Assessing Student Learning in the Data Sets and Inquiry in Geoscience Education Project (DIGS)

Daniel Zalles
SRI International

Edys Quellmalz
WestEd

Janice Gobert and Amy Pallant
The Concord Consortium

Introduction

Data Sets for Inquiry in Geoscience (DIGS) is a two-year National Science Foundation-funded project (NSF GEO #0507828) that is developing modules to supplement existing geoscience curriculum. The modules consist of weeklong curriculum units and 1-2-day performance assessments on the commonly-taught secondary-level geoscience topics of plate boundaries and climate. The modules provide extended inquiry-based investigations employing real geoscience data sets and visualizations, as well as performance assessments that provide evidence of geoscience knowledge and inquiry strategies seldom captured in traditional test formats. These units and assessments yield evidence of students' abilities to demonstrate greater understanding of the conventions and constraints of inquiry about geoscience phenomena and provide models of how the interpretation and analysis of geoscientific data sets can be scaffolded through age-appropriate tasks that facilitate high-quality student inquiry. The goals of the project are to:

- Study the impacts on student learning of Web-based supplementary curriculum modules that engage secondary-level students in projects in which students use real data sets, visualizations, and software tools to conduct investigations within two fundamental topics of study, climate change and plate boundaries.
- Develop design principles, specification shells, and prototypes of technology-based performance assessments to provide evidence of
    - students' geoscientific knowledge and inquiry skills (including data literacy skills);
    - students' ability to access, use, analyze, and interpret technology-based geoscience data sets.
- Develop scenarios based on the specification shells that describe curriculum modules and performance assessments that can be developed for other geoscience standards and curriculum programs.

Background

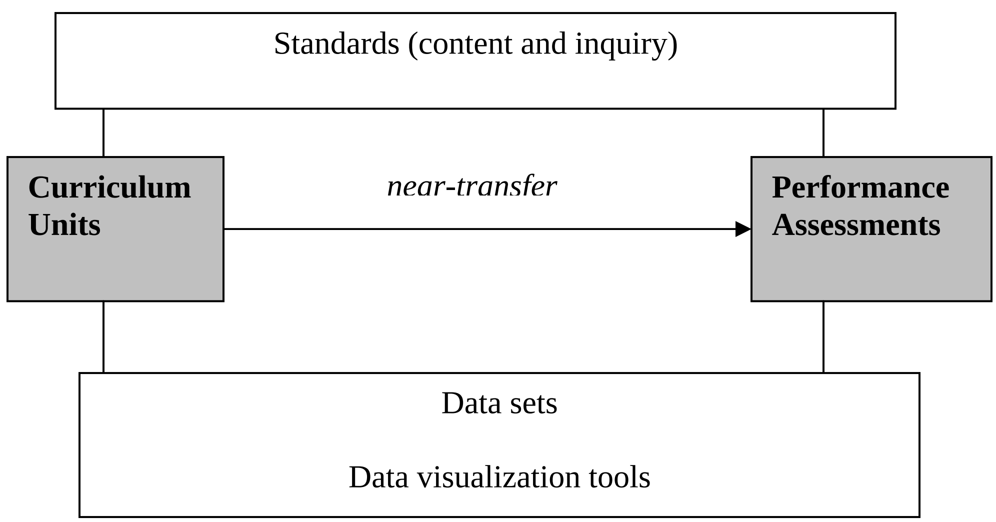
The report "Bringing Research on Learning to the Geosciences" (Manduca, Mogk, & Stillings, 2002) recommended a new program of research to invigorate and expand geoscience education. The report recommended integration of best practices in learning science with the distinctive challenges posed by using geoscience data sets and visualizations in inquiry activities (e.g., working with geologic time-referenced concepts, observing complex natural systems, using integrative and synthetic approaches).

The topic of student learning in the geosciences at the middle and secondary levels has not been well studied, particularly when compared with students' learning and conceptions in the physical sciences (Stofflet, 1994). Furthermore, previous studies have largely focused on student understanding of geoscience concepts and have not benefited from research on the teaching of inquiry in the specific geoscience disciplines that have been the focus of the project (Ault, 1994; Muthukrishna et al., 1993; Ross & Shuell, 1993; Dove, 1998; Gardner et al., 1992).Therefore, the DIGS project set forth to develop supplementary modules composed of curriculum units and performance assessments that would serve as models of resources that engage students in conducting inquiry with tools used by geoscientists.

Module Development

Figure 1 displays the structure of the DIGS modules. Students complete 4-5 day supplementary curriculum units on important geoscience topics. In the process, they examine authentic, publicly-available data sets with the help of appropriate software tools that permit students to select, simulate, and represent the data in different ways. The performance assessments present tasks that require that students transfer the inquiry skills practiced in the units to new, yet conceptually-related problems. The assessment results provide data on the students' interactions with and manipulation of the visualizations and data sets which can, in turn, be used to document achievement of inquiry skills.

Figure 1. DIGS Module Design



Design principles. The National Science Standards (National Research Council, 1996) lay out the broad conceptual knowledge and inquiry skills that define scientific literacy and provide the foundation for the module designs. The unit designs are characterized by problem-based learning tasks (Evenson & Hmelo, 2000; Hmelo-Silver, 2004). Problems are posed that require the investigation of data and do not have a single correct answer. The students construct and explain their thinking as they respond to the problems. The task structures permit measurement of distinct components of inquiry, including stating research questions, posing hypotheses, planning and conducting investigations, gathering evidence, analyzing data, considering disconfirming evidence, and communicating explanations. The performance assessments use principles of evidence-centered design, in which there is a tight coupling of knowledge and skills that the students should be able to demonstrate (the student model), characteristics of tasks that clearly reveal those skills and understandings (the task model), and the characteristics of student responses to the tasks that provide evidence that students have used the requisite knowledge and skills (the evidence model) (Messick, 1994; Mislevy et al., 2003; Pellegrino et al., 2001). Performance assessments are particularly well suited to measuring students' conceptual understandings and abilities to conduct and communicate investigations of significant, recurring problems (Baxter & Glaser, 1998; Bransford et al., 2000; Pellegrino et al., 2001).

Development of the modules. The development process began by identifying which topics should be the focus of the modules. National and state (California and Massachusetts) standards on the topics were examined, then ranges of possible unit objectives and activities, interactive technologies, publicly available data sets, and content-appropriate data visualization formats (e.g., spreadsheets and graphs, GIS, remotely-sensed images, 3-D rendering, simulations) were identified that might be used. The topics, plate boundaries and climate change, were chosen primarily because (1) the topics are widely taught in upper-level middle and/or secondary-level science curricula and (2) they provide contrasting cases of how inquiry methods are applied in different geoscience disciplines.

Analyses were then conducted to examine how the topics are addressed in typical secondary science curricula. Plate boundaries and earthquakes were found to receive fairly standard treatment in textbooks (e.g., types of plate boundaries, p and s curves, earthquake measurement practices, earthquake effects). More variability was found in how climate and climate change is treated, yet examples were found of the major concepts addressed in the climate module design (e.g., urban heat island effects, differences between weather and climate, anthropogenic influences on the composition of the atmosphere, Greenhouse Effect, feedback loops, and air temperature-monitoring).

As the topics were selected and the student tasks specified, data sets, visualizations, and software tools were identified that would be appropriate for use by secondary students. Our selection criteria were that (1) the data sets needed to be large enough to permit the investigation of change patterns in the phenomena and (2) the software needed to be flexible enough to permit student choice about what data to examine, yet have a sufficiently simple interface for students to use independently in the short time frame of the module, without extensive tutorial support. Our development criteria were to (1) build into the modules the appropriate levels of cognitive demand for the science knowledge, inquiry tasks, and technology use; (2) build upon the science knowledge addressed in the standard curricula and not introduce complex, new content, and (3) provide sufficient flexibility in the curriculum tasks to accommodate varying teacher instructional approaches.

Drafts of specification shells were developed to specify the modules' evolving designs. The shells outlined major unit and assessment activities, their parallel sequences, and their alignments with national science inquiry and content standards. The specification shells served as a reference for the design principles shaping the development of the curriculum and assessment tasks.

Technology tools. A variety of software programs for data visualization and analysis were examined. For the plate boundaries unit, a three-dimensional simulation tool called Seismic Eruption was found.[1] Freely available for downloading from the Web[2], it permits students to compare and contrast the frequencies and characteristics of real earthquakes along different types of plate boundaries around the world. For the Climate Unit, Microsoft Excel was selected as the tool with which students would create graphs to investigate air temperature change trends. Excel was chosen because it is the most ubiquitous tool in schools that permits the graphing of data sets for the trend analyses prompted in the module. The MyWorld™ geographic information system was chosen to display visualizations of specific geospatial distributions of temperature and carbon emission data sets.[3]

Technology platform for module administration. The project team considered the most cost-effective and logistically practical methods for classroom use of the tools. We decided to create separate web sites for each unit and assessment, in addition to teacher-only sites that introduced the module and presented the specification shells and alignment tables. Each web site has links to the downloadable data sets and student materials, including related readings, task directions, and response sheets. Upon completion of the project, the web site addresses will be made publicly available.

Climate module description. In the climate module, The Heat Is On: Understanding Local Climate Change, students draw conclusions about the extent to which multiple decades of temperature data about Phoenix suggest that a shift in climate is taking place there as opposed to exhibiting nothing more than natural variability. The data are from the Global Climate Historical Network data base[4]. Students also compare the changing trends in Phoenix to larger geographically-distributed temperature trends, then investigate if there is evidence of a relationship between the temperature data and data that would suggest anthropogenic influences. Students think critically about what can and cannot be known from the available data and propose a more ideal research study. The assessment requires that students apply the methods and findings from the investigation of the climate data for Phoenix to climate data for Chicago. In contrast to the curriculum unit, which primarily uses constructed-response tasks to encourage student reflection and discussion, the climate assessment tasks pose explicit selected- and constructed-response questions. For example, in the unit students are asked to construct an interpretation or conclusion, whereas in the assessment, an item may present a set of choices which students then justify.

Figure 2 displays examples of graphs of air temperature data that students examine for trends. Figure 3 shows examples of GIS images that students critically examine for

---

[1] Seismic Eruption. Version 2.1. Level 2006.05. © Alan Jones, 1996-2006
[2] http://www.geol.binghamton.edu/faculty/jones/#Seismic-Eruptions
[3] My World™, Version 4.02. Copyright © 2000-2006. Northwestern University. All Rights Reserved
[4] http://www.ncdc.noaa.gov/oa/climate/ghcn-monthly/index.php

evidence of relationships (converse or inverse) between geographic distributions of anthropogenic carbon emissions and 30-year mean temperature differences.

Figure 2. Excel graphs of air temperature data from Phoenix and Chicago

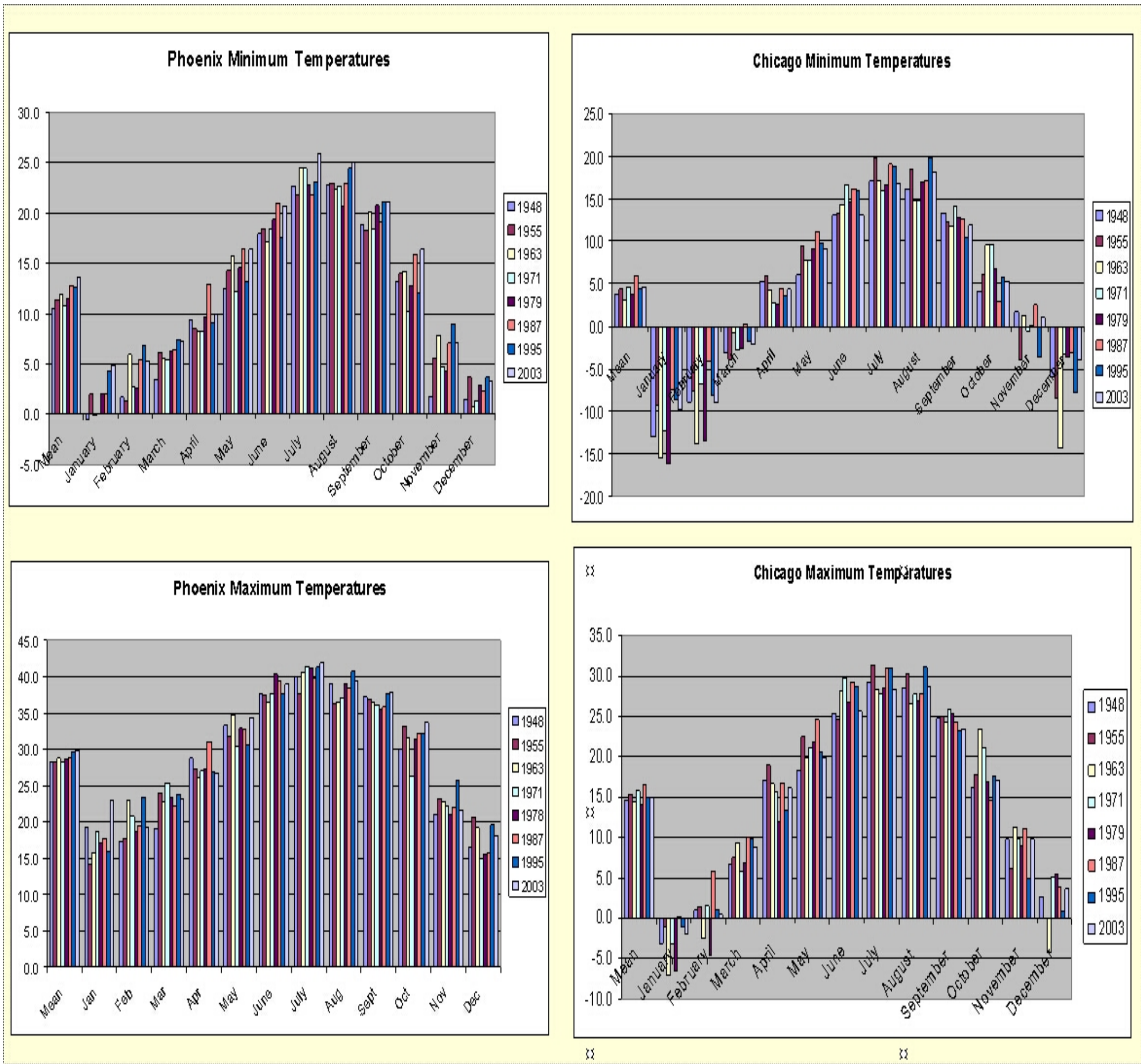
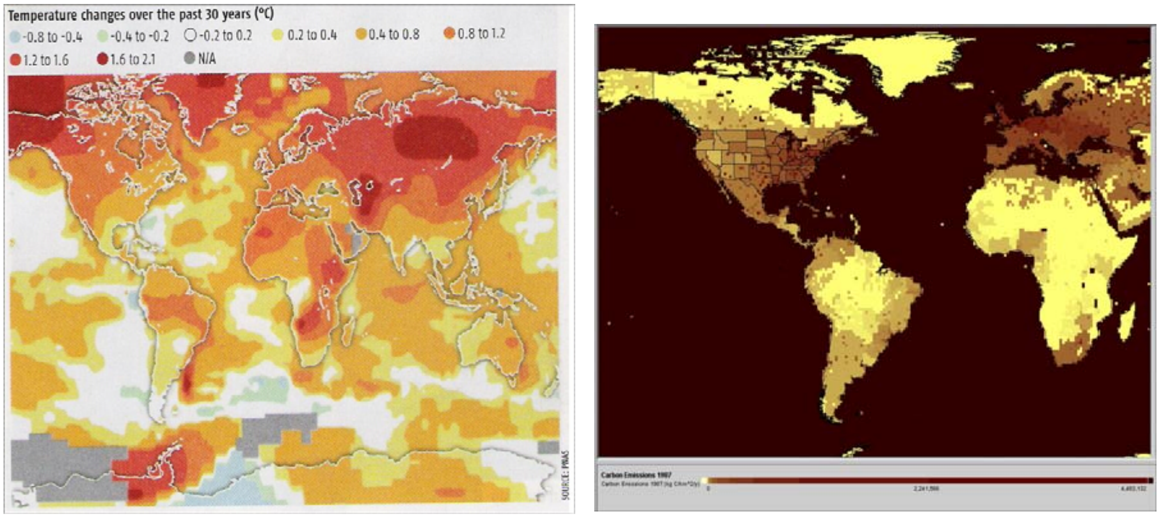


Figure 3. Global distributions of carbon emissions and temperature changes



Challenges in the development of the climate unit included:

- determining the appropriate number of data sets and visualizations to prompt the intended inquiry in the short unit and assessment time frame
- representing the science and scientific uncertainty accurately
- determining the appropriate level of technology use
- determining the appropriate level of scaffolding to help students synthesize observations from different data sets that have different ranges, relationships, and characteristics (e.g., nominal vs. ordinal)
- promoting the expression of levels of scientific uncertainty that befit the limitations of the data at the students' disposal (e.g., through self-ratings of confidence in conclusions and proposals of alternative hypotheses).

Plate boundaries module description. The plate boundaries module, On Shaky Ground: Understanding Earthquake Activity Along Plate Boundaries, engages students in use of a time-based simulation to explore earthquakes' relationship to the characteristics of plate boundaries in the Earth's crust. The tool, Seismic Eruption, simulates multiple decades of three-dimensional data about earthquakes around the world.

The unit is designed to take approximately four days and the assessment one day. The unit's components are designed to elicit the scientific inquiry abilities identified in national science standards. The students:

- hypothesize about earthquake likelihood at locations around the world
- observe earthquake patterns along divergent, convergent, and transform boundaries
- collect data and compare earthquake depth, magnitude, frequency, and location along the different plate types (convergent, divergent, transform) of plate boundaries
- analyze earthquake data sets from United States Geologic Survey database along different boundaries in data tables and in map representations
- develop visualizations of plate boundaries (create cross-sections using the Seismic

eruption tool, draw cross-sections, etc.)

- relate interactions of the plates to the emergent pattern of earthquakes.

In the assessment, the students run and analyze historical simulations of parallel earthquake data sets but on a type of plate boundary different from the one investigated in the unit.

Figure 4 displays a two-dimensional overhead view of earthquake activity between 1960 and the present in the Seismic Eruption tool, in relation to plate boundaries. Figure 5 displays a cross-sectional view of earthquake activity between 1960 and 2007 at the Mid-Atlantic ridge location specified in Figure 3, plus the key for interpreting the symbology.

Figure 4. Plate boundaries and simulated earthquake activity

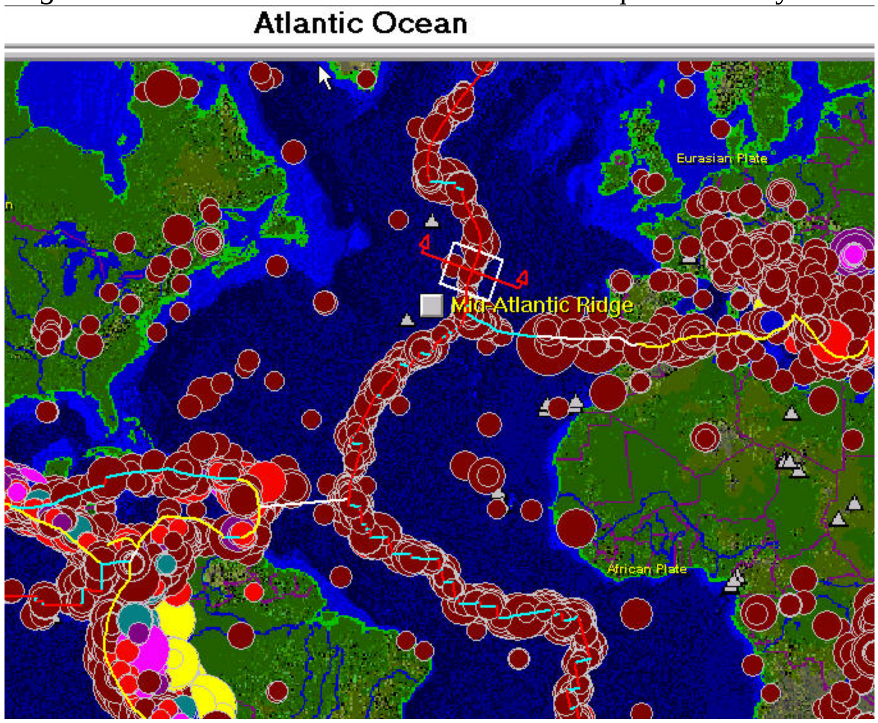
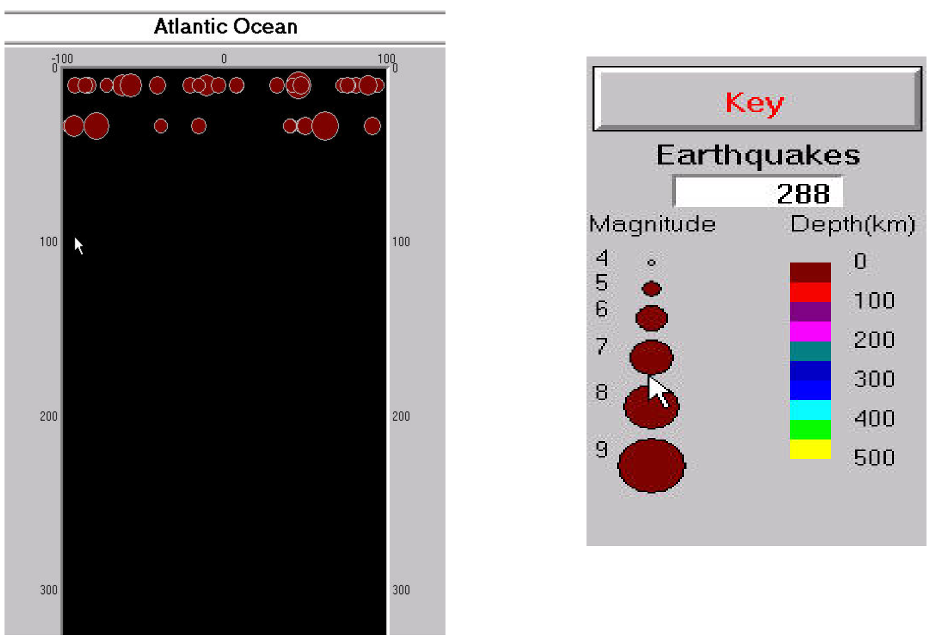


Figure 5. Cross-sectional view of earthquake simulation



Challenges in the development of the plate boundaries module included:

- incorporating into the student materials the appropriate amount of scaffolding for running the seismic eruption simulations;
- designing multiple inquiry opportunities beyond data analysis.

Technical Quality

The data being collected to document the quality of the DIGS modules includes approaches for judging curriculum quality and methods for documenting tests' technical quality recommended by research and test development standards: alignment of the assessments with national standards for science, task specifications aligned with standards, analyses of teacher and student data gathered from classroom pilot testing, and cognitive analyses of students thinking-aloud, (AERA/APA/NCME, 1999; Pellegrino, et. al, 2002; Quellmalz, et. al, 2005). Advisory panel review, external evaluator review, and sets of feasibility tests and pilot tests have been conducted to establish the technical quality of the materials. The tasks for ensuring technical quality proceeded as follows:

1. Advisors' review of initial specification shells identifying the alignment of tasks and questions with standards
2. Feasibility testing
3. Pilot testing – Round 1
4. Advisors' review of student materials
5. Pilot testing – Round 2

Advisory panel reviews. During the development process, a panel of advisors met to review module materials and alignments to standards. The panel included experts in geoscience curriculum development, assessment, and data literacy education, plus professional scientists in the content areas. Advisors also reviewed drafts of the modules between the two rounds of pilot testing.

Feasibility testing. Early feasibility tests with small groups of students were carried out to determine (1) the extent to which the tasks and questions were clear, (2) that the tasks and questions were eliciting the intended knowledge and inquiry skills, and (3) that the tasks were appropriate for the intended grade levels. Five students were tested for the climate module and four were tested for the plate boundaries module. Two of the five students trying out the climate unit worked in a pair and were observed discussing how to respond to the prompts. Other students worked alone. The students were observed responding to the prompts and were debriefed about them in interviews, task by task. The pair of students working through the climate unit were observed discussing how to respond.

Pilot testing. Once we revised the modules in response to student feasibility testing, each module was pilot tested. Round 1 of the climate module pilot test was conducted in four 11th and 12th grade environmental science classes, in October 2006, in a California Bay Area public high school. The 99 students who participated in the pilot test completed the unit in five days and the assessment in two days. The second round of pilot testing will take place in May 2007 at a different Bay Area high school.

The plate boundaries module was first pilot tested in two 9th grade classes in a public high school in suburb of Boston, Massachusetts. A second round of pilot testing was conducted in 15 classes of 8th grade students in a district near Boston in which plate boundaries is taught in 8th grade rather than in the more typical 9th grade.

In the pilot test classrooms, individual students identified by their teachers as being average achievers were observed "thinking aloud" as they responded to the assessment prompts. The think-aloud transcripts permit analysis of how well the prompts elicited the intended inquiry skills and content knowledge (Quellmalz and Haydel, 2003) and provide partial evidence of the content and construct validity of the items. Prior to pilot tests, the teachers reviewed the student unit materials and were also observed thinking aloud about the assessment prompts.

Preliminary results from the climate unit testing indicated that students were especially engaged by the opportunities provided them in the modules to make choices about what data to examine. Student challenges noted by the researchers included:

- applying knowledge of different effects of different emissions to data analyses
- differentiating between the concepts of carbon emission and carbon accumulation
- understanding how daily minimum and maximum monthly temperature readings carry different significances for understanding local climate trends
- arguing conclusions based on scientific evidence
- recognizing the importance of collecting counterfactual data when evaluating outcomes of interventions.

Preliminary results from the plate boundaries unit testing indicates that students were able to relate plate motion to the patterns of earthquakes along the different types of boundaries and were able to analyze data represented both numerically and using the visualization tools. Students were able to look at the aggregate of the data sets and reason about likelihood of earthquake occurrences around the world. Student challenges noted by the researchers included:

- understanding scale in visual representation and comparisons
- interpreting the cross-sectional representation as it related to the map view representation
- defending conclusions about earthquake patterns along plate boundaries as a result of limited data collection.

Next steps

Rubrics are being developed to score both the content knowledge of the students as well as the inquiry skills used. Illustrative examples of student work will be made available for subsequent scoring training in other classrooms. In the final phase of the project, specification shells for other scenarios will be developed that describe additional modules that could be developed on other commonly taught geoscience topics using the DIGS design principles.

References

Ault, Jr., C. R. (1984). The everyday perspective and exceedingly unobvious meaning. Journal of Geological Education, 32, 89-91.

Baxter, G. P., & Glaser, R. (1998). Investigating the cognitive complexity of science assessments. *Educational Measurement: Issues and Practice, 17*, 37-45.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Brody, M. J. (1993, April). *Student understanding of water and water resources: A review of the literature*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Dove, J. (1998). Alternative conceptions about weather. *School Science Review, 79*(289), 65-69.

Evenson, Dorothy H , Hmelo, Cindy E., (Eds.). (2000). *Problem-Based Learning: A Research Perspective on Learning Interactions*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16, 235-266.

Gardner, C. M., Simmons, P. E., & Simpson, R. E. (1992). The effects of CAI and hands-on activities on elementary students' attitudes and weather knowledge. *School Science and Mathematics, 92*(6), 334-336.

Manduca, C., Mogk, D., & Stillings, N. (2002). Bringing research on learning to the geosciences. Report from a workshop sponsored by the National Science Foundation and the Johnson Foundation. Retrieved 4/4/2007 from http://serc.carleton.edu/files/research_on_learning/ROL0304_2004.pdf

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 32*, 13-23.

Mislevy, R. J., Chudowsky, N., Draney, K., Fried, R., Gaffney, T., Haertel, G., Hafter, A., Hamel, L., Kennedy, C., Long, K., Morrison, A. L., Murphy, R., Pena, P., Quellmalz, E., Rosenquist, A., Songer, N., Schank, P., Wenk, A., & Wilson, M. (2003). *Design patterns for assessing science inquiry (PADI Technical Report 1)*. Menlo Park, CA: SRI International, Center for Technology in Learning.

Muthukrishna, N., Carnine, D., Gressen, B., & Miller, S. (1993). Children's alternative frameworks: Should they be directly addressed in science instruction? *Journal of Research in Science Teaching, 30*(3), 233-248.

National Research Council (NRC). (1996). *National science education standards*. Washington, DC: National Academy Press.

Pellegrino, J., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Quellmalz, E. S., & Haydel, A. M. (2003, April). *Using cognitive analyses to describe students' science inquiry and motivation to learn*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Ross, K., & Shuell, T. (1993). Children's beliefs about earthquakes. *Science Education*, *77*, 191-205.

Stofflett, R. (1994). Conceptual change in elementary school teacher candidate knowledge of rock-cycle processes. *Journal of Geological Education, 42,* 494-500.