer to, or farther from, another city than we previously thought. Maps allow us to transcend our direct experience of the world and think systematically about relations among multiple locations (Downs & Liben, 1993; Wood, 1992; Liben, 1999).

The purpose of this paper is to present a new perspective on the development of children's conceptions of large-scale space that considers the influences of maps and other external representations of space. The central thesis is that the development of mental representations of large-scale space is affected by the symbolic representation of spatial information on maps (Cassirer, 1944; Cassirer, Mannheim & Hendel, 1965; Ong, 1982; Liben & Downs, 1989, 1992; Olson, 1994; Hutchins, 1995; Liben, 1999, in press). Several ways in which the use of maps could influence the development of spatial cognition are discussed, and cross-cultural, historical and developmental evidence concerning the effects of maps is presented.

Overview

My claims are based on the assumption that the relation between maps and the development of spatial cognition is reciprocal in nature (Liben & Downs, 1989, 1991; Gauvain, 1993, 1995; Liben, 1999, in press). As children acquire new and more sophisticated ways of mentally representing and using spatial information, their understanding of maps improves. Likewise, children's developing conception of maps affects how they understand

Address for correspondence: Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208-2710, USA; e-mail: duttal@northwestern.edu

and conceive of spatial information. In the words of Liben and Downs (1991), 'the particular place representations encountered – just like the particular 'real' environments encountered – will shape environmental knowledge and, relatedly, behavior' (p. 147).

Researchers have investigated children's understanding and use of a variety of external spatial representations, ranging from real geographic maps to simple scale models. Although these various symbolic representations differ in important ways, they all share an important similarity: each stands for a particular place, or set of places, in the world. This work has shown that, by the age of 3, young children can understand some of the basic, symbolic relations between maps or scale models and the referent spaces (DeLoache, 1987, 1989, 1991 1995; Dow & Pick, 1992). However, young children have much more trouble acquiring or using more complex spatial information from maps. For example, 3-, 4- and 5-year-olds often cannot discriminate the location of objects in a space solely on the basis of the locations of the objects on a map. In addition, children of this age often have difficulty interpreting and using scale relations (Presson, 1982; Liben & Downs, 1992, 1993; Blades & Cooke, 1994; Blades & Spencer, 1994; Uttal, 1994, 1996; Liben & Yekel, 1996; Bence & Presson, 1997).

The difficulties that young children experience in understanding some elements of maps have been attributed, in part, to limitations in encoding, remembering or understanding the relevant spatial concepts. On this view, children's limitations in understanding maps are caused by limitations in the processing or representation of spatial information. My goal here is to show that the opposite may also be true – children's limitations in thinking about spatial information may reflect a lack of understanding of the functions and uses of maps. Learning about maps therefore may influence the development of spatial cognition.

In one sense, the notion that our conceptions of the world are influenced by the representation of the world on paper (i.e. on maps) is quite familiar. All maps distort spatial information because it is impossible to capture all aspects of a three-dimensional world on a two-dimensional sheet of paper. People's conceptions of the relevant areas often reflect these distortions. For example, because the representation of Greenland is distorted on standard Mercator projections, many people believe that this island is much larger than it really is. These kinds of errors highlight that our conceptions of large-scale space are influenced by the representation of information on maps (see Tversky, 1981).

I argue here that there is also a second and more general sense in which maps influence thinking about spatial information: exposure to maps may help to instantiate a mental model or conception of large-scale space. This conception differs fundamentally from how space is experienced when navigating in the world. For example, the map-influenced conception of space is more abstract and less tied to direct experience of specific spaces. Many of the changes that have been observed in the development of children's cognition of the large-scale environment may stem from the acquisition of the map-influenced conception of the world.

This paper is organized into eight major sections. The sections focus respectively on the following issues: (a) definitions and restrictions of scope; (b) the theoretical background for the claim that maps influence how people think about space; (c) the acquisition of spatial information from maps; (d) the cognitive consequences of acquiring spatial information from maps; (e) cross-cultural and historical evidence; (f) previous characterizations of the development of large-scale spatial cognition; (g) research evidence that supports the claim that exposure to maps affects the development of spatial cognition; and (h) suggestions for testing the theory.

Definitions and restrictions of scope

The term map has been used in numerous ways in several different fields of study, including biology, psychology, geography, mathematics and many others. The variety of definitions can lead to substantial confusion (Liben & Downs, 1991). The focus here is primarily limited to what Liben and Downs (1991) have called *prototypical maps*. These include maps shown in common atlases, road maps, and the kinds of maps that regularly appear on the walls of school classrooms. More specifically, I am focusing on the cognitive and developmental consequences of exposure to maps that represent spaces from an overhead or oblique view and that follow specific projections. Projection refers to the use of geometric techniques to capture some (but not all) aspects of the spherical shape of the earth on flat, two-dimensional paper. Although all projections necessarily involve distortions of either area or distance (Dent, 1996), they all follow spatially metric, Euclidean rules for making the compromises between preserving shape and distance information. Other kinds of maps may also affect the development of spatial cognition, but prototypical maps may have a particularly strong influence on how children's thinking about spatial information develops.

There is also a limitation on the scale of space that is discussed. Much of the focus here is on spatial cognition of environments that are typically larger than can be perceived in a single glance. This includes the space just beyond one's current vista as well as regions, countries and the entire world (Spencer, Blades & Morsley, 1989; Montello, 1993; Montello & Golledge, 1999). Smaller spaces, such as a tabletop, are not a primary focus because maps typically are not used to learn about spaces of this size.

Theoretical background

The present theoretical perspective is consistent with previous research in psychology, philosophy and geography. Many researchers have suggested that the ways in which information is represented symbolically may influence how people think about the represented information (Cassirer, 1944; Vygotsky, 1962, 1978; McLuhan, 1964; Luria, 1976; Downs, 1981; Ong, 1982; Eliot, 1987; Lave, 1988; Liben & Downs, 1989, 1993; Miller, 1989; Karmiloff-Smith, 1992; Vosniadou & Brewer, 1992; Gauvain, 1993; Olson, 1994; Hutchins, 1995; Miller & Paredes, 1996; Crosby, 1998; Liben, 1999). Examples come not only from research on children's use of maps, but also from research on other symbol systems.

One relevant example concerns the influence of literacy on metalinguistic awareness. Becoming literate may influence cognitive development in general and metalinguistic awareness more specifically. Even though young children may be able to use complex grammatical structures from an early age, they remain relatively unaware of the formal analysis of language until they are asked to read and write (Karmiloff-Smith, 1979, 1990, 1992; Bialystock, 1993, 1995; Olson, 1994; Lee & Karmiloff-Smith, 1996). The acquisition of literacy may be responsible, at least in part, for the development of explicit knowledge of grammar, sentence structure, and other aspects of language (Olson, 1994, 1996).

A second set of relevant examples comes from the domain of mathematics development. For instance, several researchers have suggested that learning about the written representation of zero (0) influences the development of concepts of nothingness (Wellman & Miller, 1986; Olson 1996). Similarly, experience in using fractions helps children to acquire a concept of division (Stern, 1993; Staub & Stern, 1997). Likewise, learning to use an abacus presents an alternative model of numerical representation that is partly visual–spatial and is based on the base-5 number system (Stigler, 1984; Stigler, Chalip & Miller, 1986; Miller & Stigler, 1991). Like written language, mathematical symbols bring into consciousness information that would otherwise remain opaque or inaccessible (Bialystock & Codd, 1996).

The influence of maps and other external representations of spatial information may be analogous to the influence of other symbol systems on cognitive development. Perhaps the most general characteristic of maps is that they allow us to acquire, inspect and think about spatial information irrespective of navigation. Children may possess the ability to represent spatial information mentally in metric terms at an early age (e.g. Landau, 1986; Bushnell, McKenzie, Lawrence & Connell, 1995; Newcombe, Huttenlocher & Learmonth, in press), but this does not mean that they understand, or have explicit access to, knowledge of space *per se*. The understanding that space exists independent of our experience of it may come about as a result of exposure to maps. In addition, maps affect *how* we think about spatial information; maps may lead people to think about space in more abstract and relational ways than they would otherwise. For these reasons, maps can be construed as *tools for thought* in the domain of spatial cognition (Vygotsky, 1978; Liben & Downs, 1993; Miller & Paredes, 1996; Staub & Stern, 1997; Liben, 1999, in press). Maps provide a cognitive tool that helps children extend their reasoning about space in a new way. Over time, children can internalize the tool and think about space in map-like ways, even if they are not looking at a map at the time.

The present analysis is not meant to imply that exposure to maps is the only influence on the development of spatial cognition or that the course of the development of large-scale spatial cognition is totally open to the influences of maps. Other factors, including the nature of a person's experience, biological constraints and motivation may affect the development of spatial cognition. My claim is that the influence of maps is sufficiently great to merit a detailed consideration.

Acquiring spatial information from maps

In this section, I consider several ways in which acquiring knowledge from maps could influence what we know about the represented information. Maps help us to conceive of the world beyond immediate experience, they make different kinds of information perceptually available, and they depict spatial information in an abstract way.

Maps help us to conceive of the world beyond immediate experience

The most general characteristic of maps regarding the acquisition of spatial information is that they make us aware of the world beyond what we have experienced. Maps allow us to look at, and study, sets of spatial relations without actually navigating through the space.

In Wood's (1993) words, 'this, essentially is what maps give us, *reality*, a reality that exceeds our vision, our reach, the span of our days, a reality we achieve no other way' (pp. 4–5). Coming to know the world through the mediated view of maps may provide an important inroad into realizing that there are alternative ways of thinking about, and mentally representing, the world beyond one's own experience. Of course, the reality that maps provide is not a faithful copy of the world because all maps distort the information that they represent.

Maps make different kinds of information perceptually available

The effects of maps are not limited only to helping us learn that there is a world beyond our immediate experience. Maps also influence *how* we acquire spatial information. Learning about space from a map differs in important ways from learning about space from navigation (Thorndyke & Hayes-Roth, 1982; Presson, Hazelrigg & DeLange, 1987; Tversky, 1996). During navigation, we constantly change the relation between where we are in space and what we are viewing (Tversky, 1996). Different features of the environment, including landmarks, alternative routes and so forth, continuously come in and out of view. In contrast, maps provide a more static view of space because substantial portions of a map can be inspected in a single glance (Blaut, 1991; Hutchins, 1995). Furthermore, the oblique or overhead point of view from which maps are typically read provides a very different perspective than we can usually gain from navigation. Moreover, because maps depict space at a relatively small scale, they can *decrease* the salience of individual features. Looking at the depiction of a particular landscape feature on a map may reveal less about the characteristics of the individual feature but more about the relation between this feature and other features.

Taken together, these characteristics of maps allow us to gain visual access to a greater number of spatial relations than would be available from direct experience. Even if one's primary focus is simply to plan a route between two locations, the representational format and scale of the map will probably bring other nearby locations into view, which can then become a focus of attention and thought. Maps thus can facilitate the *discovery* of new information (Hutchins, 1995; Liben, 1999). For example, one can immediately acquire a sense of the spatial relations among the major cities in the USA by looking at Figure 1. Information about relations among multiple locations can be extremely difficult to gain from navigation, but maps, charts and graphs make this information readily available (Larkin & Simon, 1987; Gentner & Ratterman, 1991; Shin, 1994; Uttal & Gentner, 1995; Allwein & Barwise, 1996; Gattis & Holyoak 1996). Knowledge of relations among locations may be particularly important when we are asked to reason about directions among multiple locations or to plan alternative detour routes.

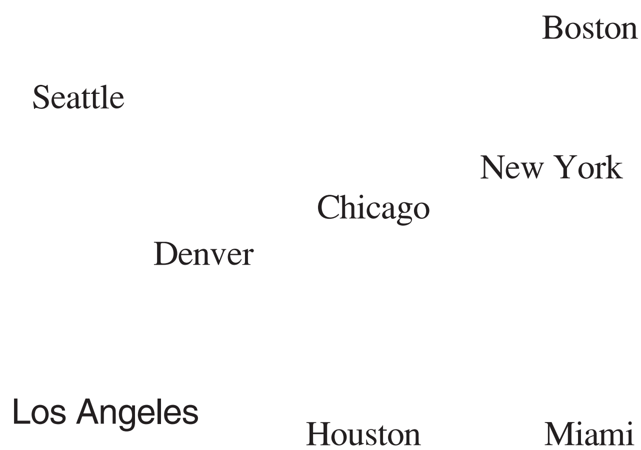


Figure 1 *A map of the (approximate) relative positions of major cities in the USA. The map highlights relations among locations.*

Maps highlight abstract space

Another important characteristic of prototypical maps is that they rely upon a geometry that is based on an *abstract* conception of space (Eliot, 1987; Jammer, 1993). An abstract conception of space assumes that space continues to exist independent of the objects that are contained within the space. Thus, space *per se* can become an object of discussion and thought, and it can be described and measured. Concepts such as kilometers, latitude, longitude etc. all imply an abstract model of space (Crosby, 1998).

That prototypical maps highlight abstract notions of space may play a particularly important role in the development of spatial cognition. As will be discussed below, one of the prevailing characteristics of young children's spatial thought is that they tend not to conceive of space in this abstract, formal way. Instead, young children know about a large-scale space primarily in terms of the particular locations that they have experienced directly. Exposure to maps therefore may play an important role in helping children to think about space in abstract ways.

In sum, maps bring into view spatial and geographic information that would otherwise remain opaque or inaccessible from direct visual experience, and moreover they facilitate thinking about the represented information. In addition, they highlight particular kinds of information, such as relations among locations. In the next section, I consider how these characteristics of maps might influence people's conceptions of large-scale space.

Cognitive consequences of map use

Taken together, the characteristics of maps that were discussed in the previous section may lead to the creation of a particular model of large-scale space. The map-influenced model may affect both how we think about spaces we already know and how we acquire information about new spaces. When adults enter a new city, they may carry with them not only actual maps but also the awareness of the possibility of 'mapping out' the city mentally. For example, knowledge of maps might provide a model that helps adults to think about locations in terms of distances, angles and multiple relations among multiple locations. In essence, experience with maps gives adults a framework of absolute space that they can use as they acquire information about new locations.

Indeed, the influence of maps and other external spatial representations can be so strong that they directly influence how we interpret reality. Hutchins (1995) provides an example from the domain of ship navigation. Navigators sometimes come to prefer the information that they gain from charts or maps over the information they gain from their own direct experience. Modern navigators,

> invest the interpretations of events in the domain of the representations with a reality that sometimes seems to eclipse the reality outside the skin of the ship. One navigator jokingly described his faith in the charted position by creating the following mock conversation over the chart: 'This little dot right here where these lines cross is where we are! I don't care if the bosun says we just went aground, we are here and there is plenty of water under the ship here.' For the navigator, the ship is where the lines of position intersect. (Hutchins, 1995, p. 115)

It is important to note, however, that the *possibility* of thinking about space in map-like terms does not mean that adults' representations of space are always map-like in nature. Several researchers (e.g. Kuipers, 1982; McNamara, 1986; Tversky, 1993, 1998; Shelton & McNamara, 1997) have consistently shown that not all spatial judgments require a map-like, abstract model of the spatial world. In particular, everyday navigation in familiar environments may well be accomplished by using representations that are organized around routes and landmarks. I do suggest, however, that the map-influenced model of the world gives adults the *possibility* of thinking about space in formal, map-like terms. The map-influenced model is useful in particular tasks, including those used by Piaget and others used to assess children's knowledge of the large-scale environment.

The influences of maps on spatial cognition and its development could operate in both a specific and a general way. At the more specific level, working with maps gives people direct experience in thinking about space in map-like ways. However, the effect need not be limited to situations in which actual maps are physically present. Over time, people may form mental models of large-scale space that are influenced by the accumulation of experience in working with maps. Maps and other symbolic representations leave their mark on cognitive processing, even after the actual symbols are removed (Miller & Paredes, 1996).

In sum, the way in which adults in modern Western societies conceive of large-scale spatial information may be affected by how this information is represented on formal maps. In the next section, I consider cross-cultural research that is consistent with this claim.

Cross-cultural and historical research on mental conceptions of large-scale space

An important implication of the present theoretical perspective is that there may be cross-cultural differences in how geographic-scale information is represented and used. My claim has been that prototypical maps influence the development of concepts of large-scale space. If this is true, then individuals who are reared in cultures that are less influenced by prototypical maps may have very different conceptions of large-scale space. Research in several non-Western cultures does suggest that the map-influenced model of geographical-scale space is not universal. Alternative conceptions of large-scale space that are less influenced by maps are often extremely sophisticated and support very impressive feats of navigation (Gell, 1983; Hutchins, 1995). Yet, they are distinctly different from modern Western (map-influenced) conceptions of large-scale space. Examining these differences highlights the effects of maps on Western conceptions of space. 'That this [i.e. modern, Western] representation of the world is a paper world, not an intuitive one, may be shown by contrasting this representation with others which are equally impressive but which are not based on maps and charts but on personal knowledge' (Olson, 1994, p. 213).

Perhaps the best documented example concerns the navigators of the Caroline Islands in the Pacific Ocean.

As Gladwin (1970), Lewis (1972) and Hutchins (1995) have noted, these individuals are able to navigate across hundreds of miles of open ocean without reference to maps or compasses. The native navigators' conceptions of space are fundamentally different from the Western, map- or chart-influenced conceptions. They imagine themselves to be part of the space. The position of the canoe is assumed to be fixed, and the navigators imagine that the stars and islands float past it. In contrast, modern, Western navigators conceive of space from above, as something that can be looked down upon and inspected, partly because of their experience in using maps (Hutchins, 1995).

A second well-documented example of sophisticated navigational abilities that are not based on the traditional Western map-like conceptions of space concerns groups of aborigines in Central and Western Australia. These individuals can travel across hundreds of miles of seemingly featureless desert, often without reference to compasses, maps or stars (Lewis, 1976; Chatwin, 1987; Berndt & Berndt 1994). In part, the navigation is accomplished by giving even small features of the desert symbolic meaning. Each navigator possesses an individual *songline*, a record of the individual's personal cosmology. Songlines connect locations in terms of myths regarding events, or *dreamings*, that took place during the creation of the particular features. One songline might include, for example, the story of the creation of a particular rock and the path that an ancient ancestor followed during the creation. In essence, the aborigines possess a very rich set of mnemonics that can provide a basis for recalling a route. When the individual goes on walkabout, he or she re-creates the record of the dreamings. The individual creation myths thus provide a rich set of landmarks that can facilitate navigation.

The aboriginal model of large-scale space is fundamentally different from the Western, map-influenced model. The aborigines do not conceive of their journeys in terms of fixed, metric distances. Rather, they remember and describe the journeys in terms of progress along a particular songline. As in the previous example regarding the Pacific navigators, the aborigines are able to navigate successfully without recourse to the Western, map-influenced model of large-scale space.

Finally, it should be noted that the Western, map-influenced model has evolved historically. Although maps have existed for perhaps as long as writing (Wilford, 1981), they were not readily available to the average person until they could be mass-produced (Brotton, 1998). In addition, abstract projections are a relatively new invention, and the map-influenced model of space may also be relatively new. Many early maps and navigational charts were essentially listings of possible sailing routes with detailed descriptions of shorelines (Wilford, 1981; Hutchins, 1995; Brotton, 1998; Crosby, 1998). In short, the historical existence of some types of maps for millennia does not imply that the modern, map-influenced model of space has existed for the same amount of time. Ong (1982) has summarized the changes that were brought about historically by adopting the map (gradually) as a model for large-scale space:

> Only after print and the extensive experience with maps that print implemented would human beings, when they thought about the cosmos or universe or 'world', think primarily of something laid out before their eyes, as in a modern printed atlas, a vast surface or assemblage of surfaces (vision presents surfaces) ready to be 'explored'. The ancient oral world knew few 'explorers', though it did know many itinerants, travelers, voyagers, adventurers, and pilgrims. (p. 73)

Although this section has focused on cross-cultural and historical differences in conceptual models of the world, there is also a general similarity among individuals from a wide variety of cultures. People in many different cultures *impose* on the external world a structure or mental model that leads to systematic organization and prediction. In modern Western societies, this model is engendered in part by exposure to maps. In non-Western societies, the model may be fundamentally different and is related to the cosmology of each culture. In both cases, however, people project a mental model onto the world and their experience of it (see Vosniadou & Brewer, 1992). Development may consist in part of the acquisition of the model that is relevant in a child's own culture.

The analysis presented thus far suggests that (a) adults in modern Western societies conceive of large-scale space partly in terms of a map-influenced model, but (b) this model is neither culturally nor historically universal. These two observations motivate the central claim of the next section: that the development of cognition of large-scale space in modern Western societies consists in part of the acquisition of a map-influenced model of the world.

Previous characterizations of the development of large-scale spatial cognition

In this section, I consider previous characterizations of the development of large-scale spatial cognition and point out how exposure to maps might influence this development. Many researchers have demonstrated that there are developmental differences in how children conceive of and mentally represent large-scale spatial information. For example, Piaget (Piaget & Inhelder, 1956; Piaget, Inhelder & Szeminska, 1960) suggested that young children's representations are based exclusively on topological relations and hence do not capture metric information such as angles and distance. Thus, on this view, the development of spatial cognition consists in part of the acquisition of new ways of mentally representing spatial information.

The results of more recent work on the development of children's mental representations of spatial information have led to an apparent paradox. On the one hand, the results of several studies suggest, contrary to Piaget's claims, that very young children can represent spatial information in sophisticated ways at a early age (Landau, 1986; Huttenlocher, Newcombe & Sandberg, 1994; Bushnell *et al.*, 1995; Newcombe & Huttenlocher, 2000). For example, Bushnell *et al.* (1995) found that 12-month-olds could find an object that was hidden under one of more than 50 irregularly shaped cushions in a large, circular hiding space. Curtains were placed around the border of the search space, and consequently there were no salient landmarks that the children could use as cues to the location of the object. These and similar results (e.g. Huttenlocher *et al.*, 1994) have been difficult to explain without claiming that the children have accurately encoded (metric) distance information that specifies the location in terms of a specific distance. Such a finding suggests that children do possess the ability to encode locations in metric terms much earlier than Piaget and others have claimed (Bremner, 1993).

Yet, despite these demonstrations of very early representation of metric information in young children, other researchers have observed substantial developmental differences in children's spatial cognition. In many cases, these developmental differences have been found in tasks that involve the cognition of large-scale space and multiple relations among multiple locations (Huttenlocher & Newcombe, 1984). It is these developmentally advanced skills, I will argue, that are affected in part by exposure to, and familiarity with, the model of space that children may gain from maps. Beginning in the late preschool years (approximately age 4 or 5), children undergo a shift in their conceptions and representations of large-scale space. Their knowledge of spatial relations becomes more abstract, and they start to think about space in terms of multiple relations among multiple locations. At approximately the same time, children begin to grasp how maps represent the large-scale world (Liben, 1999). Their understanding gradually develops throughout kindergarten and the early elementary school years. I argue here that the growing awareness of maps could be a partial cause of the observed changes in children's spatial cognition. This is not meant to imply, however, that simply looking at a map once or twice would be sufficient to improve performance. The map-mediated conception of space develops gradually. Simple exposure to a map or two is unlikely to be sufficient. Instead, the developmental process that is implied here is likely to be gradual and involve a conceptual change (e.g. Carey, 1985).

In the remainder of this section, I consider three characteristics of young children's spatial cognition that may be related to their lack of exposure to or understanding of maps: the lack of an abstract concept of space, difficulties in acquiring or using survey knowledge, and poor performance in tasks that require thinking about large-scale space.

Young children lack an absolute conception of space

Some theorists, most notably Piaget and colleagues, suggested that only older children and adults think about large-scale space in absolute terms (Piaget & Inhelder, 1956; Piaget *et al.*, 1960). On this view, young (preoperational) children's conceptions of space are fundamentally relative; they think about spatial locations in terms of how they have experienced the locations personally (Bremner, 1993).

Two examples from the work of Piaget *et al.* (1960) highlight differences in younger and older children's conceptions of large-scale space. The examples are derived from a task in which children were asked to make small-scale models of the locations of local landmarks; they were asked to arrange the objects to form a plan of their town. The descriptions of two younger children's placements highlight that the reconstructions are not constrained or influenced by an absolute conception of space:

> Miu (6; 10) puts a number of places together in what appears to be a pell-mell arrangement: the buildings and corridors of the main school, the main school playground, the kindergarten playground and entrance, the gymnasium which is a separate building, and a mound of sand on the bank of the Arve .... In short, a number of places are brought together by Miu's personal interest, while others show confusion between conceptual similarity and proximity in space. (p. 9)

> Gei (5; 11) puts the main school near the kindergarten with the two adjoining playgrounds in between, which is correct. However, while a nursery and various other buildings which have no particular interest for him

> find their right place by the main school building, places which fascinate him: the football ground and 'that house with the candy presents' (a general store which sells toys) are put near his own school – because they are near to his heart. (p. 9)

Piaget noted that, within a few years, children's placements of the objects were much more systematic and organized. In describing children of approximately ages 8–10, Piaget *et al.* noted that there was

> a marked difference between these reactions and those of subjects at lower levels. In the first place, a topographical schema is planned from the outset as a single whole and even where a subject deals with it in separate sub-groups or adds a section to his drawing, he wastes no time in making the different parts agree with each other...The end-result is always a coordinate whole ... . (p. 20)

Two aspects of the younger children's performance are noteworthy in terms of the potential influences of maps on spatial cognition. First, the reconstructions were not based on an abstract concept of space. Instead, the reconstructions reveal a relative concept of space, in which the children's personal experience is paramount. The younger children knew locations primarily in terms of what they had experienced directly or what was of special interest to them (e.g. the candy store). In this regard, the children's reconstructions are akin to the caricatures of the 'New Yorker's View of the World'. The New Yorker conflates interest and experience with actual (metric) distance: Chicago, China and California are all about the same distance from New York in this view of the world. All of these locations are relatively uninteresting to the New Yorker, and he or she simply treats them as 'away from New York'. Similarly, the young children's reconstructions conflated interest and experience with actual metric distance. The store that sold candy and toys, for example, was placed closer to the school than it should be. The children's constructions of the layout of the town should not be construed simply as reflecting distortions of otherwise abstract, metric representations. The term distortion implies a reference to another representation – a map-influenced representation. In this case, the young child's representation is fundamentally different from the map-derived model, and a comparison to this model is therefore misleading.

The second important characteristic of younger children's reconstructions is that they seem to lack knowledge of how the objects can be embedded within an overall framework. They do not seem to think that it is important to integrate the locations into a systematic plan of the town. The younger children's reconstructions of the town thus seem very different from traditional maps. In contrast, the older children's reconstructions do seem like traditional maps. The older children maintained what Piaget called the coordinate whole, and their placements of objects reflected knowledge of the relations among the locations.

The present theoretical analysis suggests that part of the explanation of the observed results is that the children lacked a map-influenced model of space. Piaget's tasks captured important developmental differences in how children think about spatial information, particularly information that cannot be perceived in a single glance. In essence, Piaget was measuring the ability of the child to think about large-scale space in map-like terms; children's reconstructions were said to reflect an absolute conception of space if the reconstructions resembled actual maps of the town. The 5- and 6-year-olds did not possess this model of the world, and hence their reconstructions did not resemble maps. Put simply, the perspective advanced here suggests that older children's reconstructions of the town look more like actual maps because they have become familiar with the way that maps typically represent spatial information. The younger children's reconstructions may not reflect a fundamental problem in representing or manipulating the relevant spatial information as much as lack of experience with the model of the world that the maps help to create.

A similar analysis can also be applied to children's comprehension of real, geographic-scale maps. Young children can understand some of the basic functions of maps, and they can comprehend some map-related concepts such as coordinate systems (Somerville & Bryant, 1985; Bryant & Somerville, 1986; Blades & Spencer, 1989). However, the early competence that young children demonstrate is not the end of the developmental story. For example, research on children's understanding of maps and aerial photographs of real, geographic-scale spaces has revealed that even elementary school children sometimes do not grasp fully the relation between a map or photograph and the represented space (Liben & Downs, 1992, 1993; Liben, 1993, 1999). For example, consider children's responses in tasks in which they are asked to identify features on maps. Typically, the researcher points to a particular feature, such a house or road, and asks the child to identify it. Children often are able to identify a substantial percentage of the depicted objects or locations (Liben & Downs, 1989, 1991; Blades & Spencer, 1994; Blades *et al.*, 1998). For example, children usually can recognize bodies of water and large buildings. However, Liben and Downs have documented several situations in which children do not appear to understand basic representational functions of maps. For example, one kindergartner said that a red line on a map could not represent a road because no roads are red in the world. Similarly, another kindergartner said that the representation of a road could not be a road because it was too narrow to accommodate a car. Likewise, another child correctly identified a large body of water (Lake Michigan on a small-scale map of Chicago), but then claimed to be able to see the lifeguard stand at the beach. Seeing the lifeguard stand would not be possible at the scale at which the map was drawn.

Children's errors are particularly interesting because they seem to indicate that they do not fully conceive of large-scale space in abstract terms (Liben & Downs, 1991; Liben, 1999). Children have not yet acquired a full understanding of how the maps stand for the represented spaces, and in addition they have not acquired the model of large-scale space that comes from using maps. Children often relate their personal experiences of items on maps to items that are usually found near the depicted items. For example, the child who claimed to be able to see a lifeguard stand on the small-scale map probably often encountered a lifeguard stand when he or she visited the beach. In interpreting map features, children recount their experiences of what is associated with the represented objects on the map. Thus there is a similarity between how children represent information that they acquire from direct experience navigating in the world and their naïve interpretation of geographic-scale maps. In both cases, young children interpret and recall features in terms of how they have experienced the information directly.

Young children have difficulty acquiring survey knowledge

The examples discussed thus far have involved children's construction or interpretation of map-like representations of the world. However, the map-derived model is not limited to how children interpret these external representations. Experience in using maps may also influence how children think about and mentally represent information gained from direct experience navigating in the world.

One relevant example concerns the acquisition of what is often called *survey knowledge*. Survey knowledge 'takes a static view from above the environment and locates landmarks with respect to each other' (Tversky, 1996, p. 479). An important characteristic of survey knowledge is what Levine, Jankovic and Palij (1982) have called *equiavailability*. This refers to knowledge of the multiple relations among multiple locations in a set of spatial objects, akin to the kind of knowledge represented in Figure 1.

Survey representations appear to be part of the repertoire of adults' representations of spatial information. For example, Taylor and Tversky (1992) found that adults could readily form survey-like representations of several hypothetical spaces based upon either route-like descriptions of the spaces or survey-like descriptions. Almost all adults performed well, including those who were read the route-like descriptions first and were not told in advance that they would need to draw a map of the space. Thus adults are capable of forming survey representations both from maps and from other sources of information, such as navigation and descriptions (see also Levine *et al.*, 1982; Landau, 1988; Taylor & Tversky, 1992; Tversky, 1996).

But the same cannot be said of young children. Although young children can navigate well through many spaces, they are less likely than older children and adults to form survey representations based on their direct experience of the spaces (Siegel & White, 1975; Hazen, Lockman & Pick, 1978; Cousins, Siegel & Maxwell, 1983; Herman, Shiraki & Miller, 1985; Anooshian & Kromer, 1986; Anooshian & Nelson, 1987). For example, Herman *et al.* (1985) investigated preschoolers' knowledge of the layout of their preschool. The children ranged from 3 to 5 years old. The researchers studied how children's mental representations changed as they learned more about the locations of objects both inside and outside the preschool. Several measures of children's survey knowledge of the space were taken across the school year. For example, in one task, the children were asked to point to locations that could not be seen from where they were currently standing. Most of the target locations were familiar or even well known to the children, and it was highly likely that the children would have experienced all of them in the course of their everyday navigation. However, the children had *not* experienced the *relations* among many of the locations directly. For example, the children might have traveled many times to both the water fountain and the playground but they had never been asked to think about the spatial relations between these two locations. The task of pointing to the unseen locations required that they transcend the knowledge that they had acquired through navigation in the preschool. Having a survey-like representation of the preschool probably would facilitate performance on this task. In general, children performed poorly on inference tasks. Other researchers have obtained similar findings in different kinds of environments (Hazen *et al.*, 1978; Cousins *et al.*, 1983).

How might maps influence the development of the ability to form and use survey representations of space, even when knowledge of the space is based on navigation? Several characteristics of survey representations bear important similarities to characteristics of maps; a 'survey [representation] is analogous to a map in many ways' (Tversky, 1996, p. 479). For example, both prototypical maps and survey representations imply a representation of space that is not directly tied to the way the information has been experienced during navigation. In addition, both maps and survey representations make accessible information about relations among multiple locations, and the relations among these locations. These similarities may indicate that maps help people to think about space in survey-like ways and to form survey representations. More specifically, experience with maps could facilitate the acquisition of survey representations in two ways. First, it could give children insight into overhead perspective that is necessary for forming a survey-like representation of the environment. Second, it could give them practice in thinking about space in terms of multiple relations among multiple locations. This insight and exposure might help children to realize that information gained from navigation can also be thought of in survey-like ways. In sum, children might gain from maps the understanding *that* and *how* information can be abstracted from direct experience and integrated into a survey-like representation.

Young children perform particularly poorly in large-scale spaces

A third important developmental difference that has often been discussed in the literature concerns the effects of the size of the space in which the tasks take place. There is often an interaction between age and the size of the space; younger children (i.e. preschoolers and kindergartners) perform much worse in relatively large spaces than in smaller spaces (Pick, 1976; Herman & Siegel, 1978; Acredolo, 1981; Pick & Lockman, 1981; Weatherford, 1982). Older children are generally less affected by the size of the space.

The work of Herman and Siegel (1978) provides a particularly relevant example of the effects of these differences. These researchers asked kindergartners, second-graders and fifth-graders to learn the locations of landmarks within a model town. The model town was placed either in a small classroom or in a large gymnasium. The children first walked through the model town several times. Then, the experimenters removed the landmarks and asked the children to place them back in the correct locations. Even kindergartners placed the objects close to the correct locations when the model town was placed in the classroom. However, in the larger gymnasium, the kindergartners performed poorly. These results are particularly interesting because they demonstrate a specific effect of the size of the space in which the task took place. The model town itself and the relative positions of the landmarks were identical in the two conditions. The only difference was the absolute size of the surrounding space.

The interaction between the effects of age and the size of the surrounding space has often been attributed to differences in the cues that are available for encoding individual locations in small-scale as opposed to large-scale spaces (Herman & Siegel, 1978; Acredolo, 1981; Huttenlocher & Newcombe, 1984). In a small-scale space, the children can rely on features of the surrounding room, such as corners and windows, to encode the locations of individual objects. In contrast, in a larger room, such as the gymnasium that Herman and Siegel used, the external framework of the room is less available. In this situation, mentally *imposing* an absolute framework or coordinate system on the locations may facilitate substantially children's memory for the locations. A mentally imposed frame of reference would allow the children to encode locations in terms of relations to other locations.

The abstract frame of reference, or a coordinate system, is precisely the kind of representational aid that may come in part from exposure to maps. Looking at maps may help children realize that there are ways of encoding spatial information that are not tied directly to specific landmarks. Both the abstract concept of space that maps entail, and the specific way in which they represent space, could help children to think about space in a way that would facilitate their performance in large-scale space.

In sum, many of the prior characterizations of the development of spatial cognition in the later preschool and early elementary school years seem to involve changes in how children think about large-scale space. Although very young children may be able to encode the locations of individual items in a small-scale space very accurately, this does not mean that they readily or spontaneously apply these skills to a large-scale space. Maps may play a critical role in helping children to think about the spaces beyond their immediate experience in absolute and survey-like terms.

Research evidence

In this section, I consider the results of specific studies that support the claim that exposure to maps, or to the kinds of information that maps can provide, may facilitate children's performance on spatial tasks. In part, the evidence consists of studies that have investigated the effects of exposure to the overhead view, regardless of the use of maps. Other studies have directly assessed the effects of using maps on children's acquisition of survey-like representations of a space.

Exposure to the overhead view

One way in which exposure to maps may facilitate the development of spatial cognition is by helping children to think about space from an oblique or overhead view. If this is true, then exposure to the overhead view might be expected to facilitate performance. The results of at least two studies are consistent with the claim that exposure to the overhead view of a spatial configuration can facilitate the development of young children's comprehension of the relations among locations.

In the first study, Rieser, Doxsey, McCarrell and Brooks (1982) investigated the effects of exposure to the overhead view on very young children's ability to navigate a detour in a simple maze. The children (ages 9–25 months) were asked to travel in the maze to find their mothers. Some of the infants were first raised to chest height to provide an overall, aerial view of the maze and of their mother. The children were then placed on the ground inside the maze. The dependent variable was whether they would go around a barrier to reach their mothers, who the infants could not see from the ground-level perspective. Exposure to the overhead view facilitated the 25-month-olds' performance. Of course, this study involved children far younger than the age at which children start to learn about maps. Nevertheless, it is noteworthy that simply exposing children to the overhead view facilitated their ability to make what is essentially a spatial inference.

Gauvain and Rogoff (1986) also demonstrated the relevance of the overhead view for acquiring survey-like representations of large-scale space. The children in this experiment were ages 6–9. They were asked to navigate a playhouse that consisted of several rooms. Half of the children were given instructions to learn the overall layout of the playhouse. These children were told to remember where all of the objects that they had seen were located. The remaining children were given route instructions; they were told to learn a specific route through the playhouse. In one room, the experimenters placed a slide from which all other rooms were visible. The older children who were asked to learn the overall layout of the playhouse spent significantly more time on top of the slide. In addition, they knew the relations among the various locations better than children who were given route instructions.

It should be noted that other research has failed to confirm that exposure to the overhead view facilitates children's performance. Liben and Yekel (1996) allowed 4- and 5-year-olds to study the layout of their classrooms from a booth located above the classroom. The children were then asked to locate stickers on maps of the classroom to indicate the location of different objects. Exposure to the overhead view did not facilitate performance. Although the cause of the differences between these results and those of Rieser *et al.* (1982) and Gauvain and Rogoff (1986) is not clear, one potentially important difference concerns the tasks that children were asked to perform. In the studies (Rieser *et al.*, 1982; Gauvain & Rogoff, 1986) that have found positive effects of exposure to the overhead view, the children were asked to use the information that they gained from the overhead view as a guide to navigation. In the Liben and Yekel study, children were asked to perform tasks on maps.

Consequences of exposure to maps

There is also evidence that exposure to maps can directly facilitate young children's acquisition of a survey-like representation of a large-scale space. For example, Uttal and Wellman (1989) demonstrated that exposure to a map could aid children's acquisition of a mental representation that included multiple relations among multiple locations. Children were asked to learn the layout of a playhouse that consisted of six contiguous rooms, as shown in Figure 2. The rooms were identified by the presence of a single toy animal in each room; thus the rooms were referred to as 'the dog's room', 'the cat's room' and so forth. Half of the children, the *map group*, learned the layouts from a simple map that showed the six rooms and the doors that connected the rooms. Small photographs were attached to the corresponding locations on the map. The children were required to memorize the map before entering the playhouse.

The remaining children, the *control group*, did not use a map. Instead, they learned *what* animals could be found in the playhouse from a series of flashcards. The experimenter presented the cards, in random order, and then asked the children to recall each of the animals. This procedure was repeated until the children had learned the list of the six animals that could be found in the playhouse. Thus, the control group knew what animals they could find in the playhouse but they did not learn the spatial locations of these animals before they entered the space.

After learning either the map or the flashcards, both groups of children were asked to navigate a circuitous route through the playhouse. Before they entered each room, the children were asked to anticipate what animal

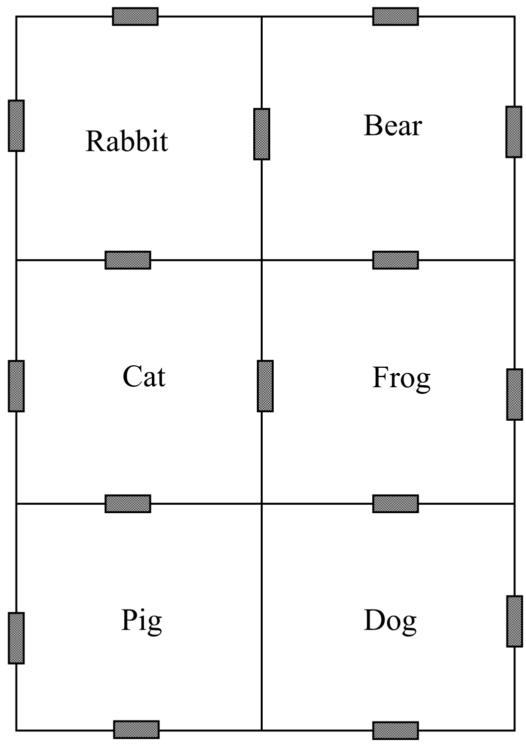


Figure 2 *The layout of the six locations used by Uttal and Wellman (1989). The shaded regions represent doors.*

would be found in the room. All of the doors to the different rooms were closed, and hence the only way for children to know which animal could be found in each room was to recall what they had learned before entering the maze – either the map that showed the locations of the animals or the flashcards that gave children only the identity of the animals. The researchers assessed children's knowledge of the layout of the playhouse by asking them to guess which animal would be found in each room. The key dependent variable was the number of animals that children could correctly identify.

Children who learned the map performed significantly better than those who learned the flashcards. Uttal and Wellman claimed that the only way children could have performed as well as they did was to form a survey-like representation of the represented space. The results of this study suggest that maps can help young children to think about space in survey-like ways. In sum, the available evidence is consistent with the claim that exposure to the unique characteristics of maps (e.g. the view from above and relations among locations) can lead to substantial improvements in performance and the acquisition of survey-like representations of the space.

Testing the theory

The results of previous studies are consistent with the notion that exposure to maps helps children to think about space in terms of integrated, survey-like representations. However, most of these studies have been conducted in relatively small spaces, and additional evidence will be needed to support the claim that exposure to maps could facilitate the development of concepts of large-scale space. In this final section I consider how the present general theoretical claims could be tested.

Providing support for the theory will require demonstrating that exposure to maps alters how children think about the spatial relations within an environment. It will not be sufficient to demonstrate that children can learn information from maps, or even that exposure to maps helps children know *more* about a given space than they knew before. Instead, if the present hypotheses are correct, then exposure to maps should lead to a gradual, qualitative change in children's reasoning about spatial locations. To provide evidence for such a change, I suggest that researchers *not* include all relevant locations on the maps that children see. If the present claims are correct, the exposure to maps should alter how children think about locations even if these specific locations are now shown on the map. My claim is not only that maps allow us to learn more about the world. I suggest also that maps alter how we think about spatial relations. If this is true, then the effect of exposure to a map should extend beyond simply knowing more about the particular locations that were shown on a map; the effect should extend as well to locations that were not shown on the map.

The proposed experiments would involve exposing children to maps of an area and then observing how this exposure influences their acquisition and thought about spatial information. The experiments will require a longitudinal design, but they will not require a great deal of time to complete. Many of the proposed experiments would follow a microgenetic approach, which involves investigating cognitive change intensively over a relatively short period of time (Siegler, 1996; Siegler & Stern, 1998). In the remainder of this section I discuss several possible venues for conducting these studies.

Familiar environments

One set of studies could be conducted in an environment with which children are reasonably familiar, such as their neighborhood. Their familiarity presumably would come largely from their direct experience navigating in the environment. Several lines of research have demonstrated developmental differences in how children of different ages represent locations within a familiar environment. Despite months or even years of exposure to the environment, younger children's mental representations are usually tied to what they have experienced directly and the routes that they have followed previously (e.g. Siegel & White, 1975; Cousins *et al.*, 1983; Herman *et al.*, 1985). Preschoolers have trouble making inferences about relations among landmarks that are not directly on routes that they have traveled. The guiding question of the proposed study would be whether exposure to a map would facilitate these kinds of judgments.

The researcher would begin by recruiting children who come from the same neighborhood because they are likely to share knowledge of the layout of this particular environment. The researcher would then assess children's knowledge of the neighborhood. These baseline assessment tasks could be similar to those used by Piaget, including a construction of maps or scale-models of the landscape. The child could also be asked to point to out-of-sight locations or to make detours in following routes. Other more sophisticated methods could also yield additional information. For example, children could be asked to rank-order the locations of landmarks in terms of distances (Kosslyn, Pick & Fariello, 1974; Newcombe & Liben, 1982). The data could then be converted to a multidimensional scaling solution to provide a map of each child's representation of the environment.

These initial measurements of children's knowledge of the environment would then be used as a baseline to assess the effects of exposing children to maps of the neighborhood. A critical question concerns the selection of landmarks to include on the map. One possibility would involve a map that showed only a few key features; several of the features that children had experienced in their daily travels would *not* be included on the map. This would allow the researcher to test whether exposure to the survey-like perspective of a map altered how children thought about the information that they had gained from direct experience. In other words, the researcher would examine not only whether the map helped children learn more about the space but also whether it helped them to think about the space in a different way.

In the map-familiarization procedure, the child would be asked to identify or describe the landmarks, following procedures similar to those of Blades *et al.* (1998) and Liben and Downs (1993). For example, the experimenter would point to the representation of the child's school and ask the child, 'Do you know what this is?' The experimenter would continue to point out landmarks on the map until the child could relate each to its referent.

After the map-familiarization phase, the experimenter and child would return to the neighborhood that was represented on the map. The experimenter would then repeat the initial assessment of the child's knowledge of the layout of the environment. The cycle could be repeated several times, probably over the course of a few weeks. If the present theoretical perspective is correct, then exposure to the maps should lead gradually to a substantial improvement in the ability to make inferences about or point to unseen locations that are not tied to specific routes that the children have experienced directly. This effect should be evident even for locations that were not shown on the map.

Unfamiliar environments

An analogous program of research could be conducted to assess how maps influence children's acquisition of information about environments with which they are initially unfamiliar. The present theoretical perspective leads to the prediction that exposure to a map should alter how children learn a new environment. Children who learn from a map, or from both direct experience and a map, should be better at inferring relations between locations. One way to approach this line of work would be to replicate the study of Herman *et al.* (1985), who investigated how children learned the layout of a preschool. However, in the proposed study, researchers would also include additional groups of children who are exposed to maps of the layout of the preschool. Specifically, one group would learn the layout by direct exposure, as in prior studies. The experimental group would augment this exposure with exposure to maps. A third, control group could be given additional exposure to particular routes, to control for the effects of added exposure that might come from the maps. Children who are exposed to the maps should be able to perform reasonably well on inference tasks that require reasoning about relations among locations that are not on the same route, whereas children who do not see the maps should perform poorly on these tasks. As suggested above, it will be useful *not* to include all of the relevant landmarks on the maps in order to demonstrate that exposure to the maps alters how children think about the spatial information.

This basic design could be applied to the learning of many other environments. Prior studies provide useful baselines regarding how children would be expected to perform without exposure to maps. The central question is whether exposure to maps helps children to augment what they learn from direct experience.

Virtual environments

The previous suggestions for testing the present theory all require assessing children's knowledge of a real environment. One potential limitation of these proposals is that it is difficult to control precisely how and when children are exposed to information about locations within an environment (although the general effects of exposure to maps should nevertheless be evident). Fortunately, emerging technologies may provide convenient and relatively inexpensive ways to address some of the research questions that have been proposed above. Virtual environments can now simulate convincingly the look and feel of real environments. In addition, they allow researchers to control precisely how and when participants see specific information (Loomis, Blascovich & Beall, 1999). Therefore researchers could test directly the claim that exposure to maps alters how children learn the layout of an environment. For example, children could be exposed to a top-down view of a space and then experimenters could observe children's virtual exploration of the space. These and other questions could be addressed more directly in a virtual environment than in a real environment.

The development of children's map-based distortions

Finally, researchers should also consider how children develop specific biases or false beliefs that stem from exposure to maps. I have mentioned that it is important to keep in mind that the map-influenced conception of the world is not an accurate copy (Kuipers, 1982; Liben, 1999). For example, many people have a map-derived misconception about the size of Greenland. Although these distortions are an inevitable consequence of our dependence on maps (Tversky, 1981, 1992), their existence also reinforces the central theme of this paper: we know the world beyond our immediate experience primarily through maps. Studying the development of false beliefs about specific geographic areas might reveal that children have begun to learn about the world from maps.

Conclusions

This paper has attempted to place the development of children's understanding of maps within a larger framework that includes research on the influences of several other symbol systems on cognitive development. Exposure to maps affects how people think about spatial information. This effect is analogous to that of written text on people's conception of language or mathematical symbols on their conception of number. The development of spatial cognition may consist in part of learning a new model of space that is tied to how information is represented on maps.

I have focused throughout on how maps might influence children's conceptions of space. However, my goal has not been to glorify maps. The changes that are brought about through learning about maps are not completely positive. As discussed above, maps create systematic biases. In addition, they reify political boundaries that may have relatively little to do with actual divisions on the ground. For example, when a person looks at a map of the former Yugoslavia, he or she can get the sense that the frontiers are rigid and stable. However, the political boundaries that are shown so clearly on maps may be much less important than the ethnic divisions that the maps do not show (see Wood, 1992). In sum, our conceptions of large-scale space are wedded to maps, for better and for worse. Perhaps the most interesting questions concern the development of this relationship.

Acknowledgements

Portions of this work were supported by Grant R29 HD 34929 from the National Institute for Child Health and Human Development and LIS Grant 97201313 from the National Science Foundation. Portions of the paper were completed while the author was a visiting researcher at the Max Planck Institute for Human Development, Berlin. I thank Dedre Gentner for her help in developing the ideas presented in this paper and for her comments on previous versions. Thanks also to Mark Blades, Gerald Clore, Judy DeLoache, Cathryn Freid, Scott Freundschuh, William Hanks, Catrin Rode, Elwin Savelsbergh, Elsbeth Stern and Linda Uttal for their comments and suggestions.

References

Acredolo, L.P. (1981). Small- and large-scale spatial concepts in infancy and childhood. In L.S. Liben, A.H. Patterson & N. Newcombe (Eds), *Spatial representation and behavior across the lifespan* (pp. 63–82). New York: Academic Press.

Allwein, G., & Barwise, J. (1996). *Logical reasoning with diagrams*. New York: Oxford University Press.

Anooshian, L.J., & Kromer, M.K. (1986). Children's spatial knowledge of their school campus. *Developmental Psychology*, **22**, 854–860.

Anooshian, L.J., & Nelson, S.K. (1987). Children's knowledge of directional relationships within their neighborhood. *Cognitive Development*, **2**, 113–126.

Bence, V.M., & Presson, C.C. (1997). The spatial basis of young children's use of scale models. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Washington, DC.

Berndt, R.M., & Berndt, C.H. (1994). *The speaking land: Myth and story in aboriginal Australia*. Rochester, VT: Inner Traditions International.

Bialystock, E. (1993). Metalinguistic awareness: the development of children's representations of language. In C. Pratt & A.F. Garton (Eds), *Systems of representation in children: Development and use* (pp. 211–233). Chichester: Wiley.

Bialystock, E. (1995). Making concepts of print symbolic: understanding how writing represents language. *First Language*, **15**, 317–338.

Bialystock, E., & Codd, J. (1996). Developing representations of quantity. *Canadian Journal of Behavioral Science*, **28**, 281–291.

Blades, M., & Cooke, Z. (1994). Young children's ability to understand a model as a spatial representation. *Journal of Genetic Psychology*, **155**, 201–218.

Blades, M., & Spencer, C. (1989). Young children's ability to use coordinate references. *Journal of Genetic Psychology*, **150**, 5–18.

Blades, M., & Spencer, C. (1994). The development of children's ability to use spatial representations. In H.W. Reese (Ed.), *Advances in child development and behavior* (Vol. 25, pp. 157–199). San Diego, CA: Academic Press.

Blades, M., Blaut, J., Davizeh, Z., Elguea, S., Sowden, S., Soni, D., Spencer, C., Stea, D., Surajpauli, R., & Uttal, D. (1998). A cross-cultural study of young children's mapping abilities. *Transactions of the Institute of British Geographers*, **23**, 269–277.

Blaut, J.M. (1991). Natural mapping. *Transactions of the Institute of British Geographers*, **16**, 55–74.

Bremner, J.G. (1993). Spatial representation in infancy and early childhood. In C. Pratt & A.F. Garton (Eds), *Systems of representation in children* (pp. 211–234). Chichester: Wiley.

Brotton, J. (1998). *Trading territories: Mapping the early modern world*. Ithaca, NY: Cornell University Press.

Bryant, P.E., & Somerville, S.C. (1986). The spatial demands of graphs. *British Journal of Developmental Psychology*, **77**, 187–197.

Bushnell, E.W., McKenzie, B.E., Lawrence, D.A., & Connell, S. (1995). The spatial coding strategies of one-year-old infants in a locomotor search task. *Child Development*, **66**, 937–958.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT/Bradford Books.

Cassirer, E. (1944). *An essay on man: An introduction to the philosophy of human culture*. New Haven, CT: Yale University Press.

Cassirer, E., Mannheim, R., & Hendel, C.W. (1965). *Philosophy of symbolic forms: The phenomenology of knowledge*. New Haven, CT: Yale University Press.

Chatwin, B. (1987). *The songlines*. New York: Viking.

Cousins, J.H., Siegel, A.W., & Maxwell, S.E. (1983). Way finding and cognitive mapping in large-scale environments: a test of a developmental model. *Journal of Experimental Child Psychology*, **35**, 1–20.

Crosby, A. (1998). *The quantification of reality: Quantification and Western society*. New York: Cambridge University Press.

DeLoache, J.S. (1987). Rapid change in the symbolic functioning of very young children. *Science*, **238**, 1556–1557.

DeLoache, J.S. (1989). Young children's understanding of the correspondence between a scale model and a large space. *Cognitive Development*, **4**, 121–139.

DeLoache, J.S. (1991). Symbolic functioning in very young children: understanding of pictures and models. *Child Development*, **62**, 736–752.

DeLoache, J.S. (1995). Early understanding and use of symbols: the model model. *Current Directions in Psychological Science*, **4**, 109–113.

Dent, B.D. (1996). *Cartography*. Dubuque, IA: William C. Brown.

Dow, G.A., & Pick, H.L. (1992). Young children's use of models and photographs as spatial representations. *Cognitive Development*, **7**, 351–363.

Downs, R.M. (1981). Maps and mappings as metaphors for spatial representation. In L.S. Liben, A.H. Patterson & N. Newcombe (Eds), *Spatial representation and behavior across the lifespan* (pp. 143–166). New York: Academic Press.

Downs, R.M., & Liben, L.S. (1993). Mediating the environment: communicating, appropriating, and developing graphic representations of space. In R.H. Wozniak & K.W. Fischer (Eds), *Development in context: Acting and thinking in specific environments* (pp. 155–181). Hillsdale, NJ: Erlbaum.

Eliot, J. (1987). *Models of psychological space: Psychometric, developmental, and experimental approaches*. New York: Springer.

Gattis, M., & Holyoak, K. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **22**, 231–239.

Gauvain, M. (1993). The development of spatial thinking in everyday activity. *Developmental Review*, **13**, 92–121.

Gauvain, M. (1995). Thinking in niches: sociocultural influences on cognitive development. *Human Development*, **38**, 25–45.

Gauvain, M., & Rogoff, B. (1986). Influence of the goal on children's exploration and memory of large-scale space. *Developmental Psychology*, **22**, 72–77.

Gell, A. (1983). How to read a map: remarks on the practical logic of navigation. *Man*, **20**, 271–286.

Gentner, D., & Ratterman, M.J. (1991). Language and the career of similarity. In S.A. Gelman & J.P. Byrnes (Eds), *Perspectives on language and thought: Interrelations in development* (pp. 225–277). New York: Cambridge University Press.

Gladwin, T. (1970). *East is a big bird: Navigation and logic on Puluwat atoll*. Cambridge, MA: Harvard University Press.

Hazen, N.L., Lockman, J.J., & Pick, H.L. (1978). The

development of children's representations of large-scale environments. *Child Development*, **49**, 623–636.
Herman, J.F., & Siegel, A.W. (1978). The development of cognitive mapping of the large-scale environment. *Journal of Experimental Child Psychology*, **26**, 389–406.
Herman, J.F., Shiraki, J.H., & Miller, B.S. (1985). Young children's ability to infer spatial relationships: evidence from a large, familiar environment. *Child Development*, **56**, 1195–1203.
Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
Huttenlocher, J.E., & Newcombe, N. (1984). The child's representation of information about location. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 81–112). Hillsdale, NJ: Erlbaum.
Huttenlocher, J.E., Newcombe, N., & Sandberg, E.H. (1994). The coding of spatial location in young children. *Cognitive Psychology*, **27**, 115–148.
Jammer, M. (1993). *Concepts of space: The history of theories in space in physics* (3rd edn). New York: Dover.
Karmiloff-Smith, A. (1979). Micro- and macrodevelopmental changes in language acquisition and other representational systems. *Cognitive Science*, **3**, 91–117.
Karmiloff-Smith, A. (1990). Constraints on representational change: evidence from children's drawing. *Cognition*, **34**, 57–83.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
Kosslyn, S.M., Pick, H.L., & Fariello, G.R. (1974). Cognitive maps in children and men. *Child Development*, **45**, 707–716.
Kuipers, B. (1982). The 'map-in-the-head' metaphor. *Environment and Behavior*, **14**, 202–220.
Landau, B. (1986). Early map use as an unlearned activity. *Cognition*, **22**, 201–223.
Landau, B. (1988). The construction and use of spatial knowledge in blind and sighted children. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi (Eds), *Spatial cognition: Brain bases and development* (pp. 343–371). Hillsdale, NJ: Erlbaum.
Larkin, J.H., & Simon, H.A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, **11**, 65–100.
Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. Cambridge: Cambridge University Press.
Lee, K., & Karmiloff-Smith, A. (1996). The development of external symbol systems: the child as a notator. In R. Gelman & T.K.-F. Au (Eds), *Perceptual and cognitive development: Handbook of perception and cognition* (2nd edn, pp. 185–211). San Diego, CA: Academic Press.
Levine, M.S., Jankovic, I.N., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General*, **111**, 157–175.
Lewis, D. (1972). *We the navigators: The ancient art of landfinding in the pacific*. Honolulu, HI: University of Hawaii Press.
Lewis, D. (1976). Observations on route finding and spatial orientation among the aboriginal peoples of the western desert region of central Australia. *Oceania*, **46**, 249–282.
Liben, L.S. (1993). Environmental cognition through direct and representational experiences: a life-span perspective. In T. Gaerling & G.W. Evans (Eds), *Environment, cognition, and action: An integrated approach* (pp. 245–276). New York: Oxford University Press.
Liben, L.S. (1999). Developing an understanding of external spatial representations. In I.E. Sigel (Ed.), *Development of mental representation: Theories and applications* (pp. 297–321). Mahwah, NJ: Erlbaum.
Liben, L.S. (in press). Thinking through maps. In M. Gattis (Ed.), *Spatial schemas in abstract thought*. Cambridge, MA: MIT Press.
Liben, L.S., & Downs, R. (1989). Understanding maps as symbols: the development of map concepts in children. In H.W. Reese (Ed.), *Advances in child development and behavior* (Vol. 22, pp. 146–201). New York: Academic Press.
Liben, L.S. & Downs, R.M. (1991). The role of graphic representations in understanding the world. In R.M. Downs, L.S. Liben & D.S. Palermo (Eds), *Visions of aesthetics, the environment, and development: The legacy of Joachim F. Wohlwill* (pp. 139–180). Hillsdale, NJ: Erlbaum.
Liben, L.S., & Downs, R.M. (1992). Developing an understanding of graphic representations in children and adults: the case of GEO-graphics. *Cognitive Development*, **7**, 331–349.
Liben, L.S., & Downs, R.M. (1993). Understanding person–space–map relations: Cartographic and developmental perspectives. *Developmental Psychology*, **29**, 739–752.
Liben, L.S., & Yekel, C. (1996). Preschoolers' understanding of plan and oblique maps: the role of geometric and representational correspondence. *Child Development*, **67**, 2780–2796.
Loomis, J.M., Blascovich, J.J., & Beall, A.C. (1999). Immersive virtual environments as a basic research tool in psychology. *Behavior Research Methods, Instruments, and Computers*, **31**, 557–564.
Luria, A. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.
McLuhan, M. (1964). *The Gutenbourg galaxy: The making of typographic man*. Toronto: University of Toronto Press.
McNamara, T.P. (1986). Hierarchies in cognitive maps. *Cognitive Psychology*, **18**, 87–121.
Miller, K.F. (1989). Measurement as a tool for thought: the role of measuring procedures in children's understanding of quantitative invariance. *Developmental Psychology*, **25**, 589–600.
Miller, K.F., & Paredes, D.R. (1996). On the shoulders of giants: cultural tools and mathematical development. In R.J. Sternberg & T. Ben-Zeev (Eds), *The nature of mathematical thinking* (pp. 83–117). Mahwah, NJ: Erlbaum.
Miller, K.F., & Stigler, J.W. (1991). Meanings of skill: effects of abacus expertise on number representation. *Cognition and Instruction*, **8**, 29–67.

Montello, D.R. (1993). Scale and multiple psychologies of space. In A.U. Frank &. I. Campari (Eds), *Spatial information theory: A theoretical basis for GIS* (pp. 312–321). Berlin: Springer.

Montello, D.R., & Golledge, R.G. (1999). *Scale and detail in the cognition of geographic information*. Report of Specialist Meeting of Project Varenius held 14–16 May 1998 in Santa Barbara, CA. University of California at Santa Barbara, *http://www.ncgia.org*.

Newcombe, N., & Huttenlocher, J. (2000). *Making space: Taking cognitive development one domain at a time*. Cambridge, MA: MIT Press.

Newcombe, N., & Liben, L.S. (1982). Barrier effects in the cognitive maps of children and adults. *Journal of Experimental Child Psychology*, **34**, 46–58.

Newcombe, N., Huttenlocher, J., & Learmonth, A. (in press). Spatial memory in 5-month-old infants. *Infant Behavior and Development*.

Olson, D.R. (1994). *The world on paper: The conceptual and cognitive implications of writing and reading*. Cambridge: Cambridge University Press.

Olson, D.R. (1996). Literate mentalities: literacy, consciousness of language, and modes of thought. In D.R. Olson & N. Torrance (Eds), *Modes of thought: Explorations in culture and cognition* (pp. 141–151). New York: Cambridge University Press.

Ong, W.J. (1982). *Orality and literacy: The technologizing of the word*. London: Methuen.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. New York: Norton.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry*. New York: Norton.

Pick, H.L., Jr (1976). Transactional-constructivist approach to environmental knowing: a commentary. In G.T. Moore & R.G. Golledge (Eds), *Environmental knowing* (pp. 185–188). Stroudsberg, PA: Dowden, Hutchinson & Ross.

Pick, H.L., & Lockman, J.J. (1981). From frames of reference to spatial representations. In L.S. Liben, A.H. Patterson & N. Newcombe (Eds), *Spatial representation and behavior across the lifespan* (pp. 39–62). New York: Academic Press.

Presson, C.C. (1982). The development of map-reading skills. *Child Development*, **53**, 196–199.

Presson, C.C., Hazelrigg, M.D., & DeLange, N. (1987). Orientation specificity in spatial memory: what makes a path different from a map of the path? *Journal of Experimental Psychology: Learning, Memory and Cognition*, **15**, 887–897.

Rieser, J.J., Doxsey, P.A., McCarrell, N.S., & Brooks, P.H. (1982). Wayfinding and toddlers' use of information from an aerial view of a maze. *Developmental Psychology*, **18**, 714–720.

Shelton, A.L., & McNamara, T.P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin and Review*, **4**, 102–106.

Shin, S.J. (1994). *The logical status of diagrams*. Cambridge: Cambridge University Press.

Siegel, A.W., & White, S.H. (1975). The development of spatial representations of large-scale environments. In H.W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10, pp. 9–55). New York: Academic Press.

Siegler, R.S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Siegler, R.S., & Stern, E. (1998). Conscious and unconscious discoveries: a microgenetic analysis. *Journal of Experimental Psychology: General*, **127**, 377–397.

Somerville, S.C., & Bryant, P.E. (1985). Young children's use of spatial coordinates. *Child Development*, **56**, 604–613.

Spencer, C., Blades, M., & Morsley, K. (1989). *The child in the physical environment: The development of spatial knowledge and cognition*. Chichester: Wiley.

Staub, F.C., & Stern, E. (1997). Abstract reasoning with mathematical constructs. *International Journal of Educational Research*, **27**, 63–75.

Stern, E. (1993). What makes certain arithmetic word problems involving comparison of sets so difficult for children? *Journal of Educational Psychology*, **85**, 7–23.

Stigler, J.W. (1984). 'Mental abacus': the effect of abacus training on Chinese children's mental calculation. *Cognitive Psychology*, **16**, 145–176.

Stigler, J.W., Chalip, L., & Miller, K.F. (1986). Consequences of skill: the case of abacus training in Taiwan. *American Journal of Education*, **94**, 447–479.

Taylor, H.A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, **31**, 261–292.

Thorndyke, P.W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, **14**, 560–589.

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology*, **13**, 407–433.

Tversky, B. (1992). Distortions in cognitive maps. *Geoforum*, **23**, 131–138.

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In A.U. Frank & I. Campari (Eds), *Spatial information theory: A theoretical basis for GIS* (pp. 14–24). Berlin: Springer.

Tversky, B. (1996). Spatial perspective in descriptions. In P. Bloom, M.A. Peterson, L. Nadel & M.F. Garrett (Eds), *Language, speech, and communication* (pp. 463–491). Cambridge, MA: MIT Press/Bradford Books.

Tversky, B. (1998). Three dimensions of spatial cognition. In M.A. Conway, S.E. Gathercole & C. Cornoldi (Eds), *Theories of memory II* (pp. 259–275). Hove, UK: Psychological Press.

Uttal, D.H. (1994).Children's and adults' reconstruction and scaling of spatial configurations. *British Journal of Developmental Psychology*, **12**, 259–275.

Uttal, D.H. (1996). Angles and distances: children's and adults' reconstruction and scaling of spatial configurations. *Child Development*, **67**, 2763–2779.

Uttal, D.H., & Gentner, D. (1995). Maps are like analogies. Paper presented at the annual meetings of the Association of American Geographers, Charlotte, NC.

Uttal, D.H., & Wellman, H.M. (1989). Children's mental representation of spatial information acquired from maps. *Developmental Psychology*, **25**, 128–138.

Vosniadou, S., & Brewer, W. (1992). Mental models of the earth: a study of conceptual change in childhood. *Cognitive Psychology*, **24**, 535–585.
Vygotsky, L.S. (1962). *Thought and language*. Cambridge, MA: MIT Press.
Vygotsky, L.S. (1978). *Mind in society*. Cambridge, MA: MIT Press.
Weatherford, D.L. (1982). Spatial cognition as a function of size and scale of the environment. *New Directions for Child Development*, **15**, 5–18.
Wellman, H.M., & Miller, K.F. (1986). Thinking about nothing: development of concepts of zero. *British Journal of Developmental Psychology*, **4**, 31–42.
Wilford, J.N. (1981). *The mapmakers*. New York: Knopf.
Wood, D. (1992). *The power of maps*. New York: Guilford Press.

Received: 26 August 1998
Accepted: 15 December 1999

Five concerns to address in geoscience education

Richard F. Yuretich
Department of Geosciences
University of Massachusetts
Amherst, MA 01003-9297

Geoscience is a latecomer to investigating how people learn in our discipline, in contrast to physics and chemistry, where the science of learning has been explored more fully. Perhaps this is because physics and chemistry were earlier faced with communicating very abstract concepts and laws that were not easily translated into words or pictures. Traditional geology has often been tarred and feathered as a “descriptive” subject, to distinguish it from the more “intellectually challenging” physical sciences, and most introductory and intermediate-level courses are very information-based. This has predetermined the teaching strategies that we have used, which have tended to rely on “show-and-tell” strategies that can work well for some topics. However, as our science deals with topics more involved than the identification of minerals, or the effects of glaciers upon the landscape, we need to ensure that our teaching strategies are up to the task. The complexities of climate modeling, mantle properties, and seismic interpretation demand that we investigate how students learn such concepts, so that our teaching strategies will be properly informed. Discussions around curriculum or courses are usually focused on the issue of content (what is to be included) rather than how we teach or why we choose that particular strategy. It is arguable that until recently the latter questions were not important, since learning any scientific subject seemed to involve the same procedures of reading, listening to lectures and perhaps reproducing experiments or investigations that formed the basis for the lectures and reading. However, as our knowledge base has grown, and as technology has altered the means that we can explore that knowledge, we need to step back and examine the whole process of learning geoscience.

The traditional image of science is usually given as “a white male in a lab coat performing experiments.” Geologists have (mercifully!) failed to fit this profile in almost all aspects, and this has been helpful in promoting geoscience to students who may not resonate with other science pursuits. The ability to leave the classroom behind for field trips and data-gathering expeditions can easily appeal to people who may not learn well in a lecture-lab format, and we should build upon this advantage. However, we really have yet to fully exploit this opportunity. When I first became acquainted with active learning and constructivism several years ago and I discussed these with geologist colleagues, the general reaction was: “Well, what’s so new about that? We have always taken our students into the field.” The act of leaving the classroom is conflated with the process of active learning itself, and perhaps this has stunted our interest in investigating why field trips can be so successful. We need to examine our current approaches to teaching geoscience and see how these address various ways of learning. Some aspects of the learning of physics or chemistry are applicable to geology as well, but other aspects – such as perception in three dimensions and visual representations such as maps or cross-sections – may require some different methods for learning effectively. I see five general areas that our discipline needs to explore more fully:

1. Action Research and Learning Goals. One of the things we need to promote is action research on the learning and teaching in our classes. In other words, we need to raise the practice of teaching to a research project in its own right. Most of us, even at research universities, spend more time teaching than in pursuit of our individual investigations. In many cases, we have extensive documentation of student outcomes, our activities in the classroom, and summative evaluations that could form the basis of ongoing research in to the effectiveness of our instructional methods. Initiating appropriate professional development workshops around these issues can stimulate interest in a careful self-analysis of the methods we currently use and the effects they are having on our students. In a related effort, we need to have a realistic set of learning goals for students in the geosciences that are more than content goals. Individual instructors and some departments have dabbled in this endeavor, but there are more similarities than differences among the programs in various colleges and universities, so a set of goals could be established for the discipline as a whole.

2. Problem-Based Learning. Problem-based or case-study learning has become the norm for many medical schools and is now being used in some undergraduate biology courses. Since the geosciences have many practical aspects, several of which involving diagnosing and solving complex problems, a case-study approach may be a very effective learning tool. However, we have only a few models that have been tried. I can envision a library of case-based investigations centering on petroleum or mineral exploration, environmental contamination or remediation, and climate systems, that could be incorporated into undergraduate curricula. What research is currently available that shows the impact of case-study learning in those subject areas where it is widely practiced? Are these results transferable to geosciences? In what ways? An expanded effort in developing case-based learning in geosciences and in evaluating its impact upon student performance and development is an important goal.

3. The Role of Field Programs. Field experiences are a hallmark of geoscience instruction, and many programs require some field training for a degree. How effective are field trips and field courses in promoting student learning? My personal experiences are that many field trips are little more than lectures at the outcrop, despite the obvious availability of materials for on-site active learning. Extended field courses, whether during the summer or during the academic year, are usually more reliant on student initiative and discovery. What are the learning goals of field courses or programs that are distinct from traditional methods? Are these goals being achieved? What would it take to make field experiences a more successful learning strategy?

4. Using Technology Wisely. Technology can be both a blessing and a curse. Computer simulations can illustrate processes or concepts that otherwise must remain in our imaginations, and the analysis of complex sets of data can be streamlined to occur in a time frame suitable for classroom instruction. Computer networks provide opportunities for interactive homework that encourages active learning. Many institutions now support flexible web platforms, such as WebCT, e-college or Blackboard, which can simplify the task of incorporating web-based instruction into a course. What data are available on how these are being used? Classroom communication systems, such as Classtalk or CPS, can augment discussion and gauge students' understanding of concepts in real time. We have little first-hand experiences with these technologies in the geosciences, and we need to see how they are being used in other subjects. On the other hand, presentation software can limit spontaneity in the classroom, and may reinforce passive listening if used in a television-like entertainment mode. What are the most effective ways that technology can be used to enhance learning in the geosciences?

5. Assessment. Assessment is an issue that we have hardly examined, and it is arguably the most important of all. How can we be sure that the methods we are using are having the desired effects? Are the goals that we have established for our courses being met? Are the students learning at the levels that we want or expect? Many of us are wedded to the traditional exam as our principal assessment tool. How can we design exams that evaluate higher order thinking? Many other techniques have been put forward to assist with formative assessment: minute papers, portfolio assessments, longer projects etc. Are these effective methods? Are there reliable data on their use and proper application? We need to more fully integrate assessment into all the aspects of our teaching, so that we can be aware of the overall success of the methods we are using.

In summary, we are at, or even somewhat past, the point to evaluate some strongly held beliefs in geoscience education. The research that is being done on the nature of learning, on the various ways that different people learn most effectively, and on how individuals construct their understanding of science from preconceptions, can help us revitalize the teaching of our favorite subject. As good scientists, we should welcome the opportunity to turn our teaching into a part of our research program.