Less talk, more action: Active Learning in Introductory Geoscience courses.

David McConnell, North Carolina State University

Introduction

Most students in STEM disciplines have expressed concern about poor teaching specifically mentioning, dull courses, disengaged instructors, and unsupportive teaching strategies (Seymour and Hewitt, 1997; see also Tobias, 1990; Strenta and others, 1994). Seymour and Hewitt (1997) suggested that a thorough revision of teaching and learning in first year STEM courses would likely improve student retention rates (Seymour and Hewitt, 1997). This view was echoed by the National Science Board (2003, p. 20), which stated that greater retention in STEM disciplines

> *…will require modification of the educational environment, particularly better teaching and advising . . . More institutional resources must be directed to improving the quality of teaching, the nature of introductory classes, the design of facilities to support new teaching methods . . .*

Of the three directions highlighted by the National Science Board, design or redesign of facilities can be effective in improving learning (Beichner and others, 2007) but requires substantial institutional and financial commitments that may be beyond many STEM programs. Secondly, introductory classes at many institutions are often an integrated prerequisite for many other classes, which limits changes in the nature of the course or curriculum. As a result, improving the quality of teaching remains as the most cost effective, tangible, and timely improvement that STEM departments may impose to improve student engagement and retention. Despite recent articles that describe effective methods for improving learning in STEM courses (Ebert-May and others, 1997; Hake, 1998; Paulson, 1999; Crouch and Mazur, 2001; Wyckoff, 2001; Oliver-Hoyo and others, 2004; Knight and Wood, 2005; Beichner and others, 2007), there is a relative paucity of data on effective pedagogies to improve learning in the geosciences. In particular, active learning strategies are still rare (10%) in introductory geoscience courses with more than 30 students (McDonald and others, 2005). While some geoscience instructors are engaging in scholarly teaching that honors best practice findings across STEM disciplines (McDonald and others, 2005; King, 2008), there has been less engagement in the scholarship of educational research to provide the empirical confirmation that these changes result in improvements in students learning.

This paper provides a description of the incorporation of research-supported practices into large, general education Earth Science classes. We describe formative assessment strategies that range from simple multiple choice questions to use of physical models. These techniques seek to directly address student conceptual understanding and may be structured to address multiple levels of cognitive development. Qualitative and quantitative analyses show that such methods are preferred by students, improve student retention, produce no decrease in content knowledge, promote deeper understanding of course material, and increase logical thinking skills (McConnell and others, 2003, 2005).

Active Learning in Large Introductory Geoscience Classes

We taught a series of a large (n=150) earth science classes at a large (student population = 24,000) Midwestern university. Similar courses are taught at most colleges and universities and feature students with a wide range of intellectual development. Many of these students do not have the skills to understand the abstract scientific concepts traditionally discussed in introductory classes. Many geological concepts will remain unlearned without appropriate activities that build on a foundation of concrete examples. The good news is that these same students can improve their thinking skills when they participate in challenging in-class collaborative learning exercises with their more intellectually sophisticated peers. While the exercises themselves are important in promoting the development of higher-order thinking skills, the group interaction also appears to be a significant contributor to the improvement of reasoning (McConnell and others, 2003, 2005).

A majority of the students in our classes were white (81%), freshmen (73%), under the age of 26 (95%) and had not declared a major (69%). Each class consisted of three, 50-minute lecture periods per week without a corresponding lab. Each student was placed in a designated four-person group to facilitate peer instruction. Throughout the lecture, the instructor used peer instruction (Mazur, 1997; Crouch and others, 2001; McConnell and others, 2003, 2006) to assess student understanding of the key concepts and recorded individual answers using an electronic personal response system.

<u>Sample Lessons</u>

The following is a description of a lesson on seafloor spreading, a key concept in understanding the theory of plate tectonics.

- Students were first asked to *sketch a view of the Earth's interior* in an effort to address their preconceptions about the structure of the Earth. This was followed by a *conceptest that showed four potential simplified cross sections through Earth (Figure 1).* Students were asked to pick the most accurate representation. The most popular choices indicated that students harbored a misconception about the size of Earth's core. As with all subsequent conceptests, the instructor summarized the reasoning behind the correct answer choice.

- Next we illustrated the relative size and positions of crust, mantle and core and emphasized the characteristics of the lithosphere and asthenosphere. This was followed a definition of a tectonic plate and brief introduction of the concept of plate tectonics. *Students were asked to consider how plate tectonics differed from the previously discussed concept of continental drift.* The instructor collected student responses and listed them on the board before analyzing the differences between the two concepts.
- We then listed the 5 key sets of observations (seafloor topography, age of seafloor, distribution of earthquakes, volcanoes, heat flow) to be considered in providing support for the concept of seafloor spreading, a precursor to plate tectonics. (Note: only the first two features would be discussed during this lecture.)
- Students were then asked to respond to a conceptest on the *relationships between age and topography of the seafloor*, drawn from their homework reading on this topic. This was followed by a *conceptest asking students to predict the profile view of the Atlantic Ocean floor (Figure 1).* On the basis of student answers to the conceptests, it was clear that the majority of students understood some basic concepts about the topography and age of the seafloor. Consequently, we were able to move relatively quickly through the next few slides and an additional conceptest focusing on the characteristics and distribution of oceanic ridges and trenches.
- Next, students were presented with a map of the age of the ocean floor and were asked to work in their groups to complete an *image analysis exercise to compare the*

patterns they observed with the seafloor topography discussed earlier. A series of student observations were recorded by the instructor who summarized the key observations for the class. (The next class began with a similar series of image analysis exercises for maps of earthquakes, volcanoes, and global heat flow.)

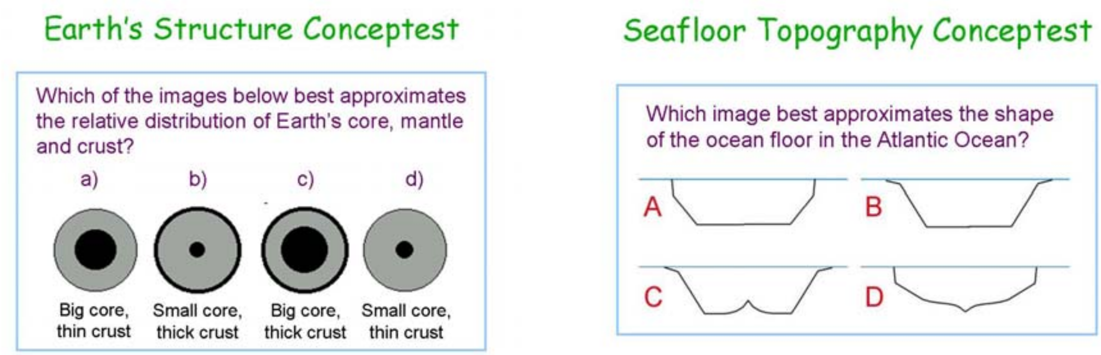

Figure 1. Sample conceptests used in "Seafloor spreading" lecture, Spring 2007. Left) Only 12% of class chose the correct interpretation of Earth's structure (choice a); 84% of the class chose an image with a small core. On the basis of the low correct response rate, we did not make this question available for peer instruction. Instead, we emphasized the relative proportions of these features in the subsequent lecture. Right) 66% of students correctly identified the exaggerated shape of the floor of the Atlantic Ocean (C) on their own; 84% responded correctly after peer instruction.

This lesson had several learning objectives such as students should be able to describe the structure of Earth, explain how compositional layers are components of mechanical layers (lithosphere, asthenosphere), define the term plate tectonics, and use seafloor topography and age to support interpretations that the plates have changed position over time (this idea was more fully detailed in the following lecture). Students were assigned tasks that required them to confront their preconceptions, allowed them to reflect on their understanding of key concepts, linked this information to previous knowledge, and asked questions requiring the use of comprehension (conceptests) and analysis (image analysis, open-ended questions). At several points during each day's lecture, students were given an opportunity for collaborative learning using a variety of exercises targeting different levels of Bloom's Taxonomy (Table 1; McConnell and others, 2003). There was no time when the lecture extended for more than 10 minutes before students were asked to assess their understanding of some aspect of their learning.

Table 1: Formative Assessment Methods used in introductory earth science class amatched to levels of Bloom's Taxonomy

Bloom's Taxonomy Level	Learning Strategy (Assessment Method)					
	Concep-tests	Venn Diagram	Image Analysis	Concept Map	Open-ended Question	Evaluation Rubric
Knowledge	•	•	•	•	•	•
Comprehension	•	•	•	•	•	•
Application	•	•	•	•	•	•
Analysis		•	•	•	•	•
Synthesis				•	•	•
Evaluation					•	•

Initial active learning classes used peer instruction and paper-and-pencil exercises but incorporated relatively few hands-on materials. Over time we gradually added rock samples, in-class demonstrations, and simple models to the exercises. Some of these models produced additional learning gains as measured using pre/post conceptests. For example, to address the common misconceptions concerning the seasons, the students used a small, inexpensive, physical model (Figure 2) to scaffold their conceptual understanding. Besides the instructor's lecture, the students worked within pre-assigned groups to complete three modeling activities. Each activity took approximately two to five minutes to complete. After each activity, the instructor used a larger version of the model to summarize the salient observations.

An initial review of the data found that the students in all but one of the classes answered a higher percentage of post-model questions than pre-model questions (Figure 3, see caption comments). Classroom observations revealed that the students would use the model to explain the underlying concepts to their peers. Students also used the model to mimic the illustrations in the book and recreated the instructor-led demonstrations. Students in all classes correctly answered a similar percentage of pre-model questions, but students who used the hands-on model correctly answered a larger percentage of post-model questions than students in the control classes (Figure 3). Students in the model-use classes correctly answered 55.6% of the pre-model conceptest questions

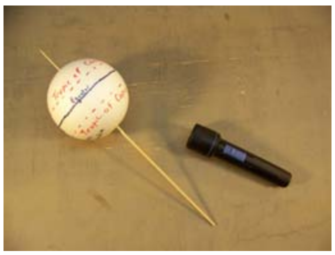

Figure 2. Student model for Earth-sun-seasons exercise. The model consisted of a small four-inch foam ball with hand-drawn lines representing the equator and tropics. The location of the university and Australia were added as geographic reference points that the instructor used when referring to either hemisphere. The axis of rotation was modeled using a wooden skewer inserted at the south pole of the model and a toothpick inserted at the north-pole. A small, inexpensive flashlight was used to simulate incoming solar energy.

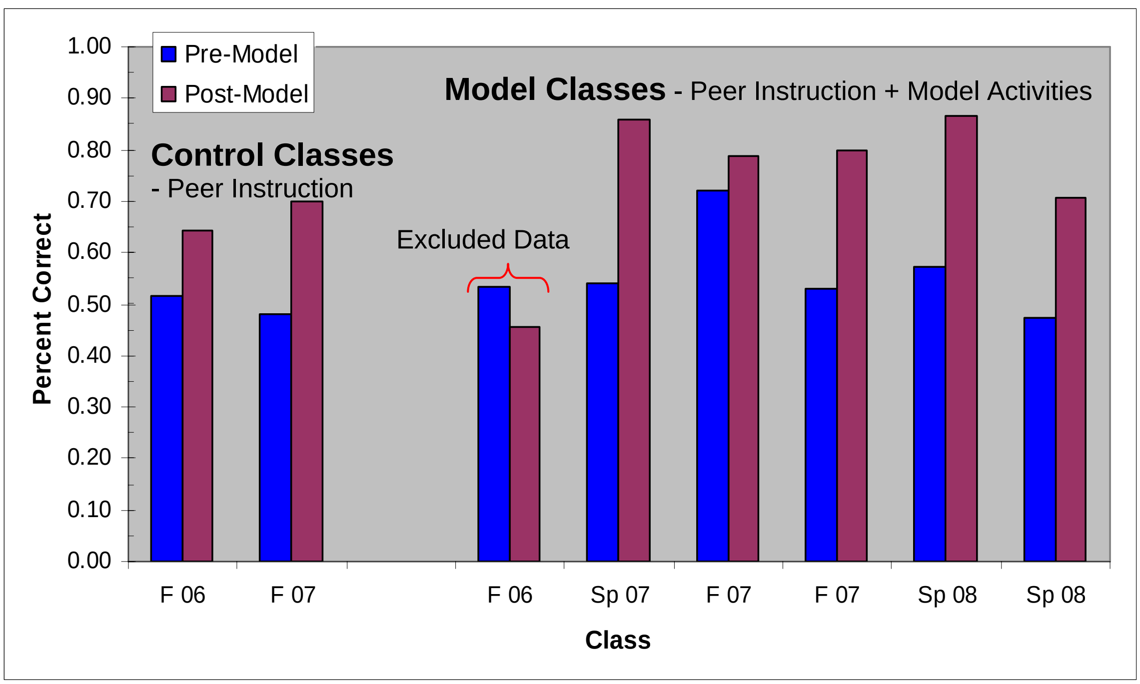

Figure 3. Eight semesters of student response data for control and model classes using the Earth-sun-seasons model (Fig. 2). 526 students answered at least one pre-model conceptest and 531 students answered at least post-model question. The F06 class data was excluded as it was the first time the students had been asked to complete a model activity and the initial set of classroom directions proved insufficient to guide student use of the model.

compared to 50.6% of the questions from the control class, however this difference was not significant at a p = 0.05 confidence level. In contrast, the students in the model-use classes correctly answered a significantly higher proportion of post-model conceptests (73.8%, p=0.002) than their counterparts in the control classes (66.6%).

Evidence for Student Learning

Our primary goal was to foster the development of higher order thinking skills and to encourage student conceptual understanding of the geosciences. Reasoning skills were measured for 741 students using the Group Assessment of Logical Thinking instrument (GALT, Roadrangka and others, 1982, 1983) as a pre- and post-test in ten sections of general education introductory geoscience courses for non-majors titled Earth Science or Environmental Geology. The GALT is a valid and reliable instrument for measuring logical thinking in student populations from sixth grade through college and consistently yields higher scores with increasing grade level (Roadrangka and others, 1982; Bitner, 1991; Mattheis and others, 1992). Success on the GALT test requires competence in five logical operations; proportional reasoning, controlling variables, combinational reasoning, probabilistic reasoning, and correlational reasoning (Roadrangka and others, 1982). The GALT instrument was administered as a pre- and post-test in multiple sections of experimental (active learning) and control (passive lecture) courses taught by several instructors teaching in similar classrooms (McConnell and others, 2005). A little more than half (57%) of the participating students were defined as concrete or transitional thinkers as assessed by the GALT and consequently were cognitively unprepared for instruction that required them to make abstract connections concerning major concepts. The active learning environment was structured to scaffold learning by beginning with exercises employing concrete examples. Abstract thought processes were modeled by their peers as they completed group exercises. Such activities give concrete and transitional students an opportunity to ask questions and build concepts in a non-threatening environment where students addressed concepts that required them to apply knowledge to new situations. Lastly, students have multiple opportunities to recognize gaps in their understanding and to take action to correct them.

Table 2. Changes to students' GALT scores by starting cognitive level. Note that these scores only represent students who completed both the pre-test and post-test. Statistically significant ($p<0.01$ or less) gains occurred for students who were initially concrete or transitional in either group. Students who began the class as abstract thinkers showed no significant gains in logical thinking.

Domain	Pre-GALT	Post-GALT	Gain	p <
Test Group (Active learning Classes, n=465)				
All	6.9 +/- 2.7	7.8 +/- 2.6	0.18	0.001
Concrete	3.1 +/- 1.1	5.2 +/- 2.4	0.24	0.001
Transitional	6.0 +/- 0.8	7.2 +/- 2.0	0.21	0.001
Abstract	9.3 +/- 1.3	9.5 +/- 1.7	0.06	0.1
Control Group (Traditional Classes, n=276)				
All	6.5 +/- 2.7	7.0 +/- 2.7	0.09	0.001
Concrete	3.0 +/- 1.0	4.2 +/- 1.6	0.14	0.001
Transitional	6.2 +/- 0.8	6.8 +/- 2.0	0.11	0.001
Abstract	9.3 +/- 1.1	9.1 +/- 1.6	-0.05	0.25

Students in both experimental and control classes increased their logical thinking skill over the course of the semester (Table 2). Normalized gains for students in the active learning classes were higher for all populations. Students in the experimental group averaged score gains were approximately twice those of the control sections (0.18 versus 0.09).

Our initial results lead us to consider whether students improvement in their GALT scores was a primarily a result of working in groups or was more attributable to working on exercises that required the application of higher order thinking skills. To investigate this question, one instructor taught two sections of the same active learning-style class using the same learning exercises and using collaborative groups in one class but not in the other. Both classes received the identical lectures and class materials. Course grades for students in the section using assigned groups were 5% higher ($p<0.01$) than in the section that employed active learning without the benefit of groups (Table 3). This result was not unexpected since the group work concept has been shown to be successful in many situations (Nelson, 1994; Paulson, 1999; Lord, 2001). Students in structured groups learn from one another and are more likely to get their questions answered as they strive to meet common goals and objectives. They become more socially connected and have more opportunities to address their misconceptions. This is particularly true in large format sections where individual attention from the instructor is at a premium.

Table 3. Course averages as a function of cognitive level when comparing sections that employ group exercises compared to those where the exercises were completed as individuals (n = 119).

Learning Strategy	Concrete	Transitional	Abstract	Overall
Active Learning (groups)	74%	77%	84%	80%
Traditional Lecture	70%	74%	79%	75%
p <	0.08	0.07	0.01	0.0005

As described in the previous section, conceptest results typically showed a shift toward a correct answer choice in class but occasionally revealed a deeply held misconception about a key concept. While conceptests provide a ready mechanism for formative assessment of student learning, instructors must seek an alternative measure of summative evaluations. One instrument in particular, the Geoscience Concept Inventory (GCI; Libarkin and Anderson, 2005) has become available in the geosciences in recent years. The GCI is a valid and reliable measure of student understanding of critical geological concepts (Libarkin and Anderson, 2005). However, unlike in other disciplines (e.g., physics, Hake, 1998), there is no identified normalized gain score that is widely accepted as indicative of a significant improvement in student learning. Generally, pre- and post-testing scores on the GCI are lower than those on instruments such as the Force Concept Inventory, resulting in smaller gains.

Libarkin and Anderson's original GCI data were reanalyzed in an effort to better constrain the range of gains that would be representative of introductory geoscience classes. We reanalyzed paired data from 27 introductory classes and found that 14 classes showed significant differences in pre and post-test scores ($p<0.01$; Figure 4a).

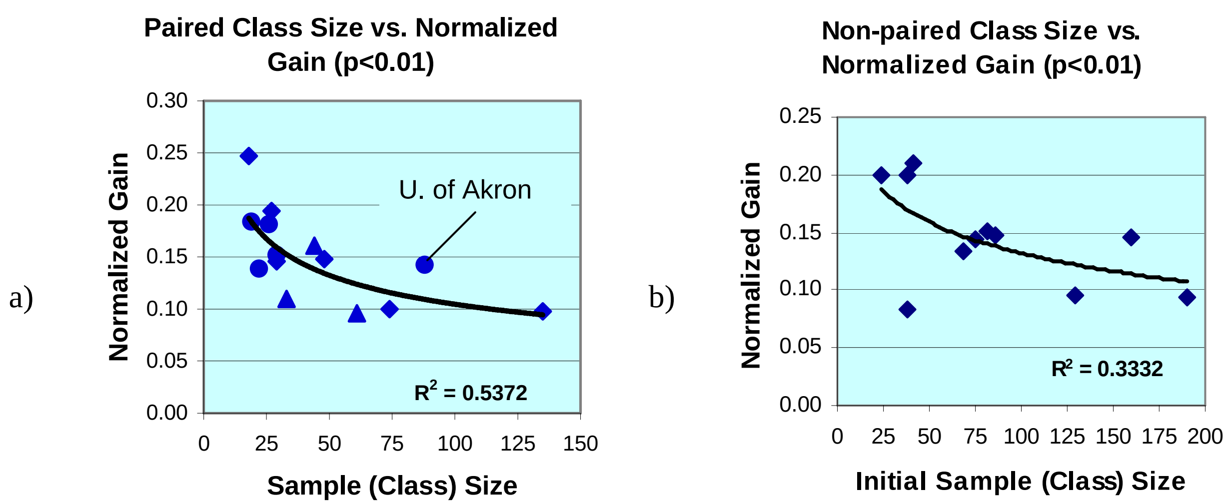

Figure 4. a) Normalized gain for paired pre- and post-test scores for the Geoscience Concept Inventory for 653 students enrolled in introductory geoscience classes at 14 institutions (original data from Julie Libarkin). Normalized gain scores decrease with increasing class size. Triangles are "traditional" classes with at least 70% of class time devoted to lecture. Circles represent "active learning" classes that had 50% or less class time devoted to lecture and incorporated other pedagogical strategies. b) Normalized gain for non-paired pre- and post-scores for GCI for students in 11 institutions. Kortz and others (2008) had a normalized gain of 0.143 in with an initial sample size of 86 students.

Normalized gain scores (g, where g = post-pre/1-pre) for all classes ranged from -0.017 to 0.33 and exhibited no clear trend when compared to environmental factors such as class size. However, when we considered only classes showing a statistically significant difference between pre and post-scores, the range of normalized gain scores diminished (0.1-0.25) and a trend of decreasing gain with increasing class size became apparent

(Figure 4a). A similar, though less robust, trend is seen with $p<0.05$ ($g=0.1-0.25$, $R^2=0.4815$) and with more widely–distributed, non-paired data ($p<0.01$, $g=0.07-0.21$, $R^2=0.3332$, Figure 4b).The GCI used as a pre- and post-test assessment for an early version of the active learning course described herein. The normalized gain in student performance in our classes at the University of Akron was 0.143, well above the trend line for other early administrations of the test in classes of similar size (Figure 4a). Although the data is limited, our analysis can be interpreted to infer some potential learning standards for small and large introductory geoscience classes. Smaller classes ($n<50$) should show statistically significant differences in pre/post GCI scores <u>and</u> normalized gain scores of 0.1-0.2. Smaller normalized gains, perhaps in a range of 0.05-0.15, may be considered reasonable for larger classes. These ranges may require some adjustment when using non-paired data (Figure 4b). For example, Kortz and others (2008) used the GCI to measure student learning in large ($n>80$) classes featuring an active learning strategy using lecture tutorials. They reported non-paired, statistically significant ($p<0.01$) GCI results that yield a normalized gain of 0.143. This result would plot slightly above the trendline for similar non-paired data (Figure 4b).

<u>Next Steps</u>

Few authors have published research results that show consistent gains in student learning across introductory geoscience courses over the last 5 years (e.g, Kortz and others, 2008). The paucity of data for introductory courses is in contrast to the rich data set available for equivalent classes in physics (e.g., Hake, 1998; Crouch and Mazur, 2001). Regardless of gains achieved using widely disseminated instruments like the

Geoscience Concept Inventory, there is no standard against which instructors can evaluate their class. Further, despite of the merits of the GCI, there is a need for additional instruments that will allow instructors to assess student responses to open ended questions that seek to foster the development higher order skills such as analysis, synthesis, and evaluation.

The good news is that there are readily available online resources such as SERC's Teaching Entry Level Geoscience site (http://serc.carleton.edu/introgeo/index.html) that provides instructors with a variety of teaching methods that they can fit to their courses and teaching styles. What is required is a consistent effort by faculty in a range of introductory courses to report data that they may already have collected to better constrain future research.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 0087894. The ideas in this paper have been improved and focused through discussions with numerous colleagues, especially David Steer, Katharine Owens, Catharine Knight, Kyle Gray, and Julie Libarkin. Some of the model data presented here is taken from a submitted article by Gray and others.

References

Beichner, R.J., Saul, J.M., Abbott, D.S., Morse, J.J., Deardorff, D.L., Allain, R.J., Bonham, S.W., Dancy, M.H., and Risley, J.S., 2007, The student-centered activities for large enrollment undergraduate programs (SCALE-UP) project, in Research-Based Reform of University Physics, E. F. Redish and P. J. Cooney, editors, *Volume 1 of Reviews in Physics Education Research*, 42 p. found at http://www.compadre.org/PER/per_reviews/media/volume1/SCALE-UP-2007.pdf, accessed October 7, 2008.

Bitner, B. L., 1991, Formal operational reasoning modes: Predictors of critical thinking abilities and grades assigned by teachers in science and mathematics for students in grades nine through twelve: Journal of Research in Science Teaching, v.28, p. 265-274.

Crouch, C.H., and Mazur, E., 2001, Peer Instruction: Ten years of experience and results. American Journal of Physics, v. 69, p. 970-977.

Ebert-May, D., Brewer, C., and Allred, S., 1997, Innovation in large lectures – teaching for active learning: Bioscience, v.47, p. 601-607

Hake, R.R., 1998, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, v. 66, p. 64-74.

King, H., 2008, The state of evidence in geoscience education research, found at http://www7.nationalacademies.org/bose/Helen%20King%20Paper%20Geoscience%20Education%20Research.pdf, accessed October 7, 2008

Knight, J.K., and Wood, W.B., 2005, Teaching more by lecturing less, Cell Biology Education, v. 4, p. 298-310.

Kortz, K.M., Smay, J.J., and Murray, D.P., 2008, Increasing learning in introductory geoscience courses using lecture tutorials, Journal of Geoscience Education, v. 56, #3, p. 280-290.

Libarkin, J.C., and Anderson, S.W., 2005, Assessment of Learning in Entry-Level Geoscience Courses: Results from the Geoscience Concept Inventory. Journal of Geoscience Education, v. 53, #4, 394-401.

Lord, T. R., 2001, 101 Reasons for using cooperative learning in biology teaching. American Biology Teacher, v. 63, p. 30-38.

Mattheis, F. E., Spooner, W. E., Coble, C. R., Takemura, S., Matsumoto, S., Matsumoto, K., and Yoshida, A., 1992, A study of the logical thinking skills and integrated process skills of junior high school students in North Carolina and Japan, Science Education, v.76, p. 211-222.

Mazur, E., 1997, Peer instruction: A user's manual: Prentice Hall, 253p.

McConnell, D. A., Steer, D. N., and Owens, K. D., 2003, Assessment and active learning strategies for introductory geology courses, Journal of Geoscience Education, v. 51, p. 205-216.

McConnell, D.A., Steer, D.N., Owens, K., & Knight, C., 2005, How students think: Implications for learning in introductory geoscience courses: Journal of Geoscience Education, v. 53, #4, p. 462-470.

McConnell, D.A., Steer, D.N., Owens, K.D., Knott, J.R., Van Horn, S., Borowski, W., Dick, J., Foos, A., Malone, M. McGrew, H., Greer, L, and Hearney, P. J. 2006. Using conceptests to assess and improve student conceptual understanding in introductory geoscience courses. *Journal of Geoscience Education*, 54 (1), 61-68.

Macdonald, R.H., Manduca, C.A., Mogk, D.W., and Tewksbury, B.J., 2005, Teaching methods in undergraduate geoscience courses: Results of the 2004 On the Cutting Edge Survey of US faculty, Journal of Geoscience Education, v. 53, #3, p. 237-252.

National Science Board, 2003. The science and engineering workforce: Realizing America’s potential. National Science Foundation, Arlington.

Nelson, C.E., 1994, Critical thinking and collaborative learning, New directions in teaching and learning, #59, p. 45-58.

Oliver-Hoyo, M.T., Allen, D, Hunt, W.F., Hutson, J., and Pitts, A., 2004, Effects of an active learning environment: Teaching innovations at a Research I institution, Journal of Chemical Education, v. 81, #3, p. 441-448.

Paulson, D. R., 1999, Active learning and cooperative learning in the organic chemistry lecture class. Journal of Chemical Education, v. 76, p. 1136-1140.

Roadrangka, V., Yeany, R., and Padilla, M., 1982, GALT, Group Test of Logical Thinking. Athens, Georgia: University of Georgia.

Roadrangka, V., Yeany, R.H., and Padilla, M.J., 1983, The construction and validation of Group Assessment of Logical Thinking (GALT), Paper presented at the annual meeting of the National Association for Research in Science Teaching, Dallas, Texas, April.

Seymour, E., and Hewitt, N.M., 1997. Talking about leaving: Why undergraduates leave the sciences. Westview Press, Boulder.

Strenta, A.C., Elliott, R., Adair, R., Matier, and Scott, J., 1994. Choosing and leaving science in highly selective institutions. Research in Higher Education, v. 35, #5, p. 513-547.

Tobias, S., 1990, They’re not dumb, they’re different: Stalking the second tier: Research Corporation, 94 p.

Wyckoff, S., 2001, Changing the culture of undergraduate science teaching. Journal of College Science Teaching: v.30, p. 306-312.