

STEM Topic: Assessment and Evaluation

Richardson, J. 2012. Concept inventories: Tools for uncovering STEM students' misconceptions. Invention and Impact: Building Excellence in Undergraduate Science, Technology, Engineering and Mathematics (STEM) Education. American Association for the Advancement of Science.

Concept inventories are receiving increased interest from STEM faculty. What are concept inventories, why the interest, and what do I need to know about concept inventories? This chapter answers these questions in the following order. In the first section, you will read a brief history of STEM concept inventory development, which should answer the question, "Why the interest in concept inventories?" In the next two sections, you will read first a short discussion on the theory of assessment as it applies to concept inventories and then a description of how to construct a concept inventory. Together, these two sections should answer the question, "What is a concept inventory?" And finally, you will read how others have used concept inventories and related tools to improve their teaching effectiveness.

Halcrow, C., Iiams, M. 2011. You can build it, but will they come? *Primus*. Vol. 21, 323-337.

Studies show that more students fail or withdraw from college mathematics courses than any other. To address this concern, the Mathematics Department at the University of North Dakota opened its Mathematics Learning Center (MLC) in the fall of 2000. In this study, the effectiveness of the MLC and the free tutoring offered for students in freshman level mathematics courses was examined.

The quantitative portion of the study examined the difference between course grades in experimental and control sections of four distinct freshman level mathematics courses. Students in the experimental sections were required to attend the Mathematics Learning Center (MLC) for one hour weekly while students in the control sections were simply informed of the availability of tutoring in the MLC. The qualitative portion of the research utilized methodologies of a phenomenological study through in-depth interviews with 13 participants. Three conclusions are offered: 1) Lower level or lower ability students are less likely to attend the MLC and seek help from tutors; 2) Once students got over their fears of engaging with tutors, they found them friendly and helpful, and believed they had greater success because of the tutoring; and 3) A positive correlation existed between time spent in the MLC and course grade for experimental section students.

Maltese, A.V., Tai, R.H. 2011. Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. Wiley Periodicals, Inc. *Science Education*. 1-31.

As the global economic crisis continues, sustaining the United States' position as a leader in research and development is a top concern of policy makers. Looking to increase the number of students pursuing degrees in STEM (science, technology, engineering, and mathematics), calls for improved mathematics and science education abound. We completed a two-part analysis to assess the school-based factors related to students choosing to complete a major in STEM. The results indicate that the majority of students who concentrate in STEM make that choice during high school, and that choice is related to a growing interest in mathematics and science rather than enrollment or achievement. These results indicate that the current policy focus on

advanced-level course taking and achievement as measures to increase the flow of students into STEM may be misguided.

Moss, K., Crowley, M. 2011. Effective learning in science: The use of personal response systems with wide range of audiences. *Computers & Education*. Vol. 56, 36-43.

This paper describes the flexibility of Personal Response Systems (PRSs), (also known as 'clickers' or electronic voting systems (EVS)), as part of strategies to support students' learning in science. Whilst variants of this technology began to appear 12 years ago, there is now a steadily increasing adoption of these systems within higher education, including science programmes, and this use has grown significantly in the last six years. They have previously been shown to offer a measurable learning benefit. Typically, someone at an institution buys these systems for learning support and they never make it out of their cases. Far less work has been done with these systems at school level. In this practitioner based paper, the broad range of practical uses for these systems is described in a variety of formal and informal learning situations – from testing the understanding of science concepts (from primary aged school children up to physics undergraduates), to undertaking evaluation of events as well as public participation in data collection for research on attitudes to careers. In addition, the data collected on such handsets can be mapped to demographic factors such as gender and age yielding further layers of analysis. Overall this is a highly flexible and transferable approach to the use of interactive technology for engaging learners of all ages as well as carrying out research.

Lewis, S.E., Shaw, J.L., Freeman, K.A. 2010. Creative exercises in general chemistry: A student-centered assessment. *Journal of Science Teaching*. Vol. 40. 48-44.

Creative exercises (CEs) are a form of assessment in which students are given a prompt and asked to write down as many distinct, correct, and relevant facts about the prompt as they can. Students receive credit for each fact that they include that is related to the prompt and distinct from the other facts they list. With CEs, students have an opportunity to demonstrate their knowledge and the opportunity to select the information that they believe is related to the prompt. In addition, CEs encourage students to connect concepts because any relevant information presented can assist them in completing the CEs. This paper describes the use of CEs in a college level chemistry class, including student answers to the CEs and a survey of students' impression of CEs.

McDermott, M.A., Hand, B. 2010. A secondary reanalysis of student perceptions of non-traditional writing tasks over a ten year period. *Journal of Research in Science Teaching*. Vol. 47, 518-539.

This study aims to add to the growing research related to the implementation of non-traditional writing tasks in classrooms to encourage science literacy. A secondary reanalysis methodology was employed to review student interviews collected as a part of several individual studies during a ten year research program. This method established an interpretive framework different than the particular frameworks guiding the individual studies. In doing so, a greater ability to generalize findings was sought. Main assertions emerging from the student responses analyzed include recognition of benefits of non-traditional writing, recognition of the need for particular task characteristics to encourage these benefits, and recognition of greater cognitive activity than is present in typical science classroom writing.

Sirum, K.L, Madigan, D. 2010. Assessing how science faculty learning communities promote scientific teaching. *Biochemistry and Molecular Biology Education*. Vol. 38, 193-202.

Although there is a need for continued pedagogical advancement in science undergraduate education, what is needed more urgently is more widespread adaptation of pedagogical practices that research has already shown to promote learning. Those practices include interactive engagement pedagogies such as active learning and inquiry-based learning. The need now is to find ways to integrate and institutionalize these evidence-based strategies for teaching science and to help science faculty learn about and implement them. Scientific Teaching Learning Communities (STLCs) create a culture that values scholarly teaching within science departments, important for bridging the gap between science and education and for improving undergraduate science learning. Evidence for the impact of STLCs on the student-learning environment was obtained through the development and use of the Participant Assessment of Learning Gains survey, an adaptation of the online Student Assessment of Learning Gains survey originally developed by Seymour et al. Data reveal how STLCs are transforming faculty behavior and directly affecting what they do in their science classrooms.

Armbruster, P., Patel, M., Johnson, E., Weiss, M. 2009. Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *Life Science Education*. Vol. 8, 203-213.

We describe the development and implementation of an instructional design that focused on bringing multiple forms of active learning and student-centered pedagogies to a one-semester, undergraduate introductory biology course for both majors and non-majors. Our course redesign consisted of three major elements: 1) reordering the presentation of the course content in an attempt to teach specific content within the context of broad conceptual themes, 2) incorporating active and problem-based learning into every lecture, and 3) adopting strategies to create a more student-centered learning environment. Assessment of our instructional design consisted of a student survey and comparison of final exam performance across 3 years – 1 year before our course redesign was implemented (2006) and during two successive years of implementation (2007 and 2008). The course restructuring led to significant improvement of self-reported student engagement and satisfaction and increased academic performance. We discuss the successes and ongoing challenges of our course restructuring and consider issues relevant to institutional change.

Corner, A., Hahn, U. 2009. Evaluating science arguments: Evidence, uncertainty, and argument strength. *Journal of Experimental Psychology: Applied*. Vol. 15, 199-212.

Public debates about socioscientific issues are increasingly prevalent, but the public response to messages about, for example, climate change, does not always seem to match the seriousness of the problem identified by scientists. Is there anything unique about appeals based on scientific evidence—do people evaluate science and nonscience arguments differently? In an attempt to apply a systematic framework to people's evaluation of science arguments, the authors draw on the Bayesian approach to informal argumentation. The Bayesian approach permits questions about how people evaluate science arguments to be posed and comparisons to be made between the evaluation of science and nonscience arguments. In an experiment involving three separate argument evaluation tasks, the authors investigated whether people's evaluations of science and nonscience arguments differed in any meaningful way. Although

some differences were observed in the relative strength of science and nonscience arguments, the evaluation of science arguments was determined by the same factors as nonscience arguments. Our results suggest that science communicators wishing to construct a successful appeal can make use of the Bayesian framework to distinguish strong and weak arguments.

Glynn, S.M., Taasobshirzai, G., Brickman, P. 2009. Science motivation questionnaire: Construct validation with nonscience majors. *Journal of Research in Science Teaching*. Vol. 46, 127-146.

This study examined how 770 nonscience majors, enrolled in a core-curriculum science course, conceptualized their motivation to learn science. The students responded to the Science Motivation Questionnaire, a 30-item Likert-type instrument designed to provide science education researchers and science instructors with information about students' motivation to learn science. The students' scores on the Science Motivation Questionnaire were reliable and related to students' high school preparation in science, GPA in college science courses, and belief in the relevance of science to their careers. An exploratory factor analysis provided evidence of construct validity, revealing that the students conceptualized their motivation to learn science in terms of five dimensions: intrinsic motivation and personal relevance, self-efficacy and assessment anxiety, self-determination, career motivation, and grade motivation. Women and men had different profiles on these dimensions, but equivalent overall motivation to learn science. Essays by all of the students explaining their motivation to learn science and interviews with a sample of the students were used to interpret Science Motivation Questionnaire scores. The findings were viewed in terms of a social-cognitive theory of learning, and directions for future research were discussed.

Ortiz-Robinson, N.L., Ellington, A.J. 2009. Learner-centered strategies and advanced mathematics: A study of students' perspectives. *Primus*. Vol. 19, 463-472.

A number of learner-centered strategies were implemented during a two semester course in real analysis that is traditionally taught in lecture format. We seek to understand the role that these strategies can have in this proof-intensive theoretical mathematics classroom and the perceived benefits by the students. Although learner-centered strategies are a welcome addition in many applied mathematics courses and are known to be successful, the literature indicates that these remain largely absent from more advanced courses [9]. In an effort to correlate student resistance and acceptance of these strategies in different classroom settings we included an applied differential equations course in the study. Student feedback was obtained for two semesters of the real analysis course and compared to the feedback obtained during one semester of the differential equations course.

Schwartz, M.S., Sadler, P.M., Sonnert, G., Tai., R.H. 2009. Depth versus breadth: How content coverage in high school science courses relates to later success in college science coursework. *Wiley Periodicals, Inc., Science Education*. Vol. 93, 798-826.

This study relates the performance of college students in introductory science courses to the amount of content covered in their high school science courses. The sample includes 8310 students in introductory biology, chemistry, or physics courses in 55 randomly chosen U.S. colleges and universities. Students who reported covering at least 1 major topic in depth, for a month or longer, in high school were found to earn higher grades in college science than did students who reported no coverage in depth. Students reporting breadth in their high school

course, covering all major topics, did not appear to have any advantage in chemistry or physics and a significant disadvantage in biology. Care was taken to account for significant covariates: socioeconomic variables, English and mathematics proficiency, and rigor of their preparatory high science course. Alternative operationalizations of depth and breadth variables result in very similar findings. We conclude that teachers should use their judgment to reduce coverage in high school science courses and aim for mastery by extending at least 1 topic in depth over an extended period of time.

King, D.B., Joshi, S. 2008. Gender differences in the use and effectiveness of personal response devices. *Journal of Science Educational Technology*. Vol. 17, 544-552.

The use of personal response devices (or “clickers”) in the classroom has increased in recent years. While few quantitative studies on the effectiveness of clickers have been published, it is generally reported that clickers have been well-received by the students who use them. Two separate populations (Winter 2006 and Spring 2006) of engineering students were given clickers to use during a general chemistry class. Clicker use was compared to student grades for each course. During both terms, a higher percentage of female students than male students “actively participated” in the lectures, where active participation as defined as answering more than 75% of the clicker questions over the course of the term. Active male students earned final grades about 10 points higher than non-active male students. Active female students, however, scored only about 5 points higher than non-active female students. Student learning was assessed by comparing performance on exam questions and clicker questions with similar content. Students who answered clicker questions correctly were 11–13% more likely to answer the corresponding exam questions correctly than were students who did not answer the clicker question. In this paper, we demonstrate the effectiveness of clicker use in the classroom and examine gender differences associated with this use.

Mestre, J.P. 2008. Learning Goals in Undergraduate STEM Education and Evidence for Achieving Them. National Academies White Paper. http://www7.nationalacademies.org/BOSE/Mestre_CommissionedPaper.pdf

In this white paper I propose a short list of learning goals in undergraduate STEM education and suggest the types of evidence that would indicate whether or not the learning goals are being achieved. Both the learning goals and proposed evidence will be accompanied by arguments and discussions about the relevance of the proposed goals given today’s context. I also discuss why certain types of evidence should carry more weight than others, where current gaps in evidence exist, and why the quality of evidence is pivotal in promoting the adoption of promising instructional practices in undergraduate STEM instruction.

Morse, D., Jutras, F. 2008. Implementing concept-based learning in a large undergraduate classroom. *Life Sciences Education*. Vol 7, 243-253.

An experiment explicitly introducing learning strategies to a large, first-year undergraduate cell biology course was undertaken to see whether awareness and use of strategies had a measurable impact on student performance. The construction of concept maps was selected as the strategy to be introduced because of an inherent coherence with a course structured by concepts. Data were collected over three different semesters of an introductory cell biology course, all teaching similar course material with the same professor and all evaluated using

similar examinations. The first group, used as a control, did not construct concept maps, the second group constructed individual concept maps, and the third group first constructed individual maps then validated their maps in small teams to provide peer feedback about the individual maps. Assessment of the experiment involved student performance on the final exam, anonymous polls of student perceptions, failure rate, and retention of information at the start of the following year. The main conclusion drawn is that concept maps without feedback have no significant effect on student performance, whereas concept maps with feedback produced a measurable increase in student problem-solving performance and a decrease in failure rates.

Park Rogers, M.A., Abell, S.K., 2008. The design, enactment, and experience of inquiry-based instruction in undergraduate science education: A case study. Wiley Periodicals, Inc. *Science Education*. Vol. 92, 591-607.

The purpose of this study was to understand one case of undergraduate inquiry-based instruction through the words and actions of students and instructors. The data sources included fieldnotes from 16 of 29 classes, two sets of student and instructor interviews (beginning and end of the semester), and a collection of artifacts, such as the laboratory manual, lecture handouts, and the course syllabus. The participants for this study included four faculty instructors and two purposively selected student groups, totaling eight students. We found the instructors' two course goals, (a) teaching students how scientists do science and (b) using an interdisciplinary approach to develop students' content knowledge of the big ideas in science, were consistent with our observations and the students' descriptions of their experience in the course. However, we observed, and the students also noted, an important feature of the course that the instructors did not describe as a course goal was its reliance on the social nature of learning. This telling case demonstrates that inquiry-based instruction is achievable in undergraduate science education. Implications are discussed for college science instructors interested in inquiry teaching.

Sampson, V., Clark, D.B. 2008. Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future direction. Published online 13 March, 2008, in Wiley InterScience (www.interscience.wiley.com).

Theoretical and empirical research on argument and argumentation in science education has intensified over the last two decades. The term "argument" in this review refers to the artifacts that a student or a group of students create when asked to articulate and justify claims or explanations whereas the term "argumentation" refers to the process of constructing these artifacts. The intent of this review is to provide an overview of several analytic frameworks that science educators use to assess and characterize the nature of or quality of scientific arguments in terms of three focal issues: structure, justification, and content. To highlight the foci, affordances, and constraints of these different analytic methods, the review of each framework includes an analysis of a sample argument. The review concludes with a synthesis of the three focal issues and outlines several recommendations for future work. Ultimately, this examination and synthesis of these frameworks in terms of how each conceptualizes argument structure, justification, and content is intended

Garin-Doxas, K., Klymkowsky, M., Elrod, S. 2007. Building, using, and maximizing the impact of concept inventories in the biological sciences: Report on a National Science Foundation-sponsored conference

on the construction of concept inventories in the biological sciences. *Life Science Education*. Vol. 6, 277-282.

The meeting "Conceptual Assessment in the Biological Sciences" was held March 3-4, 2007, in Boulder, Colorado. Sponsored by the National Science Foundation was hosted by University of Colorado, Boulder's Biology Concept Inventory Team, the meeting drew together 21 participants from 13 institutions, all of whom had received National Science Foundation funding for biology education. Topics of interest included Introductory Biology, Genetics, Evolution, Ecology and the Nature of Science. The goal of the meeting was to organize and leverage current efforts to develop concept inventories for each of these topics. These diagnostic tools are inspired by the success of the Force Concept Inventory, developed by the community of physics educators to identify student misconceptions about Newtonian mechanics. By working together, participants hope to lessen the risk that groups might develop competing rather than complementary inventories.

Hrepic, Z., Zollman, D.A., Rebello, N.S. 2007. Comparing students' and experts' understanding fo the content of a lecture. *Journal of Science Education and Technology*. Vol. 16, 213-225.

In spite of advances in physics pedagogy, the lecture is by far the most widely used format of instruction. We investigated students' understanding and perceptions of the content delivered during a physics lecture. A group of experts (physics instructors) also participated in the study as a reference for the comparison. During the study, all participants responded to a written conceptual survey on sound propagation. Next, they looked for answers to the survey questions in a videotaped lecture by a nationally known teacher. As they viewed the lecture, they indicated instances, if any, in which the survey questions were answered during the lecture. They also wrote down (and if needed, later explained) the answer, which they perceived was given by the instructor in the video lecture. Students who participated in the study were enrolled in a conceptual physics course and had already covered the topic in class before the study. We discuss and compare students' and experts' responses to the survey questions before and after the lecture.

McDaniel, C.N., Lister, B.C., Hanna, M.H., Roy, H. 2007. Increased learning observed in redesigned introductory biology course that employed web-enhanced, interactive pedagogy. *Life Sciences Education*. Vol. 6, 243-249.

Our Introduction to Biology course (BIOL 1010) changed in 2004 from a standard instructor centered, lecture-homework-exam format to a student-centered format that used Web-enhanced, interactive pedagogy. To measure and compare conceptual learning gains in the traditional course in fall 2003 with a section of the interactive course in fall 2004, we created concept inventories for both evolution and ecology. Both classes were taught by the same instructor who had taught BIOL 1010 since 1976, and each had a similar student composition with comparable biological knowledge. A significant increase in learning gain was observed with the Web enhanced, interactive pedagogy in evolution (traditional, 0.10; interactive, 0.19; $p = 0.024$) and ecology (traditional, 0.05; interactive, 0.14; $p = 0.000009$) when assessment was made unannounced and for no credit in the last week of classes. These results strengthen the case for augmenting or replacing instructor-centered teaching with Web-enhanced, interactive, student centered teaching. When assessment was made using the final exam in the interactive course, for credit and after studying, significantly greater learning gains were made in evolution

(95%, 0.37, $p < 0.0001$) and ecology (143%, 0.34, $p < 0.000003$) when compared with learning gains measured without credit or study in the last week of classes.

Oakley, B.A., Hanna, D.M., Kuzmyn, Z., Felder, R.M. 2007. Best practices involving teamwork in the classroom: Results from a survey of 6435 engineering student respondents. *IEEE Transactions on Education*. Vol. 50, 266-273.

A teamwork survey was conducted at Oakland University, Rochester, MI, in 533 engineering and computer science courses over a two-year period. Of the 6435 student respondents, 4349 (68%) reported working in teams. Relative to the students who only worked individually, the students who worked in teams were significantly more likely to agree that the course had achieved its stated learning objectives ($p < 0.001$). Regression analysis showed that roughly one-quarter of the variance in belief about whether the objectives were met could be explained by four factors: 1) student satisfaction with the team experience; 2) the presence of instructor guidance related to teamwork; 3) the presence of slackers on teams; and 4) team size. Pearson product-moment correlations revealed statistically significant associations between agreement that the course objectives had been fulfilled and the use of student teams and between satisfaction with teams and the occurrence of instructor guidance on teamwork skills. These and other results suggest that assigning work to student teams can lead to learning benefits and student satisfaction, provided that the instructor pays attention to how the teams and the assignments are set up.

Preszler, R.W., Dawe, A., Shuster, C.B., Shuster, M. 2007. Assessment of the effects of student response systems on student learning and attitudes over a broad range of biology courses. *Life Science Education*. Vol. 6, 29-41.

With the advent of wireless technology, new tools are available that are intended to enhance students' learning and attitudes. To assess the effectiveness of wireless student response systems in the biology curriculum at New Mexico State University, a combined study of student attitudes and performance was undertaken. A survey of students in six biology courses showed that strong majorities of students had favorable overall impressions of the use of student response systems and also thought that the technology improved their interest in the course, attendance, and understanding of course content. Students in lower-division courses had more strongly positive overall impressions than did students in upper-division courses. To assess the effects of the response systems on student learning, the number of in-class questions was varied within each course throughout the semester. Students' performance was compared on exam questions derived from lectures with low, medium, or high numbers of in-class questions. Increased use of the response systems in lecture had a positive influence on students' performance on exam questions across all six biology courses. Students not only have favorable opinions about the use of student response systems, increased use of these systems increases student learning.

Quitadamo, I.J., Kurtz, M.J. 2007. Learning to improve: Using writing to increase critical thinking performance in general education biology. *Life Sciences Education*. Vol. 6, 140-154.

Increasingly, national stakeholders express concern that U.S. college graduates cannot adequately solve problems and think critically. As a set of cognitive abilities, critical thinking skills provide students with tangible academic, personal, and professional benefits that may ultimately address these concerns. As an instructional method, writing has long been perceived

as a way to improve critical thinking. In the current study, the researchers compared critical thinking performance of students who experienced a laboratory writing treatment with those who experienced traditional quiz-based laboratory in a general education biology course. The effects of writing were determined within the context of multiple covariables. Results indicated that the writing group significantly improved critical thinking skills whereas the nonwriting group did not. Specifically, analysis and inference skills increased significantly in the writing group but not the nonwriting group. Writing students also showed greater gains in evaluation skills; however, these were not significant. In addition to writing, prior critical thinking skill and instructor significantly affected critical thinking performance, whereas other covariables such as gender, ethnicity, and age were not significant. With improved critical thinking skill, general education biology students will be better prepared to solve problems as engaged and productive citizens.

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., Heaney, P.J. 2006. Using conceptests to assess and improve student conceptual understanding in introductory geoscience courses. *Journal of Geoscience Education*. Vol. 54, 61-68.

Conceptests are higher-order multiple-choice questions that focus on one key concept of an instructor's major learning goals for a lesson. When coupled with student interaction through peer instruction, conceptests represent a rapid method of formative assessment of student understanding, require minimal changes to the instructional environment and introduce many of the recognized principles of effective teaching that enhance student learning. In this study, instructors from several different institutions developed over 300 conceptests for the geosciences. These instructors then used this suite of concept questions in a wide range of classroom settings, including large introductory general education Earth Science courses for non-majors at open enrollment institutions, smaller physical geology classes suitable for majors at private colleges, and in introductory geology laboratory settings. Results of pre- and post-class Geoscience Concept Inventory (GCI) testing and qualitative feedback from students and instructors showed that conceptests increased attendance, improved student satisfaction, and enhanced student achievement. Participating instructors found implementation of conceptests into their classes straightforward and required less than 30 minutes of preparation per class. The conceptest question database is available on-line for geoscience instructors.

Carlson, P.A., Berry, F.C., Voltmer, D. 2005. Incorporating student peer-review into an introduction to engineering design course. 35th ASEE/IEEE Frontiers in Education Conference, October, Indianapolis, IN.

We report on a project to improve the teaching of engineering design at the junior level. Peer review of student work is an integral part of collaborative learning and reform-driven engineering education. Yet successfully implementing this pedagogical technique requires significant amounts of instructor and class time. Furthermore, if adequate formative assessment does not emerge from peer review, the experience may devolve into "busy work" in the eyes of the student. Here, we give early results from an NSF-funded study using Calibrated Peer Review (a web-delivered, collaborative learning environment) to enhance learning in engineering design.

Fox, M.A., Hackerman, N. Editors, National Research Council. 2003. Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics. Free Executive Summary. <http://www.nap.edu/catalog/10024.html>.

Economic, academic, and social forces are causing undergraduate schools to start a fresh examination of teaching effectiveness. Administrators face the complex task of developing equitable, predictable ways to evaluate, encourage, and reward good teaching in science, math, engineering, and technology. Evaluating, and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics offers a vision for systematic evaluation of teaching practices and academic programs, with recommendations to the various stakeholders in higher education about how to achieve change. What is good undergraduate teaching? This book discusses how to evaluate undergraduate teaching of science, mathematics, engineering, and technology and what characterizes effective teaching in these fields. Why has it been difficult for colleges and universities to address the question of teaching effectiveness? The committee explores the implications of differences between the research and teaching cultures-and how practices in rewarding researchers could be transferred to the teaching enterprise. How should administrators approach the evaluation of individual faculty members? And how should evaluation results be used? The committee discusses methodologies, offers practical guidelines, and points out pitfalls. Evaluating, and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics provides a blueprint for institutions ready to build effective evaluation programs for teaching in science fields.

Zusho, A., Pintrich, P.R. 2003. Skill and will: The role of motivation and cognition in the learning of college chemistry. *Int. Jour. Science Education*. Vol. 25, 1081-1094.

This study investigated how students' level of motivation and use of specific cognitive and self-regulatory strategies changed over time, and how these motivational and cognitive components in turn predicted students' course performance in chemistry. Participants were 458 students enrolled in introductory college chemistry classes. Participants' motivation and strategy use were assessed at three time points over the course of one semester using self-report instruments. Results showed an overall decline in students' motivational levels over time. There was also a decline in students' use of rehearsal and elaboration strategies over time; students' use of organizational and self-regulatory strategies increased over time. These trends, however, were found to vary by students' achievement levels. In terms of the relations of motivation and cognition to achievement, the motivational components of self-efficacy and task value were found to be the best predictors of final course performance even after controlling for prior achievement.

Hall, S.R., Waitz, I., Brodeur, D.R., Soderholm, D.H., Nsr, R. 2002. Adoption of active learning in a lecture-based engineering class. 32nd ASEE/IEEE Frontiers in Education Conference. November, Boston MA.

Three years ago, the Department of Aeronautics and Astronautics at MIT expanded its repertoire of active learning strategies and assessment tools with the introduction of muddiest-point-in-the-lecture cards, electronic response systems, concept tests, peer coaching, course web pages, and web-based course evaluations. This paper focuses on the change process of integrating these active learning strategies into a traditional lecture-based multidisciplinary course, called Unified Engineering. The description of the evolution of active learning in Unified Engineering is intended to underscore the motivation and incentives required for bringing about

the change, and the support needed for sustaining and disseminating active learning approaches among the instructors.

Kardash, C.M., Wallace, M.L. 2001. The perceptions of science classes survey: What undergraduate science reform efforts really need to address. *Journal of Educational Psychology*. Vol. 93, 199-210.

Nine-hundred twenty two undergraduates completed an 80-item survey that assessed their perceptions of undergraduate science classes. Factor analysis of the items yielded 6 factors: (1) Pedagogical Strategies, (2) Faculty Interest in Teaching, (3) Student Interest and Perceived Competence in Science, (4) Passive Learning, (5) Grades as Feedback, and (6) Laboratory Experiences. Women differed significantly from men on the Pedagogical Strategies, Passive Learning, Grades as Feedback, and Laboratory Experiences factors. Correlational analyses and evidence from distinct groups supported the survey's construct validity. Students reported room for improvement of the science faculty's pedagogical practices. From the students' perspective, how information is taught appears to be at least as much of a concern as what information is taught.

Brickhouse, N.W., Dagher, Z.R., Letts, W.J., Shipman, H.L. 2000. Diversity of students' views about evidence, theory, and the interface between science and religion in an astronomy course. *Journal of Research in Science Teaching*. Vol. 37, 340-362.

Arguments for teaching about the nature of science have been made for several decades. The most recent science education policy documents continue to assert the need for students to understand the nature of science. However, little research actually explores how students develop these understandings in the context of a specific course. We examine the growth in students' understanding about the nature of astronomy in a one-semester college course. In addition to student work collected for 340 students in the course, we also interviewed focus students three times during the course. In this article we briefly describe class data and discuss in detail how five students developed their ideas throughout the course. In particular, we show the ways in which students respond to instruction with respect to the extent to which they (a) demand and examine evidence used for justifying claims, (b) integrate scientific and religious views, and (c) distinguish between scientific and nonscientific theories.

Caprariis, P.D., Barman, C., Magee, P. 2001. Monitoring the benefits of active learning exercise in introductory survey courses in science: An attempt to improve the education of prospective public school teachers. *The Journal of Scholarship of Teaching and Learning*. Vol. 1, 11.

Introduction of collaborative, active-learning exercises in a traditional lecture-based Environmental Geology course produced measurable changes in student learning. Oral surveys used as part of an assessment strategy suggest that students in the class use material from the exercises in responding to questions long after the subjects were covered in class. In addition, the variance of the grade distribution of the final examination suggests that learning is more uniform across the class than in previous semesters. Implementation of this approach is not limited to small classes; a single instructor can monitor a class of approximately 60 students as they work through the exercise.

Wright, J.C., Millar, S.B., Kosciuk, S.A., Penberthy, D.L., Williams, P.H., Wampold, B.E. 1998. A novel strategy for assessing the effects of curriculum reform on student competence. *Journal of Chemical Education*. Vol. 75. 986-993.

In this paper, we describe a new assessment strategy that was designed to determine whether such changes in student skills are observable by independent and unbiased observers. The methods were developed by representatives of the University of Wisconsin Chemistry faculty to assess reform success in ways that they would believe. The strategy is applicable to course comparisons that are often found in curriculum reform projects where the project design does not allow the controlled course settings that are sought for educational research. Two sections of a large analytical chemistry course for first-year undergraduates were assessed. One section was taught using methods that focused on lectures that carefully led the student to mastery of the course material using methods that encouraged student questions and participation. The other section was taught using cooperative learning methods that emphasized group work and self-discovery (12). These sections are labeled responsive lecturing (RL) and structured active learning (SAL). The SAL approach had been developed in 1992 and refined during the subsequent spring semesters. Unbiased external evaluation judged that both sections represented best practice for each method.

House, J.D. 1994. Student motivation and achievement in college chemistry. *Int'l J. of Instructional Media*. Vol. 21, 1-6.

The initial assessment of learner characteristics is an important component of instructional design models; however, limited attention has been given to the importance of students' initial motivation. The purpose of this study was to investigate the relationship between students' achievement expectancies and academic self-concept and their subsequent achievement in college chemistry. There were three main findings from this study. First, several specific learner characteristics were significant predictors of achievement. Second, students' achievement expectancies and academic self-concept were more significant predictors of performance than were students' prior achievement and their prior instructional experience. Finally, prior achievement was the only variable to significantly enter the regression equation for predicting students' earning the highest grade possible vs. earning lower grades. These findings suggest that students' academic self-concept and achievement expectancies are significant predictors of overall grade performance in chemistry while prior achievement was the only significant predictor of high grade performance.

Novick, L.R., Holyoak, K.J. 1991. Mathematical problem solving by analogy. *J. of Experimental Psychology*. Vol. 17, 398-415.

We report the results of 2 experiments and a verbal protocol study examining the component processes of solving mathematical word problems by analogy. College students first studied a problem and its solution, which provided a potential source for analogical transfer. Then they attempted to solve several analogous problems. For some problems, subjects received one of a variety of hints designed to reduce or eliminate the difficulty of some of the major processes hypothesized to be involved in analogical transfer. Our studies yielded 4 major findings. First, the process of *mapping* the features of the source and target problems and the process of *adapting* the source solution procedure for use in solving the target problem were clearly distinguished: (a) Successful mapping was found to be insufficient for successful transfer and (b)

adaptation was found to be a major source of transfer difficulty. Second, we obtained direct evidence that schema induction is a natural consequence of analogical transfer. The schema was found to co-exist with the problems from which it was induced, and both the schema and the individual problems facilitated later transfer. Third, for our multiple-solution problems, the relation between analogical transfer and solution accuracy was mediated by the degree of time pressure exerted for the test problems. Finally, mathematical expertise was a significant predictor of analogical transfer, but general analogical reasoning ability was not. The implications of the results for models of analogical transfer and for instruction were considered.