
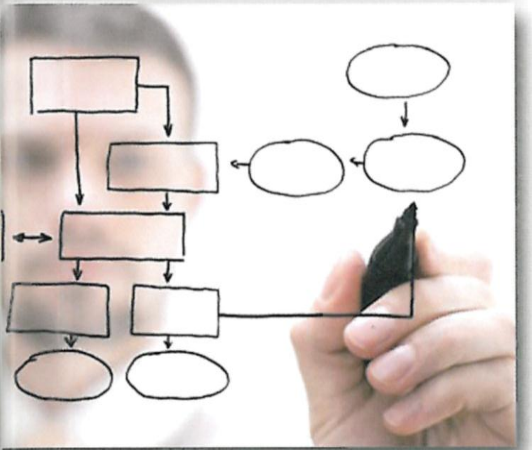

Assessment in the Disciplines
Volume 3

Assessment in Engineering Programs: Evolving Best Practices

Edited by
William E. Kelly

Sponsored by

AIR Association for Institutional Research

American Society for Engineering Education
ASEE

CHAPTER 1

EVOLUTION OF ENGINEERING ASSESSMENT

William E. Kelly
American Society for Engineering Education
Washington, DC

Introduction

The vision of the engineer of 2020 presented by the National Academy of Engineering (NAE) in their 2004 report goes well beyond the ABET outcomes, as it should. One of the keys, according to the authors of the NAE report, is life-long learning—one of the current ABET outcomes and one that has been particularly difficult to assess and evaluate. The NAE authors state that "...to be individually/personally successful, the engineer of 2020 will learn continuously throughout his or her career, not just about engineering but also about history, politics, business, and so forth."

Assessment is increasingly focusing on student learning, including the ability of students to assess their own learning. Ultimately, professionals must have the ability and the motivation to assess where there are gaps in their knowledge and skills and to develop, implement, and evaluate appropriate learning strategies to address those gaps.

An overall goal of accreditation is to assure minimum levels of quality in programs and to promote continuous quality improvement in programs. The preface to the current ABET Engineering Criteria (ABET, 2007) states that

> These criteria are intended to assure quality and to foster the systematic pursuit of improvement in the quality of engineering education that satisfies the needs of constituencies in a dynamic and competitive environment. (p. 1)

The current ABET criteria are not prescriptive as to what students should learn, but they do require a process of goal setting, evaluation of achievement of goals, assessment of outcomes, and quality improvement. As a minimum, each program must assess all of the outcomes listed in ABET Criterion 3 (a) through (k) or also known as 3(a–k) (ABET, 2007).

ABET recognizes that to realize its vision, it must be a leader in promoting assessment and continuous improvement. To this end, ABET recently completed a longitudinal study of the impact of EC2000 on engineering education and published *Engineering Change: A Study of the Impact of EC2000* that documents the positive impact that the new engineering criteria already are having on graduates (ABET, 2006).

The ABET longitudinal study also documents the broad support that the new criteria and the philosophy of assessment and improvement have in the engineering education community. The issue for the engineering education community now is how best to do assessment and how to do it effectively and efficiently.

ABET and others recognize that the processes for assessment and improvement must be sustainable for a range of educational institutions. Therefore, assessment must not be too burdensome for faculty and also must be perceived by faculty and administrators as adding value to their programs. It also is reasonable for faculty members to expect that good assessment practices will help them improve their teaching effectiveness and ultimately, to improve student learning.

ABET provides the criteria, but it is less able to define what constitutes good practice for assessment and improvement. It does, however, provide guidance through activities such as workshops for faculty on engineering assessment practice (ABET, 2008a). It also provides guidance on assessment planning through a section on the ABET web site maintained by Gloria Rogers.

Prados, Peterson, and Lattuca (2005) provide an overview of the initial impact of EC2000 as well as a comprehensive discussion of the factors that led to the changes in the accreditation criteria and accreditation processes. They also outline some of the challenges that ABET has experienced in transitioning to EC2000 and sustaining the change. ABET accredits programs; this process provides assurance to customers—students, parents, and employers—that graduates of each program have achieved the minimum competencies in the specified program field.

There were a number of drivers to change engineering accreditation that converged in the late 1980s, including industry concerns that graduates' competencies might not be preparing them for the new global economy. Employers expressed doubts that the strong technical skills that were the norm as a result of changes made to engineering curricula in the 1950s were still sufficient, if they had ever been. There needed to be more attention in engineering programs to the "soft skills" such as communications and team work. There were also strong feelings, notably among engineering deans, that the ABET processes had become too burdensome and could even be a barrier to needed innovation in engineering education.

ABET responded with both new, simplified criteria and a new performance or outcomes-based approach to accreditation. Instead of providing the recipe for engineering education, the approach is to focus on outcomes with ABET providing a minimum set in the Criteria. ABET also accepted the challenge to change its own processes, including training an entire new cadre of program evaluators and team chairs, and, finally, assessing the impact of EC2000 on engineering education.

There have been numerous sessions and papers at American Society for Engineering Education (ASEE) regional and national meetings dealing with assessment. ASEE started early with its white paper on assessment (American Society for Engineering Education, 1996). A search of the ASEE 2007 annual conference web site using *assessment* as the key word turned up over 125 papers with the word *assessment* in the paper title. This compares with nine papers in 1996, the earliest year for which papers are available on the ASEE web site. At the 2007 annual conference, there were also nine sessions with the word *outcomes* in the session title.

In 2005, Olds, Moskall, and Miller published a review of the state of assessment in engineering education as reflected in articles published in the ASEE *Journal of Engineering Education* (JEE). Assessment was reviewed in its role supporting engineering education research rather than as it is commonly applied by engineering faculty to assess achievement of the ABET Criterion 3(a–k) outcomes. Although most programs are primarily interested in processes to assess student achievement of outcomes, some are also interested in assessing the impact of program changes on student learning and also on outcomes such as student retention. Assessment design is divided into descriptive and experimental approaches, and examples from recent JEE articles are given and briefly discussed and summarized in two tables. Even programs only interested in assessing the ABET outcomes should find this paper a useful summary of where we are. Longer term, when programs look at evaluating achievement of program objectives, including the assessment of program changes on achievement of objectives, they will find their perspective on assessment to be more aligned with assessment for research.

So far, there is has been much less published on program quality assurance and improvement. There has been some experimentation by the International Organization for Standardization (ISO) with ISO 9001 and with Malcolm Baldridge in higher education, but there appear to have been few or no recent attempts to apply either to engineering programs in the U.S. On her "Assessment Planning" web site, Gloria Rogers notes that for industry and for education, it is difficult to separate assessment and quality assurance and, ultimately, quality improvement (Rogers, 2008).

At the ASEE annual meeting in 2007, there were only six papers with *quality* in the title that appeared to deal with program quality improvement. The paper by Prados et al. previously mentioned is one of the few recent JEE articles that discusses quality improvement. This is likely to change as the new general criteria now have a specific criterion dealing with quality improvement.

One of the issues that engineering programs must deal with is workload; thus, faculty members and program administrators want assessment methods that are both effective and efficient and, of course, acceptable to ABET. There is also the issue of data collection and processing, and this is where institutional-wide support services such as institutional research offices can and are helping. Universities have a need for, and an interest in, defining and managing assessment and improvement processes that will serve a wide variety of accrediting agencies and institutional needs.

In 2004, ABET looked at some of the potential barriers to full implementation of EC2000. One of the conclusions was that

> Workload, documentation, and assessment tools continue to provide frustration for constituents. Sustaining the change relies on sustaining the level of commitment and enthusiasm—the level of momentum—both on campus and at ABET. We understand this and are working to continually improve it.

Standards and assessment are a fact of life for elementary and middle school teachers and increasingly for high school teachers as states set and enforce learning standards. One purpose of standards is to document and promote best practice. The Joint Committee on Standards for Educational Evaluation (JCSEE) has published *The Student Evaluation Standards* (2003). Also, the ISO published a working agreement on applying ISO 9001 to higher education that could be useful to programs in organizing their improvement processes (Kelly, 2007). At this point, there are no international standards for assessment in higher education.

In the remainder of this chapter, I will outline some of what has been reported about assessment and improvement in the Association for Institutional Research (AIR) business and mathematics volumes and then provide an overview of this volume.

Previous Association for Institutional Research (AIR) volumes

AIR is the professional development organization that supports institutional research efforts in postsecondary education. Campus institutional research offices collect and manage data, which are used for a variety of campus planning and management activities. Increasingly, institutional research offices provide support, particularly for data collection and management, for campus-wide and program-level assessment activities. AIR instituted the series "Assessment in the Disciplines" specifically to support institutional researchers and faculty in campus assessment activities (Association for Institutional Research, 2008).

AIR has published three volumes on assessment practice thus far. The first two volumes focus on assessment in business schools, and the third volume deals with assessment in mathematics. This engineering volume is the fourth in the series.

Business schools are accredited at the school or college level in contrast to engineering programs, which are accredited at the program level. However, accreditation of general engineering programs at schools where there are multiple tracks in engineering probably comes close to business schools accreditation.

In 2003, the Association to Advance Collegiate Schools of Business (AACSB) approved and began to implement new standards for accrediting business schools (Association to Advance Collegiate Schools of Business, 2008). These standards require business schools to provide direct evidence of student progress in meeting learning goals (Trapnell, 2005). The operative words here are "direct evidence."

Martell and Calderon (2005) present what is intended to be a primer on assessment for business schools. In their introduction to the first AIR business volume, they provide a summary of what business schools are doing with assessment today. Some of their comments should resonate with engineering faculty and administrators. For example, they advise business deans to focus on direct assessment methods for assessing student learning; they note that surveys have their place in academic planning and management but not as evidence of student learning. They also advise deans to keep things simple. They point out that good program assessment does not have to meet the standards of academic rigor expected for peer-reviewed publication, but it does have to be effective—the judges of effectiveness ultimately being the users and the accrediting team.

Since the AACSB standards are relatively new, it is to be expected that business schools would be and are struggling with how to do direct assessment and provide appropriate evidence, a situation not too different from what engineering programs are dealing with. Martell and Calderon note that requirements for assessment data for business schools are consistent with those of regional and other professional accreditation bodies (e.g., ABET) and some state legislatures. There are many good examples in the two business volumes that should be useful to engineering faculty members and administrators charged with organizing assessment and improvement processes at the school and college level in engineering.

Anyone familiar with the evolution of engineering accreditation over the last ten years or so would find the discussion in business familiar. Martell and Calderon's comment that the AACSB requirements for assessment data are consistent with regional and other professional accrediting bodies and some state legislatures suggests, as noted earlier, that universities will increasingly define best assessment practices to demonstrate student learning for multiple audiences.

Mathematics knowledge and skills are extremely important in engineering. The ABET curricula requirement for mathematics and science is 32 credits, and a number of the program criteria imply a high level of mathematical performance for graduates. Assessment of mathematics readiness—related to performance—is also increasingly recognized for its importance in advising students studying, and potentially interested in studying, engineering. According to Adelman (1998), "the highest level of mathematics studied in secondary school is strongly correlated with bachelor's completion in any field." This is particularly true with respect to persistence and success in engineering.

Mathematics and the AIR mathematics volume are different from business and engineering in that mathematics programs themselves are not separately accredited. However, mathematics is an important part of all engineering programs, and thus there would be expected to be common assessment issues. Several of the papers in the mathematics volume are directly applicable to engineering. Also, the Mathematical Association of America (MAA) is actively supporting assessment with its "Supporting Assessment in Undergraduate Mathematics" (SAUM) program and has case histories available online (Mathematical Association of America, 2006). A relevant case history is one describing assessment of the core mathematics program at West Point. All cadets take the same four-course sequence in mathematics where the first course is in discrete dynamical systems with an introduction to calculus (Heidenberg & Huber, 2006).

Assessment in Engineering

ABET is a good source of assessment materials and a resource to check frequently is the ABET web page (Rogers, 2008). For programs undergoing a review there is no substitute for participation in the ABET annual meetings. The fall ABET meeting is now largely devoted to assisting programs in preparing self-study reports and the summer EAC meeting for deans provides the latest information on practices in place for the current visit cycle (e.g., what are the latest issues, how evaluators are looking at materials, how to present materials).

There is no substitute for experience, and all engineering administrators should consider volunteering as an ABET evaluator (ABET, 2008c). One of the keys to ABET's past and future success is having a cadre of dedicated and effective evaluators. ABET is focusing its current improvement efforts on improving the performance of program evaluators (ABET, 2008d). ABET's strategy has been not only to change the criteria, now essentially common for all of the commissions, but to improve the accreditation process itself and, for the long term, to contribute to sustaining improvement of engineering education broadly.

ABET has a long history of conducting faculty workshops, first to introduce faculty to EC2000 and currently to provide information on ways to do assessment. Online webinars are now available.

Rose-Hulman Institute of Technology has been holding an annual symposium on assessment since 1997. The 2008 symposium was entitled "Best Assessment Processes," which suggests that assessment for the engineering community has evolved to the point that good assessment practices can be specified.

The ASEE annual and regional meetings are excellent places to find out how faculty members in engineering and engineering technology are actually doing assessment on campus. As mentioned earlier, the ASEE white paper on assessment is still a good resource, and the ASEE *Journal on Education* is another resource. However, as suggested by Martell and Calderon, the key is direct, effective, and simple assessment.

What void can this AIR volume fill? Clearly there is no shortage of materials on assessment. In soliciting papers for this volume, the editor sought coverage of a range of what faculty and programs are actually doing that appeared to be successful. Past, and then current, ABET EAC members were asked to suggest assessment activities that should be highlighted in the volume.

Engineering Assessment

The ABET longitudinal study is unique in the assessment literature. A summary of the study is available on the ABET web site, and a copy of the complete report can be purchased from ABET (ABET, 2006). The full report is over 400 pages long, and there is much that could be gleaned from the report about what seems to be working for programs with respect to assessment and improvement. Two of the chapters in this volume analyze what was learned about the impact of EC2000 and suggest ways that this information can be used to improve program- and campus-level assessment.

In Chapter 2, Volkwein et al. provide an overview of the *Engineering Change* study. What impact is EC2000 having on programs and, ultimately, on student learning? The criteria do not affect student learning directly but only indirectly through changes that programs make to satisfy the ABET criteria. Engineering assessment is focused specifically on the Criterion 3(a–k) outcomes which effectively define student learning for engineering programs.

For student experiences and learning outcomes, results are reported in Chapter 2 as changes in programs, changes in student experiences, and changes in student learning outcomes. Improvements in student learning outcomes as measured by surveys of graduates from 1994 and 2004 indicate slightly higher competencies based on graduates' self assessment for all of the Criterion 3(a–k) outcomes. Figure 1 is replotted from data in Figures 7, 8, and 9 in Chapter 2.

Nine competencies were defined that relate to the 11 ABET outcomes. In Figure 1, *design and problem-solving skills* (Des&Form) combine ABET outcomes (c) and (e) and *societal and global issues (*Impact&Know*)* combine ABET outcomes (h) and (j). See Chapter 2 for more details on the factor analysis that is the basis for this reduction.

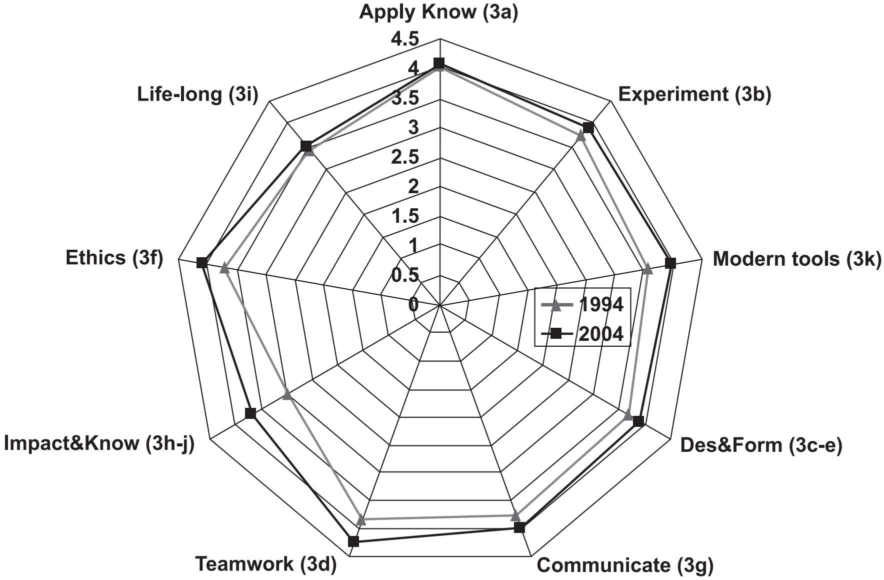

Figure 1. Comparison of graduates' self assessments for outcomes (a) through (k).

Interestingly, graduates see the most improvement in their competency with respect to societal and global issues (Impact&Know). With respect to learning, Volkwein et al. conclude that overall, EC2000 graduates (2004) are slightly better prepared than pre-EC2000 (1994) graduates. In Figure 1, it is clear that graduates improved in all categories except Criterion 3(a) where they remained at the same level. Overall, soft skills have improved but not at the expense of technical preparation.

The most prescriptive specification of what engineering graduates should know is embodied in the Fundamentals of Engineering (FE) examination that many engineering schools are reportedly using for outcomes assessment. Although none of the papers in this volume deal specifically with the use of the FE exam for outcomes assessment, Chapter 7 by Estes et al. outlines the role it plays in assessment at West Point, and Briedis comments on using the FE for assessment in Chapter 8. In a recent paper, Lawson (2007) reviews some of the problems in using the FE for assessment but notes that it will likely continue to be used and that users should just be aware that good assessment practice always involves multiple measures.

Standards and assessment are a fact of life for elementary and middle school teachers and increasingly for high school teachers as states set learning standards. There has also been an effort by the Joint Committee on Standards for Educational Evaluation (JCSEE) "to develop standards to help ensure ethical, useful, feasible, and accurate evaluation of students." The Joint Committee is accredited by the American National Standards Institute (ANSI) to develop standards and the "Student Evaluation Standards" are an American National Standard (ANS). Chapter 3 by Gullickson and Gullickson shows how these standards can be related to the ABET requirements for assessment and evaluation. A case history is presented to illustrate the use of the standards in engineering assessment.

Not all of what students learn is learned or expected to be learned in the classroom, and student competitions can be an integral part of a program's strategy to demonstrate achievement of the ABET outcomes. There are a number of competitions that have the potential to contribute to the overall educational experience, and programs recognize this. In Figure 6 in Chapter 2, it can be seen that this is the area where the out-of-class experiences increased the most between 1994 and 2004 graduates. Chapter 4 by Waldman et al. describes the United States Environmental Protection Agency (EPA) P3 program and what participation in the program can contribute to student learning for engineering programs (2007).

Although the P3 competition is open to all students, according to Waldman et al., the majority of students have been from engineering schools. Educating the next generation of engineers and scientists in sustainability is important to EPA, and there is a need to assess how programs like P3 can contribute to this. Waldman et al. describe the P3 competition as an example of a "significant learning" experience as defined by Fink (2003). Fink's taxonomy includes foundational knowledge (hard sectors) but also human dimensions (soft sectors). Fink advocates spiral rather than linear development, a concept touched on by Robson et al. in Chapter 10 in this volume.

Although cooperative education is no longer separately accredited by ABET, it is still a distinctive part of many engineering programs. ABET's approach with EC2000 was that the outcomes for a program with a cooperative education component should be the same as for traditional programs, and therefore no special accommodation was needed. On the other hand, there are unique learning and assessment opportunities, and Chapter 5 deals with assessment of a cooperative education program. In Chapter 5, Mozrall et al. describe the system at the Rochester Institute of Technology (RIT) for collecting and analyzing data for cooperative education. RIT is a leader in cooperative education, and all of its eight colleges are involved. The engineering programs at RIT are five years in length with a required co-op experience after the first two years.

For RIT, EC2000 provided the incentive to look carefully at how they assessed co-op, which prior to EC2000 had been evaluated with a traditional open-ended question process. Faculty and administrators at the college and institution level recognized that their assessment processes should evolve to take advantage of the unique opportunities that co-op offers to gather meaningful data on achievement of the ABET Criterion 3 outcomes and also for program improvement. As a result, assessment processes evolved from paper and pencil to web-based built around Criterion 3(a–k) outcomes. The general survey is generic with respect to (a–k) so that it can be used by all engineering programs; programs can tailor the survey to meet specific program criteria outcomes as appropriate. The process has evolved so that information can be gathered on student achievement as assessed by students and also on students' perception of the opportunities employers provide to demonstrate achievement of (a–k).

An example of an actual student evaluation is included, indicating how both quantitative and qualitative data are gathered. Performance metrics have been developed and are used to track student achievement. One of the advantages described by the authors is that the system allows faculty to add material to the curriculum (e.g., specific engineering tools) and actually see the improvement in student co-op experiences as assessed by both students and employers before the students graduate.

The system at RIT has evolved into a rich source of data for both program assessment and improvement. The clear institutional support for co-op, continuous improvement of the co-op program, and the fact that co-op is mandatory for students and employers appear to be important factors in the success of the new assessment system.

Chapter 6 by Charles Malmborg describes a flexible, data-driven tool for program-level assessment developed and used by the industrial and management engineering (IME) program at Rensselaer Polytechnic Institute (RPI). The tool is described as a decision-support system that allows the program's accreditation system to be used to examine different strategies for improvement with respect to the 12 program outcomes. Five different measures including portfolios are used, and results are recorded in database systems that are available online. The decision-support system allows the chair and the faculty assessment committee convenient access to the data to study trends, achievement of performance goals and so on. This is an example of a program applying its own engineering expertise to designing and developing an engineering system that best meets the needs of its customers—in this case, themselves.

Although grades per se are not generally accepted as assessment, all faculty members know intuitively that assessment and grading are closely related. Thus, the question is how to integrate assessment with the requirement to evaluate student work and performance. Embedded assessment is one way, and Chapter 7 by Estes et al. describes how the existing grading system at West Point is used to support direct assessment.

Assessment is part of the institutional culture at West Point, and the program assessment required for ABET fits into an institutional-wide framework. The program-level efforts are designed to take advantage of the institutional-level data collection system. Data on civil engineering graduates can be extracted from data for all graduates and used to evaluate program objectives. Web-based end-of-course surveys are used in all courses allowing comparison of performance across all USMA programs.

Objective 3 for all graduates of West Point is that "Graduates communicate effectively." Based on survey results, the civil engineering program meets its objective to ensure graduates communicate effectively (4/5) with 5 high and 4 indicating the objective is met, although there is room for improvement. This program objective directly relates to ABET Outcome 3(g) and civil engineering program Outcomes 9 and 10 that graduates will write and speak effectively.

For all outcomes, USMA CE faculty members rate specific courses for their contribution to the outcome with (5) being a "very large contribution." Practitioners typically participate in evaluating capstone and independent study projects and provide outside input on Outcomes 9 and 10 (communications).

Estes et al. describe a process for assessing performance for each outcome that combines senior survey data at the program and institutional level. There is also a capstone design assessment where grading yields a measure of student performance for all outcomes. A measure of coverage of each outcome is also obtained for the capstone; low coverage means that the primary assessment must be obtained somewhere other than in the capstone course.

The concept behind embedded assessment is to take advantage of what faculty are already doing to improve student learning and use these efforts to collect direct evidence and demonstrate achievement of ABET outcomes. Chapter 8 by Daina Briedis describes how a large research university is using embedded assessment to improve the effectiveness of its assessment and, at the same time, manage faculty workload.

Two reasons given as to why grading generally does not work for outcomes assessment is that it is norm-based and typically integrates more than one outcome. Briedis makes the case that course grading schemes can be designed to accomplish some assessment. To do this, student performance metrics must be developed, for example, for written communication skills. Typically, a rubric is prepared and may even be shared with students. Faculty member can then assess student performance and make a judgment as to performance. The results of this assessment can then be used in the grading scheme to calculate grades which are typically norm-based—students are ranked relative to others in the class. When used in a capstone course with multiple evaluators, some calibration of the performance metrics is possible; then faculty can set minimum performance levels. Once developed, rubrics for outcomes such as communication skills can be used throughout a program.

Assessment at the Colorado School of Mines is moving into its third generation, and Barbara Olds provides a retrospective in Chapter 9. Although longitudinal portfolios provided good information, their use was just not sustainable. However, the data collected with portfolios led to changes in the way communications skills are developed on campus.

In 2004, Virginia Tech renamed its Engineering Fundamentals department the Department of Engineering Education and expanded its mission to include research. One of the primary missions of the new department is to carry out "rigorous research in the area of engineering education." Chapter 10 by Robson et al. focuses on assessment of a two-credit introduction to engineering course taken by all engineering and computer science students in the first semester. Approximately 1,500 students take this course each fall, and it provides a unique opportunity to study questions such as what factors affect retention.

Robson et al. discuss assessment of the factors affecting success in engineering defined by cumulative grade-point average after four semesters and retention, where students are considered "retained" if they are still enrolled in engineering after the 4th semester. The authors describe five primary assessment measures ranging from a new student survey to focus groups. This chapter is an example of how assessment methodology can be used beyond ABET (a–k) for program improvement and to study important issues such as factors contributing to student success and retention. These are examples of the kinds of questions that would benefit from rigorous research.

Are the ABET Criterion 3(a–k) outcomes valid? In Chapter 11, Volkwein et al. provide an analysis of the results of the employer survey conducted for the *Engineering Change* study. With respect to Criterion 3(a–k), Volkwein et al. found that for the employers surveyed, the ABET outcomes are important. Their top ranking goes to communications Criterion 3(g) with 91% rating communications highly important or essential and 8% rating it moderately important for a combined 98%. Problem-solving skills also rate very high. The outcomes rated lowest, at least for new graduates as the question was asked, are outcomes (j) and (k) dealing with contemporary issues and the social context of engineering.

The Grinter Report in 1954 largely defined engineering curricula as they are delivered today. Specifically, the report recommended curricular requirements for mathematics and science and engineering science that are essentially the same as the requirements in EC2000. The report was reprinted in the *Journal of Engineering Education* and is available online (Grinter Report, 1954).

After the engineering colleges reviewed a draft of the Grinter Report, the committee invited industry input. Industry responded and surveyed "smaller organizations at the operating and manufacturing level"(where I would expect to find recent graduates) and found:

> The returns indicated no criticism of the technical competence of engineers but raised questions concerning (1) the adequacy of their background in basic science, and humanistic fields and (2) concerning their capacity for effective communication. (p. 85)

In their appendix for Chapter 11, Volkwein et al. report the employers' ratings of the importance of each of the 11 EC2000 outcomes. In Figure 2, I have added the employers' ratings of the importance of the outcomes to the graduates' self assessment for 1994 and 2004. To allow direct comparison, I have used the geometric average of the employers' rating of Criterion (h–j) and (c–e) as shown.

In Figure 3, I have plotted the gap between employers' ratings of the importance of the outcomes and the 2004 graduates' self assessments at graduation. Again, I have used the geometric average for outcomes (h) and (j) and (c) and (e) to allow direct comparison. The greatest gap is in the ability of the graduates to communicate (g).

It is interesting, and probably significant, that fifty years after the Grinter Report, employers still feel that communications skills are very important and that there is room for improvement. Figures 2 and 3 suggest that although there has been improvement with EC2000, there is still a real or at least perceived gap for employers.

Communications and problem-solving skills are important skills for all university graduates, not just for engineers. Boyer (1987) stated that "The foundation for a successful undergraduate experience is proficiency in the written and spoken word." A recent study entitled *Reinventing Undergraduate Education: Three Years After the Boyer Report*, reports the results of a 2001 survey of 123 research universities review of the status of undergraduate education (Boyer Commission, 2003). The data suggest that

> ... professional programs, such as business and engineering, are outstripping the arts and sciences departments in important areas such as written and oral communications.

Specifically, with respect to writing skills, the conclusion is that

> writing skills are a priority; course requirements are increasing. But writing is often taught in ways that diminish its importance in the eyes of students. The courses are often taught by teaching assistants and adjuncts, not professors. Furthermore, if professors do not require extensive written work in their majors, students will not think writing skills matter for their professional life. Students too often feel that passing the writing course is the goal; they do not always understand that the ability to write well is a survival skill.

Certainly more can be done in engineering, but the *Engineering Change* data suggests that progress is being made. Moreover, the literature suggests that engineering and business programs are doing a relatively good job!

Some Closing Thoughts

Assessment and improvement methodologies for engineering education are evolving. The ABET Criteria have provided a driver and a motivator for assessment and improvement, and there are many other incentives and pressures to improve engineering education. Real success with assessment and improvement will only come when these processes have become part of the culture of engineering education.

The goal of engineering education is student learning, and good assessment and teaching are tools to achieve this goal. In this volume, we focus on the assessment component of the learning system. It should be clear from reading the chapters in this volume that there are many ways to design an assessment system to meet specific program needs and constraints; the key, however, is faculty engagement.

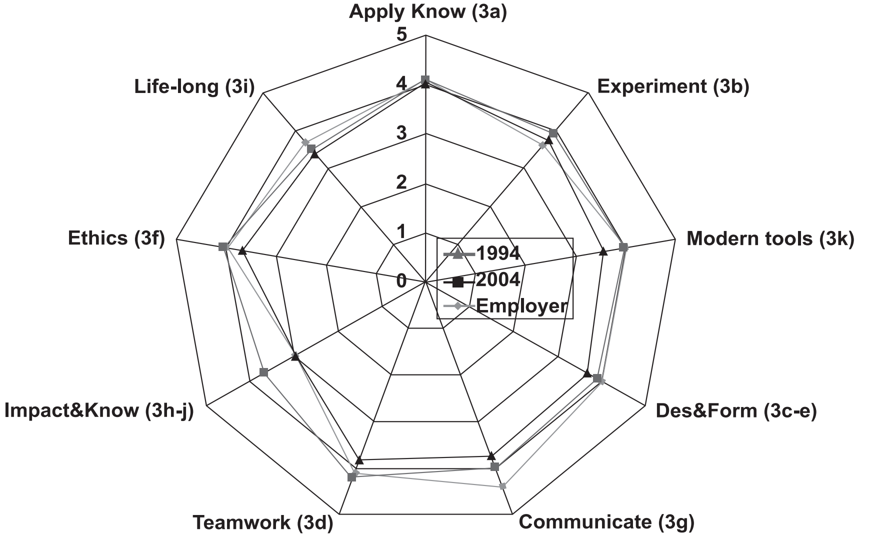

Figure 2. Graduates' self assessments of engineering skills and employers' rating of importance.

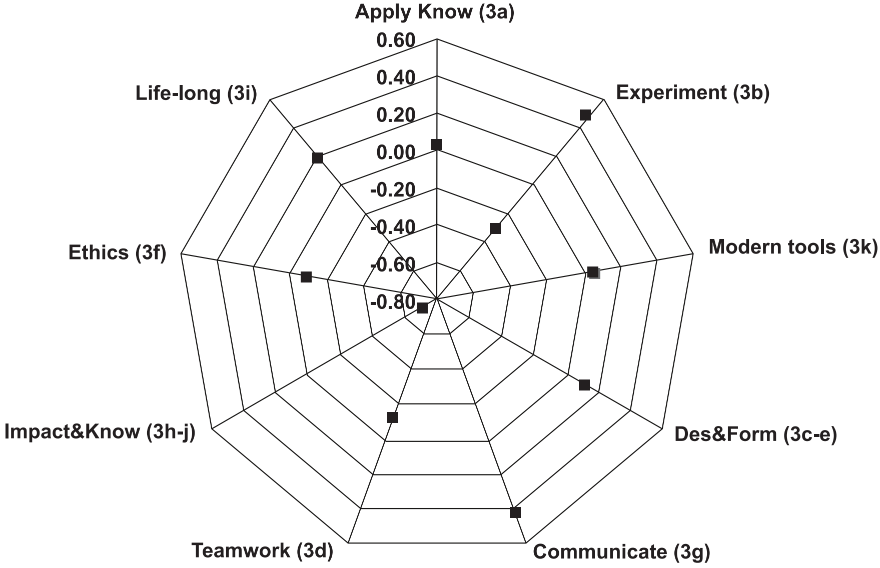

Figure 3. Gap between employers' rating of importance and 2004 graduates' assessment of achievement.

References

ABET. (2004). *Sustaining the change, a follow-up report to the Vision for Change*. Retrieved April 25, 2008, from http://www.abet.org/Linked%20Documents-UPDATE/White%20Papers/Sustaining%20the%20Change-Web.pdf

ABET. (2006). *Engineering change: A study of the impact of EC 2000*. Retrieved April 25, 2008, from http://www.abet.org/Linked%20Documents-UPDATE/White%20 Papers/Engineering%20Change.pdf

ABET. (2007). *Criteria for accrediting engineering programs effective for evaluations during the 2007–2008 cycle*. Retrieved April 25, 2008, from http://www.abet.org/Linked%20Documents-UPDATE/Criteria%20and%20PP/E001%2008-09%20 EAC%20Criteria%2011-30-07.pdf

ABET. (2008a). *Faculty workshop on assessing outcomes*. Retrieved April 25, 2008, from http://www.abet.org/workshop.shtml

ABET. (2008b). *Assessment planning with Gloria Rogers, Ph.D*. Retrieved April 25, 2008, from http://www.abet.org/assessment.shtml

ABET. (2008c.). *How do I get involved*? Retrieved April 25, 2008, from http://www.abet.org/volunteer.shtml

ABET. (2008d). *Partnership to advance volunteer excellence (PAVE)*. Retrieved April 25, 2008, from http://www.abet.org/pave.shtml

Adelman, C. (1998). *Women and men of the engineering path: A model for analysis of undergraduate careers*. Washington, DC: U.S. Government Printing Office. http://www.ed.gov/offices/OERI/PLLI/pubs-postsec-engineering.html . Not able to retrieve actual report but see the following for extracts of report and tables http://nsf.gov/statistics/nsf99338/c3/tt03-08.htm and http://nsf.gov/statistics/nsf99338/c3/tt03-09.htm

American Society for Engineering Education. (1996). *Assessment white paper*. Retrieved April 25, 2008, from http://www.asee.org/resources/beyond/assessmentwhitepaper.cfm

Association for Institutional Research. (2008). *Assessment in the disciplines*. Retrieved April 25, 2008, from http://www.airweb.org/?page=204

Association to Advance Collegiate Schools of Business. (2008). *Accreditation*. Retrieved April 25, 2008, from http://www.aacsb.edu/accreditation/

Boyer Commission. (2001). *The Boyer Commission on educating undergraduates in the research university, reinventing undergraduate education: Three years after the Boyer Report*. Retrieved April 25, 2008, from http://www.sunysb.edu/pres/pdfs/0210066-Boyer%20Report%20Final.pdf

Boyer, E. L. (1987). *College: The undergraduate experience in America*. New York: Harper & Row.

Fink, L. D. (2003). *Creating significant learning experiences: An integrated approach to designing college course*. New York: Jossey-Bass.

Grinter Report. (1954). Retrieved April 25, 2008, from http://www.asee.org/resources/upload/The-Grinter-Report-PDF.pdf

Heidenberg, A., & Huber, M. (2006). *Assessing the use of technology and using technology to assess*. Retrieved April 25, 2008, from http://www.maa.org/saum/cases/heidenberg-huber1105-saum.pdf

Joint Committee on Standards for Educational Evaluation. (2003). *The student evaluation standards: How to improve evaluation of students*. Thousand Oaks, CA: Corwin Press.

Kelly, W. E. (2007). Certification and accreditation in civil engineering, *Journal of Professional Issues in Engineering Education and Practice*, *133,* 181–187.

Lattuca, L. R., Terenzini, P. T., & Volkwein, J. F. (2006). *Engineering change: A study of the impact of EC2000. Final Report*. Baltimore, MD: ABET, Inc.

Lawson, W. E. (2007). Reliability and validity of FE exam scores for assessment of individual competence, program accreditation, and college performance. *Journal of Professional Issues in Engineering Education and Practice*, *133,* 320–326.

Martell, K., & Calderon, T. G. (Eds.). (2005). *Assessment of student learning in business schools: Best practices each step of the way.* Tallahassee, FL: Association for Institutional Research.

Mathematical Association of America. (2006). *Supporting assessment in undergraduate mathematics (SAUM)*. Retrieved April 25, 2008, from http://www.maa.org/saum/

National Academy of Engineering. (2004). *The engineer of 2020: Visions of engineering in the new century*. Washington, DC: National Academy Press.

Olds, B. M., Moskall, B. M., & Miller, R. C. (2005). Assessment in engineering education: Evolution, approaches and future collaborations. *Journal of Engineering Education*, *95*(1), 13–25.

Prados, J. W., Peterson, G. D., & Lattuca, L. R. (2005). Quality assurance of engineering education through accreditation: The impact of Engineering Criteria 2000 and its global influence. *Journal of Engineering Education*, *95*(1), 165–184.

Rogers, G. (2008). Assessment planning with Gloria Rogers, Ph.D. Retrieved April 25, 2008, from http://www.abet.org/assessment.shtml

Trapnell, J. E. (2005). Foreword. In K. Martell & T. G. Calderon (Eds.), *Assessment of student learning in business schools: Best practices each step of the way* (pp. iii–iv). Tallahassee, FL: Association for Institutional Research.

United States Environmental Protection Agency. (2007). *EPA P3: People, prosperity and the planet student design competition for sustainability*. Retrieved April 25, 2008, from http://es.epa.gov/ncer/p3/index.html

CHAPTER 2

MEASURING THE IMPACT OF ENGINEERING ACCREDITATION ON STUDENT EXPERIENCES AND LEARNING OUTCOMES[1]

J. Fredericks Volkwein
Lisa R. Lattuca
Patrick T. Terenzini
Center for the Study of Higher Education
Pennsylvania State University

Introduction

Throughout the history of the accreditation process, accreditors have responded to changing contexts and pressures from inside and outside the academy by modifying their processes. For example, in response to the increasing cost burden associated with regional and professional accreditation reviews, agencies have encouraged institutions to embed these reviews in ongoing institutional processes such as strategic planning or program review. Acknowledging the growing consensus that student learning outcomes are the ultimate test of the quality of academic programs, accreditors have also refocused their criteria, reducing the emphasis on quantitative measures of inputs and resources and requiring judgments of educational effectiveness from measurable outcomes (Volkwein, Lattuca, Caffrey, & Reindl, 2003). The Council for Higher Education Accreditation (CHEA), which recognizes individual accreditation agencies, endorses assessment of student learning outcomes as one dimension of accreditation: "Students, parents, and the public ... want to know what the learning gained in these [academic] programs will mean in the marketplace of employment and in their lives as citizens and community members" (CHEA, 2003, p. 4).

---

[1]This chapter is based on a study that was supported by a grant from the Accrediting Board of Engineering and Engineering Technology (ABET), and upon an article that is published in *Research in Higher Education*, 48 (March 2007), 129–148. The authors acknowledge and thank the other members of the research team who participated in the development of the design, instruments, and databases for this project: Dr. Linda C. Strauss, senior project associate; Suzanne Bienart, project assistant; and graduate research assistants Betty J. Harper, Vicki L. Baker, Robert J. Domingo, Amber D. Lambert, and Javzan Sukhbaatar. The opinions expressed here do not necessarily reflect the opinions or policies of ABET, and no official endorsement should be inferred.

Assessment of student outcomes may assist higher education accreditors and institutions to answer the increasingly fervent calls for accountability from state and federal legislators. Among accreditors, the trend toward assessment of student outcomes as a criterion for accreditation has gained considerable momentum, but requirements vary by accreditation agency. Regional accreditors generally require that institutions or programs conduct assessment as a condition for accreditation; for some example standards, see the Middle States Commission on Higher Education (2006), New England Association of Schools and Colleges (2005), North Central Association of Colleges and Schools (2003), Northwest Commission on Colleges and Universities (2006), Southern Association of Colleges and Schools (2004), and the Western Association of Schools and Colleges (2005).

Some professional accreditation agencies have taken outcomes assessment a step further by identifying specific learning outcomes to be achieved by accredited programs. For instance, ABET Inc., formerly the Accreditation Board for Engineering and Technology, specifies 11 undergraduate learning outcomes for all baccalaureate engineers, regardless of engineering specialty. Similarly, the Accreditation Council for Graduate Medical Education identified six general competencies (e.g., patient care, medical knowledge, interpersonal skills, and communication) in their accreditation criteria (Batalden, Leach, Swing, Dreyfus, & Dreyfus, 2002). In response to these changes, engineering and medical programs have started to align their curricula with the outcomes stipulated by their respective criteria (see Batalden, et al., 2002; Lattuca, Terenzini, & Volkwein, 2006).

Discussions about the effectiveness of accreditation as a quality assurance tool might be less contentious if there were clear evidence regarding the impact of the process on institutions, academic programs, and graduates. Surprisingly, despite the centrality of the process in higher education, there is little systematic research on the influence of accreditation on programs or learning. Anecdotal accounts of institutional and program responses to new accreditation standards are abundant over the past several years in the proceedings of the American Society for Engineering Education, but there are only a handful of studies that examine the impact of accreditation across institutions or programs. Moreover, these studies typically focus simply on documenting institutional responses. For example, in a study of the impact of changes in accreditation standards for accounting programs, Sinning and Dykxhoorn (2001) found programs (a) working to identify the skills and knowledge base required for employment in the field and (b) developing educational objectives reflecting these skills. Similarly, a survey of 21 mechanical engineering programs conducted by the American Society of Mechanical Engineers (ASME) found that the implementation of EC2000 in these programs "created an environment in which the entire program faculty was involved in the process of establishing program educational objectives and student outcomes and assessment processes" (Laurenson, 2001, p.20). Systematic studies of the impact of accreditation processes on both changes in educational programs and student learning are, to our knowledge, non-existent.

The opportunity to study the impact of an outcomes-based accreditation model on educational processes and student learning arose in 2002 when ABET engaged Penn State's Center for the Study of Higher Education to assess the impact of its new accreditation criteria for undergraduate engineering programs. The new criteria were expected to stimulate significant restructuring of curricula, instructional practices, and assessment activities in engineering programs (ABET, 1997) and the EC2000 impact study, entitled *Engineering Change* (Lattuca, et al., 2006), assesses the extent to which the expected restructuring has occurred and its influence on the 11 student learning outcomes specified in EC2000.

ABET's reform efforts began in the 1990s, as the agency responded to criticisms from two key stakeholder groups. Employers voiced concerns regarding the mismatch between industry needs and the skill sets of graduates of engineering programs. Engineering faculty and administrators countered that ABET's prescriptive accreditation criteria were barriers to curricular and pedagogical innovation (Prados, Peterson, & Aberle, 2001). With funding support from the National Science Foundation and industries represented on the ABET advisory council, ABET conducted a series of workshops to generate ideas about needed change and build consensus among different constituencies. Recommendations from the workshops, published in *A Vision for Change* (ABET, 1995), became catalysts for the development of new criteria, which were circulated to the engineering community a few months later. Following a period of public comment, the ABET Board of Directors approved *Engineering Criteria 2000* (EC2000) in 1996 and released the publication in 1997.

The new criteria radically altered the evaluation of undergraduate engineering programs, shifting the emphasis from curricular specifications to student learning outcomes and accountability. Under EC2000, engineering programs must define program objectives to meet their constituents' needs. To ensure accountability, each program is required to implement a structured, documented system for continuous improvement that actively and formally engages all of its constituents in the development, assessment, and improvement of academic offerings. Programs must also publish specific goals for student learning and measure their achievement to demonstrate how well these objectives are being met (Prados, 1995). ABET was one of the first accrediting bodies to adopt a philosophy of continuous improvement for accreditation and the first to submit that process to scholarly evaluation.

ABET piloted the EC2000 criteria in 1996–1997. After a three-year transition period (1998–2000)—during which programs could choose to undergo review using either the old or new criteria—EC2000 became mandatory in 2001. In this paper, we use data from the *Engineering Change* study to answer two research questions regarding the influence of EC2000 on engineering programs. The first research question focuses on the overall impact of EC2000, asking

- Did the new EC2000 accreditation standards have the desired impact on engineering programs, student experiences, and student outcomes?

The second research question explores the potential influence of the phased implementation of the EC2000 standards over a five-year period, assuming that this may have affected rates of change in programs and student learning. As noted, institutions undergoing an accreditation review between 1998 and 2000 had the option of meeting the new EC2000 standards "early" (i.e., before they became mandatory) or to "defer" an EC2000 review until the next cycle. The decision to undergo EC2000 review early or to defer to the next cycle may have reflected program readiness to meet the requirements outlined in the new criteria. Early adopters of EC2000 accreditation, for instance, may have already introduced outcomes assessment or continuous improvement principles into their programs. Institutions that deferred EC2000 and stood for review under the old criteria may have had one or more programs that were not well positioned to respond to these new requirements. "On-time" programs, the ones that came up for review after EC2000 became mandatory in 2001, may be more like early programs in terms of their readiness to provide evidence of assessment and continuous improvement practices. The second research question for this study therefore asks,

- Do engineering programs in the three review cycle groups (pilot/early, on-time/required, and deferred) differ in program changes, student experiences, and/or student outcomes?

In other words, are programs that deferred adoption of EC2000 significantly different from those programs that did not?

Conceptual Framework of the *Engineering Change* Study

The conceptual model guiding the study (see Figure 1) summarizes the logic of the study's design, assuming that if implementation of the EC2000 evaluation criteria were having the desired effect, several changes in engineering programs would be evident:

- Engineering programs would make changes to align their curricula and instructional practices with the 11 learning outcomes specified by EC2000.
- Changes in program faculty culture would be evident as faculty members engaged at a higher rate than before EC2000 in activities such as outcomes assessment and curriculum revision.
- Faculty and program administrators would adjust program practices and policies regarding salary merit increases, tenure, and promotion criteria to give greater recognition to the kinds of teaching and learning required by EC2000.
- Changes to the quality and character of student educational experiences inside and outside the classroom would be visible.
- All these changes in curricula, instructional practices, faculty culture and student experiences would influence student learning (defined as improved student performance on measures of the 11 EC2000 learning outcomes).
- Employers would report improvements in the knowledge and competencies of the engineering graduates they have hired since implementation of EC2000.

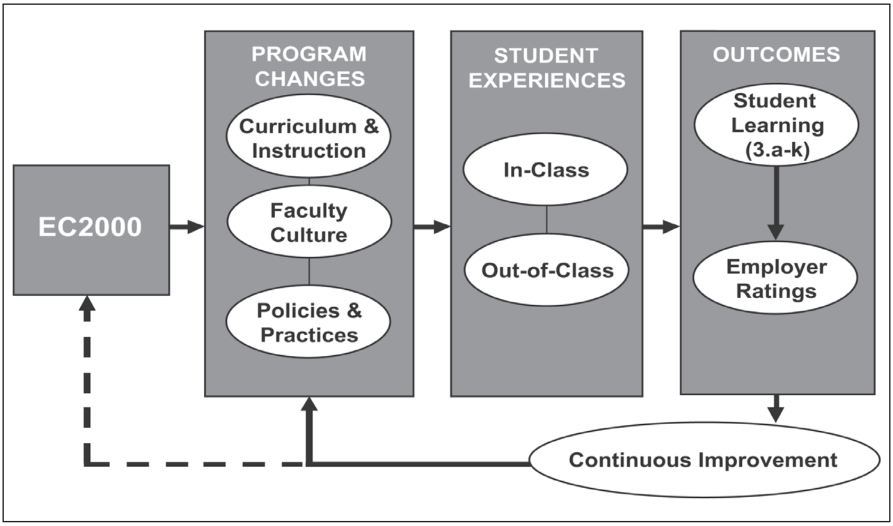

Figure 1. Conceptual framework.

Data Sources and Instruments

The *Engineering Change* study examined accredited engineering programs in selected engineering fields within a representative sample of institutions. The population of programs for the EC2000 study was defined to be those programs accredited since 1990 in seven targeted disciplines; the seven disciplines targeted for the *Engineering Change* study are aerospace, chemical, civil, computer, electrical, industrial, and mechanical engineering. For the past five years, these seven fields have annually accounted for about 85 percent of all undergraduate engineering degrees awarded in the U.S. Of the 1,241 ABET-accredited engineering programs in the targeted disciplines, 1,024 met the 1990 accreditation specification. The project team selected programs for participation in the study based on a two-stage, disproportionate, stratified random sample with a 7x3x2 design. The sample is "disproportionate" because the team over-sampled the smaller disciplines (e.g., aerospace and industrial) to ensure an adequate number of responses for discipline-specific analyses. The sample is stratified on three criteria: (a) the targeted seven disciplines; (b) the three accreditation review cycles (pilot/early, on-time, deferred); and (c) whether the programs and institutions participated in a National Science Foundation Engineering Education Coalition during the 1990s. During the 1990s, the National Science Foundation funded ten, multi-institution Engineering Education Coalitions to design and implement educational innovations and to disseminate findings and best practices in an effort to encourage curricular and instructional reform. Wankat, Felder, Smith and Oreovicz (2002) assert that NSF support has "probably done more to raise awareness of the scholarship of teaching and learning in engineering than any other single factor" (p. 225). To round out the sample, four EC2000 pilot institutions (first reviewed in 1996 and 1997), and several Historically Black Colleges and Universities (HBCUs) and Hispanic Serving Institutions (HSIs) were added. The final sample included 203 engineering programs at 40 institutions (see Table 1).

Table 1

Institutions Participating in Engineering Change: A Study of the Impact of EC2000

Doctoral Institutions

Arizona State University
Case Western Reserve University
Clemson University
Cornell University
Georgia Institute of Technology
Howard University
Illinois Institute of Technology
Iowa State University
Lehigh University
Marquette University
Massachusetts Institute of Technology
Ohio State University
Princeton University
Purdue University
Syracuse University
Temple University
Youngstown State University
Texas A&M University
University of Arkansas
University of California, Los Angeles
University of Florida
University of Illinois at Chicago
University of Michigan
University of Missouri-Columbia
University of Notre Dame
University of the Pacific
University of Texas at Arlington
University of Texas at Austin
Virginia Polytechnic Institute and State University
Western Michigan University
Worchester Polytechnic Institute

Master's Institutions

California State Polytechnic University, Pomona
California State University, Sacramento
Embry Riddle Aeronautical University
North Carolina A&T University
Tuskegee University

Bachelor's and Specialized Institutions

South Dakota School of Mines and Technology
Tri-State University
Union College
United States Military Academy

Data Collection and Analysis

<u>Data Collection</u>

Engineering Change collected data in 2003–04 from several sources, including graduating seniors in 2004, engineering graduates from the same programs in 1994, faculty members and program chairs from these same programs, deans of the participating institutions, and employers (data from deans and employers were not used in the analyses reported here). The Survey of Seniors in Engineering Programs solicited information on basic demographic information, level of participation in out-of-class activities related to engineering education, student-learning outcomes associated with the 11 EC2000 outcomes criteria, classroom experiences, and plans for the future. The companion instrument, the Survey of Engineering Alumni, asked 1994 graduates to report on these same educational experiences and their learning outcomes at the time of their graduation.

The Survey of Engineering Program Changes collected program-level information from program chairs, including changes over time in curricular emphases associated with the EC2000 learning outcomes, levels of faculty support for assessment and continuous improvement efforts, and changes in institutional reward policies and practices. The companion instrument, the Survey of Faculty Teaching and Student Learning in Engineering, collected complementary information on changes in curricula and instructional practices at the course-level, participation in professional development activities, assessments of student learning, and perceptions of changes in the faculty reward system.

In the fall of 2003, survey instruments were sent to 2,971 faculty members and 203 program chairs. In spring, 2004, surveys were sent to the population of 12,144 seniors nearing graduation as well as the 15,734 1994 graduates of the targeted 40 campuses (instruments are available at http://www.ed.psu.edu/cshe/abet/ instruments.html).

Survey responses were received from 4,543 graduating seniors, 5,578 alumni, 1,272 faculty members, and the chairs of 147 engineering programs on 39 campuses. The student and alumni surveys yielded 4,330 (36% response rate) and 5,336 (34%) usable cases respectively. The faculty and program chair respondents yielded usable cases for 1,243 faculty members (42%), and 147 program chairs (representing 72% of the targeted programs).

After conducting analyses of missing data on the student, alumni, and faculty databases, we eliminated respondents who submitted surveys for which more than 20% of the variables were missing. We then imputed values (using expected maximization estimation in SPSS) for missing data in the usable cases. We did not impute personal characteristics such as male/female, race/ethnicity, and discipline, and we did not impute missing data for the program chairs due to the smaller number of cases in the database.

Variables

Three sets of control variables are used in our multivariate analysis: (a) precollege characteristics of graduating seniors and alumni; (b) institutional characteristics; and (c) engineering program characteristics. Precollege characteristics include age, male/female, SAT/ACT scores, transfer status, race/ethnicity, family income, parents' education, high school GPA, and citizenship. Institutional and program characteristics are type of control, institutional type (based on Carnegie Classification, 2001), institutional size and wealth, participation in an NSF Coalition, and engineering discipline.

Table 2
Sample Programs by Accreditation Cycle

Cycle Groups	Programs (*n*)	Pre-EC2000	Mixed	Post- EC2000
Pilot	18	1990–1991	1996–1997*	2002–2003*
Early	69	1992–1994	1998–2000*	2004–2006*
Deferred**	48	1992–1994	1998–2000	2004–2006*
On-time	68	1994–1996	2001–2003*	2007–2009*
Total	203			

* Programs reaccredited during these years are reviewed under the new EC2000 standards.
** These 48 programs postponed for six years their review under the new EC2000 Standards.

Two sets of dichotomously coded independent variables are used in these analyses—student cohort (1994 or 2004) and EC2000 review cycle (pilot/early, on-time, or deferred). See Table 2 for the breakdown of sample programs by review cycle. We are especially interested in comparing the performance indicators of the 48 "deferred" programs with the other groups, because those who deferred had not been reviewed under the new EC2000 standards at the time of our data collection in 2004.

The dependent variables in our multivariate analyses are indicators of program changes, student experiences, and learning outcomes. Table 3 summarizes these 14 scales and one single-item measure that assess changes in program characteristics before and after EC2000. Eight of these scales (four each from the program chair and faculty datasets) measure changes in curricular emphasis on topics associated with the EC2000 learning outcomes (e.g., changes in curricular emphasis on Foundational Knowledge and Skills). Two scales measure changes in emphasis on Active Learning or Traditional Pedagogies. Changes in Faculty Culture were assessed by two scales tapping faculty participation in professional development activities related to instructional improvement and engagement in projects to improve undergraduate engineering education. An additional item asked program chairs to assess changes in faculty support for continuous improvement efforts. Finally, two scales (one for faculty and one for program chairs) measured changes in perceptions of the degree to which the faculty reward system emphasizes teaching. As seen in Table 3, the Alpha reliabilities for these 14 scales range from .90 to .49, with only three falling below .72.

Table 3
Variables Used in ANCOVA Analyses

Program Change Variables [a]

Foundational knowledge as reported by faculty (3-item scale, alpha=.85)
Sample item: Compared to the first time you taught the course how, if at all, has the emphasis on applying knowledge of mathematics, sciences, and engineering changed?
Foundational knowledge as reported by program chairs (5-item scale, alpha=.74)
Sample item: Over the past decade, how, if at all, has your program's emphasis on basic engineering science changed?
Professionalism & societal issues as reported by faculty (4-item scale, alpha=.79)
Sample item: Compared to the first time you taught the course how, if at all, has the emphasis on professional ethics changed?
Professionalism & societal issues as reported by program chairs (5-item scale, alpha=.79)
Sample item: Compared to the first time you taught the course how, if at all, has the emphasis on professional ethics changed?
Communication skills as reported by program chairs (3-item scale, alpha=.79)
Sample item: Over the past decade, how, if at all, has your program's emphasis on verbal communication changed?
Project skills as reported by faculty (4-item scale, alpha=.72)
Sample item: Compared to the first time you taught the course how, if at all, has the emphasis on project management changed?
Project skills as reported by program chairs (4-item scale, alpha=.79)
Sample item: Over the past decade, how, if at all, has your program's emphasis on engineering design changed?
Applied engineering skills as reported by faculty (3-item scale, alpha=.52)
Sample item: Compared to the first time you taught the course how, if at all, has the emphasis on modern engineering tools changed?
Active learning as reported by faculty (7-item scale, alpha=.80)
Sample item: Compared to the first time you taught the course how, if at all, has the emphasis on assignments or exercises focusing on application changed?
Traditional pedagogy as reported by faculty (2-item scale, alpha=.49)
Sample item: Compared to the first time you taught the course how, if at all, has the emphasis on lectures changed?
Assessment for improvement as reported by program chairs (single item)
Over the past decade, have there been any changes in your program's use of assessment information for ongoing, systematic efforts to improve program quality?
Instructional development as reported by faculty (4-item scale, alpha=.73)
Measured on a 3-point rating scale of less, same, or more
Sample item: Compared to five years ago, how often have you participated in seminars or workshops on assessing student learning?
Undergraduate education projects as reported by faculty (5-item scale, alpha=.64)
Measured on a 3-point rating scale of less, same, or more
Sample item: Compared to five years ago, how often have you participated in a project to improve undergraduate engineering education?
Emphasis on teaching in rewards as reported by faculty (3-item scale, alpha=.90)
Sample item: Over the past decade, how has your program's emphasis on teaching in recruiting and hiring changed?
Emphasis on teaching in rewards as reported by program chairs (3-item scale, alpha=.89)
Sample item: Over the past decade, how has your program's emphasis on teaching in recruiting and hiring changed?

[a] Measured on a 5-point scale ranging from *significant decrease* to *significant increase* unless otherwise noted.

Table 3 (continued)

Student Experience Variables [b]
Collaborative learning (7-item scale, alpha=.90)
Sample item: How often did you discuss ideas with classmates?
Instructor interaction and feedback (5-item scale, alpha=.87)
Sample item: How often did you interact with instructors as part of the course?
Clarity and organization (3-item scale, alpha=.82)
Sample item: How often were assignments and class activities clearly explained?
Program encouragement for openness (4-item scale, alpha=.74)
Sample item: In your major, how often did engineering courses encourage you to examine your beliefs and values?
Perceived program climate (4-item scale, alpha=.57)
Sample item: In your major, how often did you observe the use of offensive words, behaviors, or gestures directed at students because of their identity?
Internship/co-op (single item)
Measured on a 5-point rating scale ranging from *none* to *more than 12 months*
Student design competition (single item)
Measured on a 5-point rating scale ranging from *none* to *more than 12 months*
Professional society involvement (single item)
Measured on a 4-point rating scale ranging from *not at all* to *highly*

Student Learning Outcomes Variables [c]
Applying math and science (2-item scale, alpha=.74)
Sample item: Please rate your ability to apply knowledge of mathematics.
Experimental skills (4 items scale, alpha=.89)
Sample item: Please rate your ability to analyze evidence or data from an experiment.
Applying engineering skills (4-item scale, alpha=.94)
Sample item: Please rate your ability to apply engineering tools.
Design and problem solving (6-item scale, alpha=.92)
Sample item: Please rate your ability to design solutions to meet needs.
Communication skills (4-item scale, alpha=.86)
Sample item: Please rate your ability to convey ideas in writing.
Group skills (3 item scale, alpha=.86)
Sample item: Please rate your ability to work with others to accomplish team goals.
Societal and global issues (5-item scale, alpha=.92)
Sample item: Please rate your ability to understand contemporary issues.
Ethics and professionalism (5-item scale, alpha=.87)
Sample items: Please rate your ability to understand the engineering code of ethics.
Life-long learning (3-item scale, alpha=.78)
Sample item: To what extent are you able to learn and apply new technologies and tools?

[b] Measured on a 4-point rating scale ranging from *almost never* to *almost always* unless otherwise noted

[c] Measured on a 5-point scale ranging from *no ability* to *high ability*

Three scales measure students' in-class experiences: Clarity and Organization; Collaborative Learning, and Instructor Interaction and Feedback. Two other scales represent students' perceptions of Program Openness to Ideas and People and Program Climate. Additional out-of-class experiences are measured by single items that assess (a) the degree to which students or alumni were active in a student chapter of a Professional Society, and (b) the number of months students and alumni spent in each of the following: Internship/Cooperative Education, International Travel, Study Abroad, and Student Design Competition.

The learning outcomes scales are derived from a series of principal components analysis of 36 survey items. The research team developed these items through an iterative process designed to operationalize the 11 learning outcomes specified in EC2000. For each item, respondents indicated their level of achievement with regard to a particular skill on a 5-point scale (1 = *No Ability* and 5 = *High Ability*). The principal components analysis produced a nine-factor solution that retained more than 75 percent of the variance among the original 36 survey items. (See Table 3 for sample items and scale reliabilities, and see Strauss and Terenzini [2005] for the full factor structure and description of the item development process.)

As shown in Table 3, the majority of the 28 multi-item scales developed for this study are quite strong and have alpha reliabilities above .80. Only four scales fall below the conventional .70 standard. Such scales can be retained for analyses if they have substantive or theoretical value, which these do.

<u>Analytical Procedures</u>

Data were analyzed using analysis of covariance (ANCOVA) with multiple covariates to control for graduates' precollege characteristics, program, and institutional traits. In the initial analysis, means for the pilot/early, on-time, and deferred EC2000 cycle groups were compared across eight student in- and out-of-class experience scales, nine student outcome scales, 14 program change scales, and one single-item program change variable. The Bonferroni correction was applied in order to mitigate the effects of making comparisons across multiple groups by controlling overall error rate. In the second phase of the analysis, the 1994 and 2004 engineering graduates' mean scores were compared on student experience and outcome variables. Effect sizes were calculated in order to determine the magnitude of the differences between the student cohorts. In the third and final phase of analysis, a series of pairwise multiple comparisons were used to determine the mean differences in program changes, student experiences and outcomes among the three EC2000 cycle groups.

Results of the EC2000 Study

We first examined the data to see if the new EC2000 accreditation standards are having the desired impact on engineering programs, student experiences, and student outcomes. We summarize here the major findings consistent with the logic of the conceptual model in Figure 1.

Changes in Engineering Programs between 1994 and 2004

According to program chairs and faculty members, engineering program curricula changed considerably following implementation of the EC2000 criteria. Both program chairs and faculty members report increased emphasis on nearly all of the professional skills and knowledge sets associated with EC2000 Criterion 3(a) to (k). Three-quarters or more of the chairs report moderate or significant increases in their programs' emphases on communication, teamwork, use of modern engineering tools, technical writing, life-long learning, and engineering design. Similarly, more than half of the faculty respondents report a moderate to significant increase in their emphasis on the use of modern engineering tools, teamwork, and engineering design in a course they taught regularly (These results are not shown here but are available from the first author).

EC2000's focus on professional skills might be expected to lead to changes in teaching methods as faculty members provide students with opportunities to learn and practice their teamwork, design, and communication skills. Consistent with that expectation, half to two-thirds of the faculty report that they have increased their use of active learning methods, such as group work, design projects, case studies, and application exercises, in a course they teach regularly (see Figure 2).

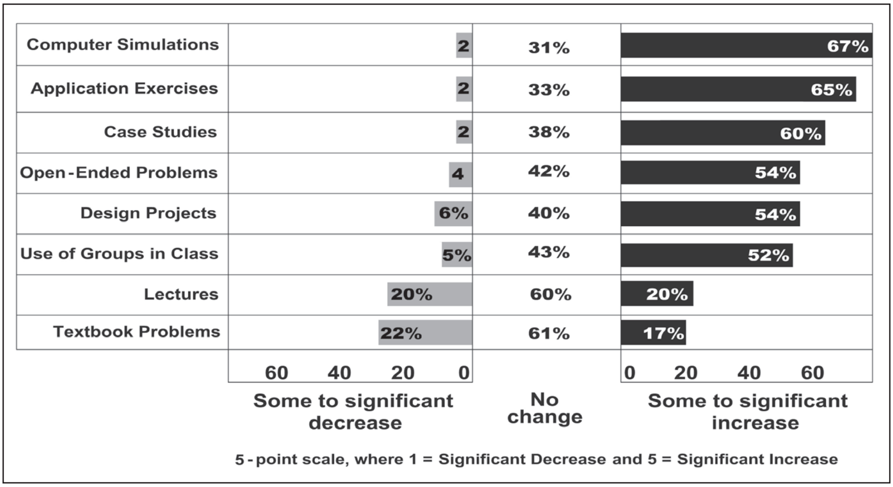

Figure 2. Faculty reports on teaching methods.

EC2000 also requires that engineering programs assess student performance on the (a) to (k) learning outcomes and use the findings for program improvement. Program chairs report high levels of faculty support for these practices (see Figure 3). About three quarters of chairs estimate that either more than half or almost all of their faculty supported curricular development and revision efforts (76%) and systematic efforts to improve (73%). Seventy percent report moderate to strong support for the assessment of student learning, and about two-thirds report similar levels of support for data-based decision-making.

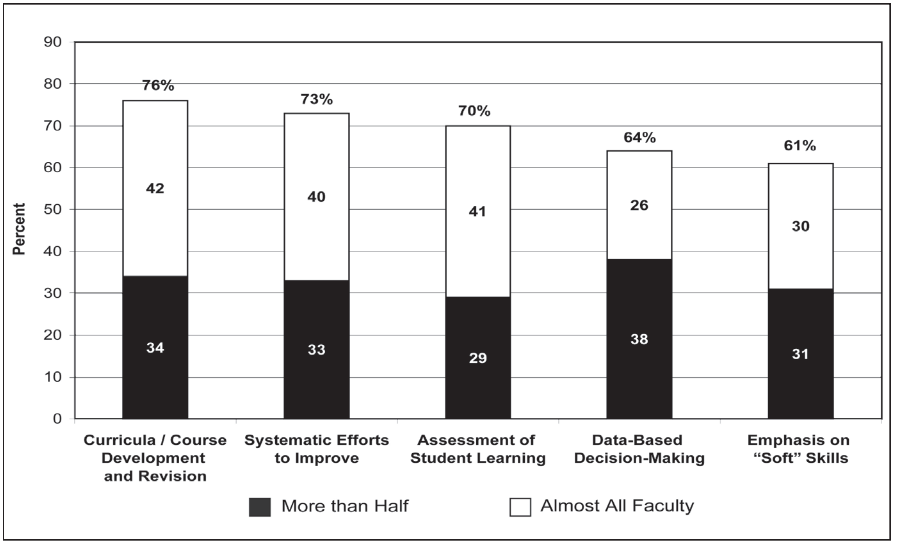

Figure 3. Program chairs' reports of faculty support for EC2000 initiatives.

Faculty generally corroborate this finding: As shown in Figure 4, nearly 90% of the faculty respondents report some personal effort in assessment, and more than half report moderate to significant levels of personal effort in this area. For the most part, moreover, faculty members do not perceive their assessment efforts to be overly burdensome: Nearly 70% think their level of effort was "about right."

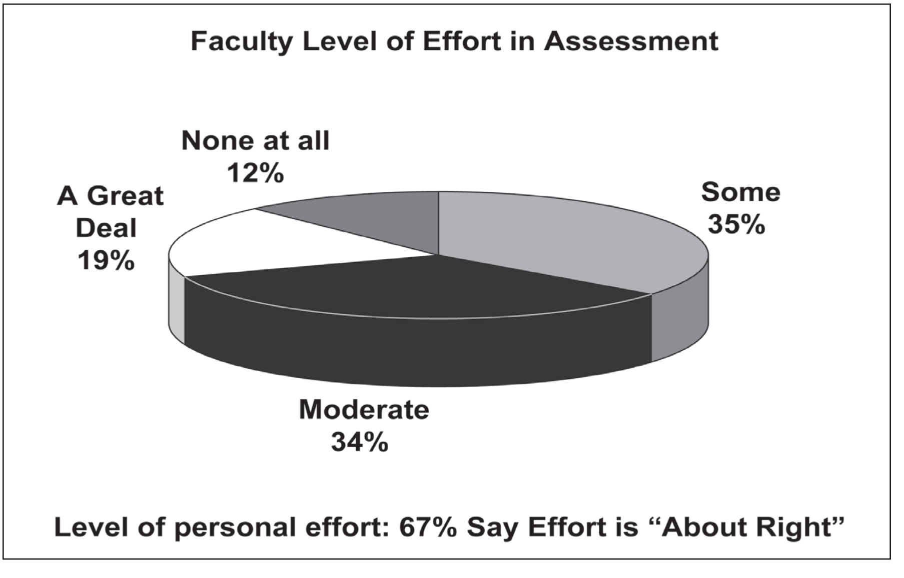

Figure 4. Faculty level of effort in assessment.

One of the most important influences on faculty work in colleges and universities is the institutional reward system, which can encourage or discourage attention to teaching. The EC2000 accreditation criteria require that engineering programs be responsible for the quality of teaching, learning, and assessment, but do faculty members believe that their institutions value their contributions in these areas when making decisions about promotion, tenure, and merit-based salary increases? About half of the program chairs and faculty surveyed see no change in their institution's reward system over the past decade. About one third of the program chairs, however, report an increase over the past decade on the emphasis given to teaching in promotion, tenure, and salary and merit decisions. In contrast, roughly one-quarter of faculty respondents believe the emphasis on teaching in their reward systems decreased in the same time period. Senior faculty members, however, tend to report increased emphasis on teaching in promotion and tenure decisions whereas untenured faculty are more likely to report decreased emphasis.

Changes in Student Experiences

Have the curricular and pedagogical changes reported by program chairs and faculty had a measurable impact on the educational experiences of engineering undergraduates? The evidence suggests they have. Indeed, the experiences of the 2004 graduates differ in a number of ways from those of their counterparts of a decade earlier. The direction of the changes, moreover, is consistent with what one would expect if EC2000 were putting down roots.

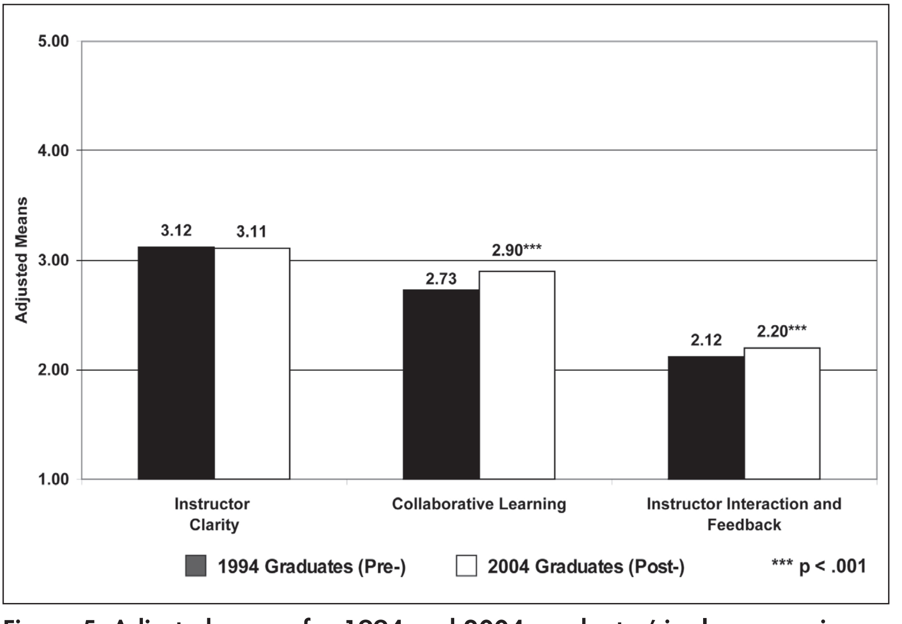

Figure 5. Adjusted means for 1994 and 2004 graduates' in-class experiences during their engineering programs.

As shown in Figures 5 and 6, compared to their counterparts of a decade earlier and controlling for an array of individual and institutional characteristics, 2004 graduates reported:

- More collaborative work and active engagement in their own learning;
- More interaction with instructors and instructor feedback on their work;
- More study abroad and international travel experiences;
- More involvement in engineering design competitions; and
- More involvement in professional society chapters.

Although the differences tend to be small, six of these eight differences between pre- and post-EC2000 graduates are statistically significant.

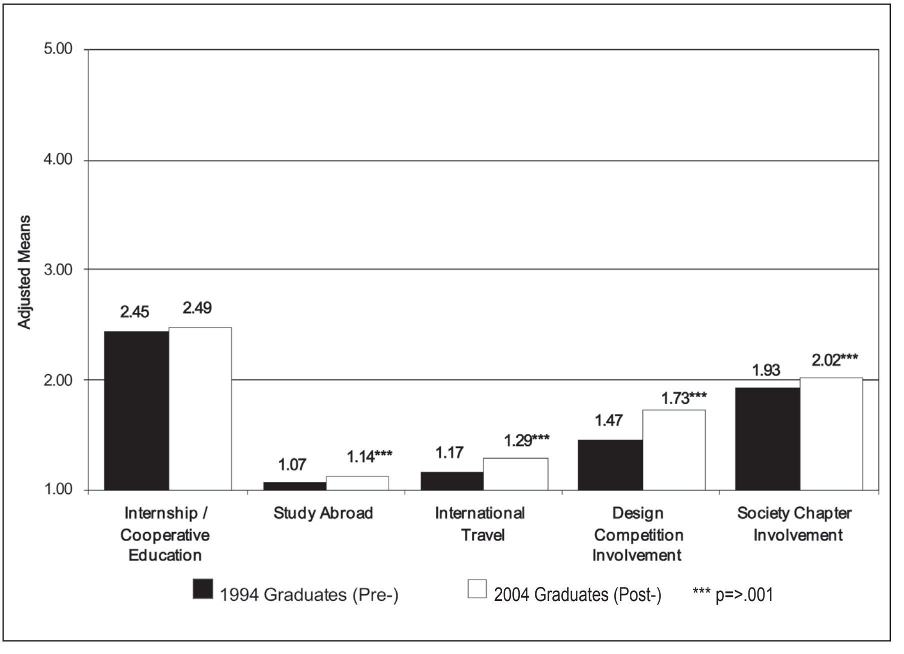

Figure 6. Adjusted means for 1994 and 2004 graduates' out-of-class experiences during their engineering programs.

Changes in Learning Outcomes

As noted in the methodology section, a factor analysis on the battery of items reflecting EC2000 Criterion 3 (a) to (k) learning outcomes yielded nine scales. These nine scales became the dependent variables in a series of multivariate ANCOVA analyses that allowed us to examine the differences between 1994 and 2004 graduates. These analyses controlled for the differences in student pre-college traits, as well as for the characteristics of the institutions and engineering programs they attended.

Each of the nine outcomes scales are based on self-reported ability levels at the time of graduation. A growing body of research over the past 30 years has examined the adequacy of self-reported measures of learning and skill development as proxies for objective measures of the same traits or skills. Although results vary depending on the traits and instruments examined, these studies report correlations of .50 to .90, between self-reports and actual student grades, SAT scores, the ACT Comprehensive Test, the College Basic Academic Subjects Examination, and the Graduate Record Examination. For a review of this literature see Volkwein (2005). Moreover, most social scientists place a great deal of confidence in aggregated scores as an indication of real differences between groups under the conditions prevailing in this study (Kuh, 2005).

Figures 7–9 show the results of the multivariate ANCOVA analyses. We see significant gains between 1994 and 2004 in graduates' ability to:

- Apply knowledge of mathematics, science, and engineering;
- Use modern engineering tools necessary for engineering practice;
- Use experimental skills to analyze and interpret data;
- Design solutions to engineering problems;
- Function in groups and engage in teamwork;
- Communicate effectively;
- Understand professional and ethical obligations;
- Understand the societal and global context of engineering solutions; and
- Recognize the need for, and engage in life-long learning.

In all cases, the differences are consistent with what one would expect under the assumption that EC2000 is having an impact on engineering education. All differences, moreover, are statistically significant ($p < .001$), with effect sizes ranging from +.07 to +.80 of a standard deviation (mean = +.36). Five of the nine effect sizes exceeded .3 of a standard deviation, an effect size that might be characterized as "moderate."

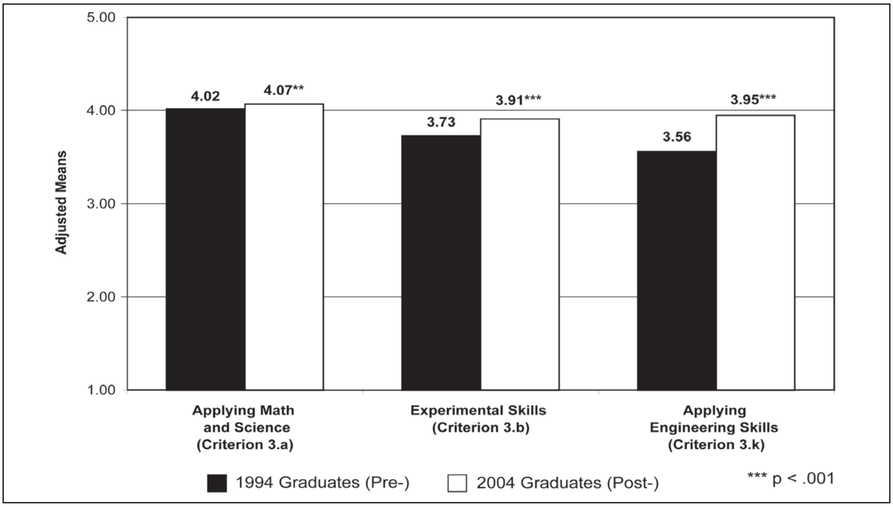

Figure 7. Adjusted means for 1994 and 2004 graduates' reports of their competence in mathematics, science, and engineering science skills.

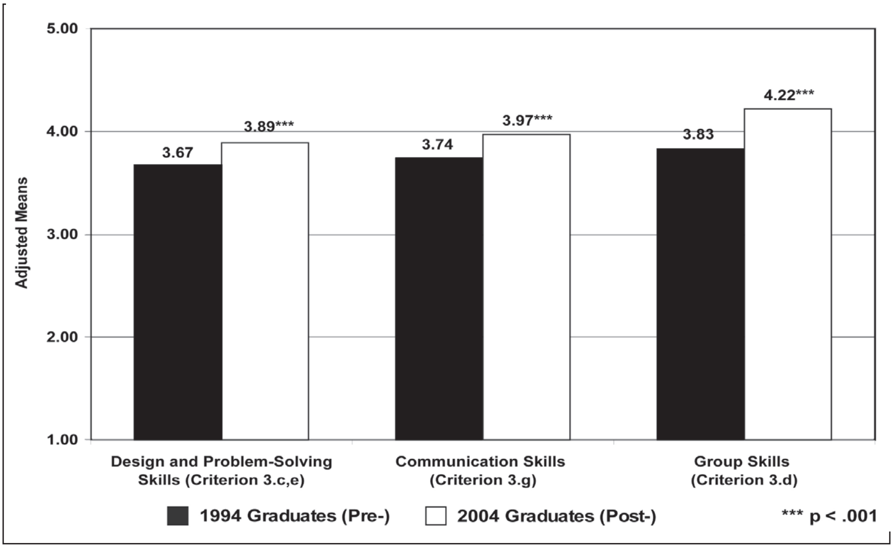

Figure 8. Adjusted means for 1994 and 2004 graduates' reports of their competence in project-related skill areas.

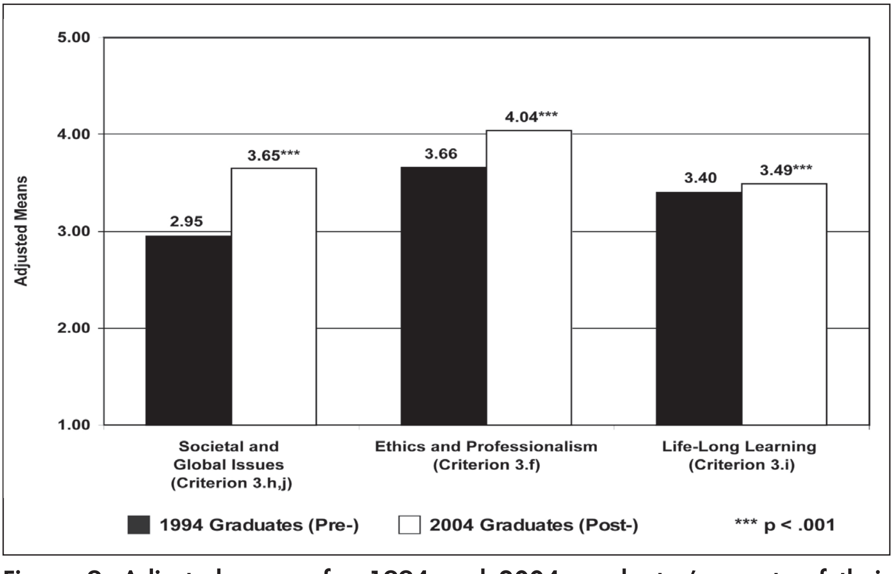

Figure 9. Adjusted means for 1994 and 2004 graduates' reports of their competence in the contexts and professionalism cluster.

The largest differences between 1994 and 2004 graduates are in four areas: Awareness of societal and global issues that can affect (or be affected by) engineering decisions (effect size = +.80 of a standard deviation), applying engineering skills (+.47 sd), group skills (+.47 sd), and awareness of issues relating to ethics and professionalism (+.46 sd). The smallest difference is in graduates' abilities to apply mathematics and sciences (+.07 sd). Despite that small but statistically significant difference, this finding is particularly noteworthy because some faculty members and others have expressed concerns that developing the professional skills specified in EC2000 might require devoting less attention to teaching the science, math, and engineering science skills that are the foundations of engineering. This finding indicates not only that there has been no decline in graduates' knowledge and skills in these areas, but that more recent graduates report slightly better preparation than their counterparts a decade earlier. The evidence suggests not only that implementation of EC2000 is having a positive impact on engineering education, but also that gains are being made at no expense to the teaching of basic science, math, and engineering science skills.

Differential Impact by Year of Accreditation Review

Having documented the EC2000-driven changes between 1994 and 2004, we next examined engineering programs reviewed earlier and later in the accreditation cycle. Are programs that deferred the adoption of EC2000 significantly different from those programs that did not? More specifically, do they differ in program changes, student experiences, and/or student outcomes?

As we note in the methodology discussion above, the ABET six-year accreditation review cycle means that the 203 programs in our study stood for review under the new EC2000 standards at different points in time. After the 18 pilot programs (at four institutions) were reviewed in 1996–97, those engineering programs scheduled for re-accreditation in 1998, 1999, and 2000 were given a choice to undergo review either under the new EC2000 standards or the old standards, thus deferring the EC2000 review for six years. Beginning in 2001, all programs were required to adhere to the EC2000 standards in order to be reaccredited. At the time of data collection in 2004, 18 programs had been reviewed twice under EC2000 (the pilots), 69 programs had opted to be reviewed early (in 1999–2000), 48 programs had elected to postpone their EC2000 review for six years and instead be reviewed under the old standards, and 68 programs underwent EC2000 review as required (shown in Table 2). Hence a key question: Do the engineering student experiences and outcomes in the 48 "deferred" programs differ significantly from the 155 others in 2004?

Congruent with our conceptual model, we first examined the 15 program measures (described in Table 3 above) and found relatively uniform changes in program curricula, in pedagogical practices, and in the general faculty culture (see Table 4). Changes in the faculty and program emphasis on communications skills, project skills, and assessment for improvement were especially strong but relatively uniform across all program groups, even those that had not yet experienced an EC2000 re-accreditation review. Only one of the 15 variables resulted in statistically significant differences: faculty in the Deferred group reported greater emphasis on teaching in the rewards structure than faculty in the other two groups.

Table 4
Adjusted Program Change Means, Standard Deviations, and Significant Differences among the Three Accreditation Groups

Program Changes	Means[a]	Standard Deviations	Significant Differences Between Groups
<u>Curriculum and Instruction</u>			
Foundational Knowledge (Faculty)			
Early	3.130	.524	ns
On-time	3.126	.601	ns
Deferred	3.167	.586	ns
Foundational Knowledge (Chairs)			
Early	3.231	.462	ns
On-time	3.147	.396	ns
Deferred	3.272	.505	ns
Professionalism & Societal Issues (Faculty)			
Early	3.471	.580	ns
On-time	3.448	.596	ns
Deferred	3.551	.578	ns
Professionalism & Societal Issues (Program Chairs)			
Early	3.899	.541	ns
On-time	3.827	.512	ns
Deferred	3.865	.426	ns
Communication Skills (Chairs)			
Early	4.079	.623	ns
On-time	4.097	.627	ns
Deferred	4.037	.485	ns
Project Skills (Faculty)			
Early	3.509	.598	ns
On-time	3.525	.643	ns
Deferred	3.509	.634	ns
Project Skills (Chairs)			
Early	4.078	.651	ns
On-time	4.271	.591	ns
Deferred	4.299	.561	ns
Applied Skills (Faculty)			
Early	3.628	.622	ns
On-time	3.586	.594	ns
Deferred	3.627	.632	ns
Active Learning (Faculty)			
Early	3.554	.622	ns
On-time	3.575	.661	ns
Deferred	3.548	.589	ns
Traditional Pedagogy (Faculty)			
Early	3.205	.649	ns
On-time	3.096	.648	ns
Deferred	3.197	.608	ns

$*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$

[a] Adjusted for institutional control (public/private), NSF Coalition participation, Carnegie Classification, wealth, and size.

Program Changes	Mean[a]	Standard Deviations	Significant Differences Between Groups
Faculty Culture			
Assessment for Improvement (Program Chairs)			
Early	4.335	.639	ns
On-time	4.159	.659	ns
Deferred	4.263	.646	ns
Instructional Development (Faculty)			
Early	2.144	.455	ns
On-time	2.142	.484	ns
Deferred	2.182	.474	ns
UG Education Programs (Faculty)			
Early	2.085	.376	ns
On-time	2.086	.397	ns
Deferred	2.090	.373	ns
Policies and Practices			
Emphasis of Teaching in Rewards (Faculty)			
Early	3.183	.783	ns
On-time	3.214	.780	Deferred sig. > On-time
Deferred	3.294	.639	Deferred sig. > Early
Emphasis of Teaching in Rewards (Program Chairs)			
Early	4.560	1.251	ns
On-time	3.969	1.442	ns
Deferred	4.014	1.281	ns

*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$

[a] Adjusted for institutional control (public/private), NSF Coalition participation, Carnegie Classification, wealth, and size.

Next, we examined the 2004 in-class and out-of-class student experiences shown in Table 5 and found almost the same uniformity among the three groups. For each of the eight indicators of EC2000-relevant experiences and nine indicators of student learning, we tested for significant differences among the three groups, and only six of the 51 tests proved statistically significant at <.05. As shown in Table 5, students in the Deferred programs report significantly greater engagement in collaborative learning than those in the Early and On-time programs, and they also report greater gains in ethics and professionalism than students in the On-time group. On the other hand, those in the Deferred group report less involvement in internships and cooperative education experiences than both the Early and On-time groups, and lower gains in experimental skills than the On-time group. There were no significant differences among the three cycle groups in the 13 other learning experiences and outcomes (e.g., instructor clarity, interaction and feedback, design competition, program openness and diversity, and participation in a student chapter of a professional society). We assumed that the students in the Early and On-time groups would report better EC2000-like experiences and greater learning gains than those in the Deferred group. Instead we find that 45 of the 51 differences in means among the three groups on the 17 indicators are not significant, and three of the six significance tests favored students in the Deferred programs and three favored students in the other two groups.

Table 5
2004 Adjusted Mean Differences in Student Experiences and Outcomes among the Three Accreditation Groups

	Mean Differences[a] and Significance		
	Early – On Time	On Time – Deferred	Early – Deferred
Student Experiences			
In-Class			
Instructor Clarity	-.007	-.011	-.018
Collaborative Learning	-.040	-.073 *	-.113 ***
Instructor Interaction & Feedback	.041	-.024	.017
Out-of-Class			
Internship/Cooperative Education	-.023	.245 ***	.222 ***
Design Competition	.065	.040	.105
Professional Society Chapter	.030	-.118	-.088
Program Diversity Climate	.000	.034	.033
Program Openness to Ideas and People	.017	.018	.034
Student Outcomes			
Applying Math & Science	.035	-.022	.013
Design & Problem Solving	-.059	.059	.000
Experimental Skills	-.024	.087 **	.064
Group Skills	.000	-.036	-.035
Ethics & Professionalism	.045	-.080 *	-.035
Communication Skills	.027	-.036	-.009
Societal & Global Issues	.006	-.028	-.022
Applying Engineering Skills	.010	.018	.027
Life-long Learning	-.008	.011	.003

*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$

[a] Adjusted for student transfer status, age, male/female, US citizenship, race/ethnicity, mother's education, father's education, family income, test scores, high school GPA, discipline, and for institutional control (public/private), NSF Coalition participation, Carnegie Classification, wealth, and size.

Somewhat puzzled by these results, we next looked back at the 1994 data to compare the 1994 and 2004 programs. We already knew from the faculty and program chairs that the changes in programs and curricula across the three groups were surprisingly uniform (as shown in Table 4), so we then examined the 1994 ANCOVA-adjusted means on the eight student experiences and nine outcomes for the three groups. To our surprise, we found that (a) the 1994 Deferred group had the lowest mean on 14 of the 17 variables; and (b) the gains between 1994 and 2004 and the associated effect sizes are greatest for the Deferred group on 13 of the 17 variables (table not included but available from first author). Thus, we appear to have a homogenization or catch-up effect: the three groups were much more disparate in 1994, but much more alike in 2004.

This conclusion is supported by the analysis shown in Table 6, which shows the 51 ANCOVA-adjusted mean differences among the three groups on the 17 indicators of student experiences and outcomes in 1994. The means for the Early and/or On-time groups exceed the means of the Deferred group on 14 of the 17 measures of student experiences and outcomes; and the lag by the Deferred group is especially significant in instructor clarity, interaction and feedback, design competition, diversity climate, design, and problem solving, experimental skills, engineering skills, and life-long learning. Only in Professional Society Chapter participation did the Deferred group exceed the On-time (but not the Early) group. In summary, the Deferred group means in 1994 are significantly lower than the other two groups in eight of the 17 student experiences and outcomes, but by 2004 the deferred group means are lower in only two of the 17 indicators and higher in two.

Table 6
1994 Adjusted Mean Differences in Student Experiences and Outcomes among the Three Accreditation Groups.

	Mean Differences[a] and Significance		
	Early – On Time	On Time – Deferred	Early – Deferred
Student Experiences			
In-Class			
Instructor Clarity	-.047 *	.099 ***	.052 *
Collaborative Learning	.018	.005	.023
Instructor Interaction & Feedback	-.010	.145 ***	.135 ***
Out-of-Class			
Internship/Cooperative Education	.039	.002	.041
Design Competition	.042	.057	.100 **
Professional Society Chapter	.140 ***	-.095 *	.044
Program Diversity Climate	-.029	.049 *	.019
Program Openness to Ideas and People	-.012	.020	.008
Student Outcomes			
Applying Math & Science	.008	.025	.033
Design & Problem Solving	-.046	.120 ***	.073*
Experimental Skills	-.043	.102 **	.059
Group Skills	.053	-.020	.033
Ethics & Professionalism	-.017	-.012	-.030
Communication Skills	.032	.000	.032
Societal & Global Issues	.013	.058	.071
Applying Engineering Skills	-.014	.099 **	.085**
Life-long Learning	.004	.052	.056*

*$p \le 0.05$, **$p \le 0.01$, ***$p \le 0.001$

[a] Adjusted for student transfer status, age, male/female, US citizenship, race/ethnicity, mother's education, father's education, family income, test scores, high school GPA, discipline, and for institutional control (public/private), NSF Coalition participation, Carnegie Classification, wealth, and size.

Conclusions

The 2004 evidence suggests a rather uniform level of post-EC2000 performance across most of the measures and most of the groups in our study. This contrasts with the 1994 evidence showing that the Early and On-time groups exceeded the means of the Deferred group on the majority of the 17 measures of student experiences and outcomes. We suspect that, for one reason or another, the programs and institutions in the deferred group were lagging significantly behind the others in the 1990s, and they appeared to know it. Thus, they elected to stand for re-accreditation under the old ABET standards and to defer the EC2000 review for six years so they could catch up.

The EC2000 Study provides the education research community with evidence of a connection between changes in accreditation and the subsequent improvement of programs, curricula, teaching, and learning in undergraduate programs. Our findings reveal that engineering programs have changed substantially since the implementation of EC2000. In general, programs increased their emphasis on curriculum topics associated with EC2000 as well as the emphasis on active learning strategies. Engineering faculty are more engaged in assessment activities and in professional development activities related to teaching and assessment. The nature of the student educational experience has also changed. Compared to alumni of a decade ago, today's engineering graduates are engaged in more collaborative learning activities in the classroom, interact more with faculty, and are more actively involved in co-curricular activities such as engineering design competitions and student chapters of professional organizations. Today's graduates also report significantly higher levels in all of the student learning outcomes specified in the EC2000 criteria, and in some cases, the differences between 1994 and 2004 graduates are substantial.

All of this evidence suggests that EC2000 is working as expected. We stress, however, that these changes are not attributable solely to EC2000. Other analyses have demonstrated that programmatic changes are due *in part* to EC2000 but that additional external and internal influences also shape engineering programs (for analyses and discussion, see Lattuca, et al., 2006; Lattuca, Strauss, and Sukhbaatar, 2004). Rather than the sole influence, EC2000 accreditation is an important driver in a set of convergent factors (including faculty initiatives, external funding for projects to improve teaching and learning, and employer feedback) that influence educational activities and learning in engineering programs.

The current study, however, provides additional convincing evidence supporting the important role that accreditation has played in engineering education. When we began this study, we expected to find that engineering programs varied in their responsiveness to EC2000. We assumed that programs that deferred their EC2000 accreditation review year would demonstrate less change than those that adopted EC2000 early. We expected to find lower levels of curricular and instructional change and different kinds of student classroom and out-of-class experiences in the deferred populations. We also expected that graduates of deferred programs would report lesser levels of preparation on the 11 EC2000 learning outcomes. Our analyses, however, indicate that, although the deferred programs were indeed significantly behind the EC2000 early adopters in 1994, by 2004 these differences among Early/On-time and Deferred programs had largely disappeared.

Our findings suggest that EC2000 was indeed a catalyst for change. By 1996, all engineering programs knew that an EC2000 review lay ahead of them—and they all began to respond in kind. The deferred programs bought time by using the old criteria for reviews between 1998 and 2000, but they did so, apparently, knowing that they lagged behind the others. As these analyses indicate, deferred programs made considerable progress in meeting the goals expressed in EC2000 by 2004, actually catching up to programs that were, in 1994, better prepared for an EC2000 review.

A remaining question concerns the generalizability of these findings to regional accreditation and to professional accreditation in other fields. Regional, or institutional, accreditation often requires the use of outcomes assessment, but does not specify particular learning outcomes for all institutions. The diversity of institutional missions presumably precludes the identification of a single set of learning outcomes for all colleges and universities. Yet, the American Association of Colleges and Universities (AAC&U) recently published the results of its decade-long national effort to identify liberal learning outcomes for all college graduates. By undertaking studies of high school and college students, faculty, employers, civic leaders, and accrediting bodies, AAC&U has galvanized a national consensus and identified a set of college student outcomes that are needed by today's citizens. Reported in *Taking Responsibility for the Quality of the Baccalaureate Degree* (2004), these outcomes, such as written and verbal communication, inquiry and critical thinking, quantitative literacy, teamwork, and information literacy, are highly consistent with the professional outcomes specified in EC2000. Moreover, the AAC&U efforts also have discovered goal congruence among the regional and specialized accrediting agencies. The specialized accrediting agencies that participated in the AAC&U Project on Accreditation and Assessment (PAA) were unanimous in their belief that a strong liberal education was essential to success in each of their professions (AAC&U, 2004). For example, seven of the eleven EC2000 learning outcomes for engineering and technology graduates are visible in almost any list of desired general education outcomes in any major at most leading institutions. Hence, academic autonomy and differential mission may not be the barrier it has been presumed to be—and the identification of some shared learning outcomes for all graduates does not negate the need to specify unique outcomes that respond to additional important institutional values.

In the case of professional accreditation, it is reasonable to wonder whether accreditors' specification of learning outcomes will result in the kinds of consistent changes in programs and student outcomes we see in engineering. Like the regional accreditors, ABET enjoys almost universal compliance; 96 percent of all engineering programs in the United States are ABET-accredited. The number of programs seeking accreditation, however, varies by professional field, and in some fields there are multiple accreditors. Widespread change may depend on the perceived importance of accreditation in each field. Some institutions may simply decide that the return on investment from specialized accreditation is not enough to justify overhauling their educational programs.

Despite variations in the types of accreditation and their application in particular fields, the study of the impact of EC2000 provides useful evidence that can inform discussions of accountability and quality assurance in higher education. Additional analyses of these data will provide further information on the role that accreditation plays in program improvement, but the current study at least answers the nagging question about what influence, if any, accreditation has on educational quality and student learning in fields like engineering. Importantly, the study provides a model for evaluations in other professional fields and might be replicated by other agencies as they seek to understand—and establish—their impact on undergraduate and graduate education.

References

ABET. (1995). The *vision for change: A summary report of the ABET/NSF industry workshops*. Retrieved April 26, 2008, from http://www.abet.org/papers.shtml

ABET. (1997). *Engineering criteria 2000*. Baltimore, MD: Author.

Association of American Colleges and Universities (AAC&U). (2004). *Taking responsibility for the quality of the baccalaureate degree*. Washington, DC: Author. Retrieved April 26, 2008, from http://aacu-secure.nisgroup.com/acb/stores/1/product1.cfm?SID=1&Product_ID=83

Batalden, P., Leach, D., Swing, S., Dreyfus, H., & Dreyfus, S. (2002). General competencies and accreditation in graduate medical education. *Health Affairs, 21*(5), 103.

Bollag, B. (2005, December 9). Congress stops giving heartburn to accreditors. *The Chronicle of Higher Education 52*(16), A1, A22–23.

Bollag, B. (2006, April 14). Federal panel floats plan to overhaul accreditation. *The Chronicle of Higher Education 52*(32), A1, A18–19.

Carnegie Foundation for the Advancement of Teaching. (2001). *The Carnegie classification of institutions of higher education* (2000 ed.). Menlo Park, CA: Author.

Council for Higher Education Accreditation (CHEA). (2003). *Statement of mutual responsibilities for student learning outcomes: Accreditation, institutions, and programs*. Washington, DC: Author. Retrieved April 26, 2008, from www.chea.org/pdf/StmntStudentLearningOutcomes9-03.pdf

Kuh, G. D. (2005). Imagine asking the client: Using student and alumni surveys for accountability in higher education. In J. C. Burke & Associates (Eds.), *Achieving accountability in higher education: Balancing public, academic, and market demands* (pp. 148–172). San Francisco: Jossey-Bass.

Lattuca, L. R., Strauss, L. C., & Sukhbaatar, J. (2004, November). *The impact of program accreditation on curriculum & instruction: Findings from a national study*. Paper presented at the Annual Meeting of the Association for the Study of Higher Education, Kansas City, MO.

Lattuca, L. R., Terenzini, P. T., & Volkwein, J. F. (2006). *Engineering change: A study of the impact of EC2000. Executive Summary and Final Report.*. Baltimore, MD: ABET Inc. Available from http://www.abet.org/order.shtml

Laurenson, R. M. (Ed.). (2001). *Initial assessment of the impact of ABET/EC2000 implementation using mechanical engineering programs as the pilot study group: A project of the ASME Committee on Engineering Accreditation.* NY: American Society of Mechanical Engineers.

Middle States Commission on Higher Education. (2006). *Characteristics of excellence in higher education: Eligibility requirements and standards for accreditation*. Philadelphia, PA: Author. Retrieved April 26, 2008, from http://www.msche.org/publications_view.asp?idPublicationType=1&txtPublicationType=Eligibility+Requirements+and+Standards+for+Accreditation

New England Association of Schools and Colleges, Commission on Institutions of Higher Education. (2005). *Standards for accreditation*. Bedford, MA: Author. Retrieved April 26, 2008, from http://cihe.neasc.org/standards_policies/standards/standards_html_version

North Central Association of Colleges and Schools, The Higher Learning Commission. (2003). *Handbook of accreditation* (3rd ed.). Chicago, IL: Author. Retrieved April 26, 2008, from http://www.ncahlc.org/index.php?option=com_content&task=view&id=37&Itemid=116

Northwest Commission on Colleges and Universities. (n.d.). *Accreditation standards*. Redmond, WA: Author. Retrieved April 26, 2008, from http://www.nwccu.org/Standards%20and%20Policies/Accreditation%20Standards/Accreditation%20Standards.htm

Prados, J. W. (1995, August). *Engineering education and accreditation reform: How did we get here, and where are we going?* Presentation at ABET Synthesis workshop, Baltimore, MD.

Prados, J. W., Peterson, G. D., & Aberle, K. B. (2001, September). *A Vision for change: The transformation of US educational quality assurance in engineering*. Paper presented at the SEFI (European Society for Engineering Education) Conference, Copenhagen, Denmark.

Sinning, K. E., & Dykxhoorn, H. J. (2001). Processes implemented for AACSB accounting accreditation and the degree of faculty involvement. *Issues in Accounting Education*, *16*(2), 181–205.

Southern Association of Colleges and Schools, Commission on Colleges. (2004). *Principles of accreditation: Foundations for quality enhancement*. Decatur, GA: Author. Retrieved April 26, 2008, from http://www.sacscoc.org/principles.asp

Strauss, L. C., & Terenzini, P. T. (2005, June). *Assessing student performance on EC2000 Criterion 3.a–k*. Proceedings from ASEE Annual Conference, Portland, OR.

Volkwein, J. F. (2005). *On the correspondence between objective and self-reported measures of student learning outcomes*. Resource paper. Retrieved April 26, 2008, from http://www.ed.psu.edu/cshe/abet/resources.html

Volkwein, J. F., Lattuca, L. R., Caffrey, H. S., & Reindl, T. (2003). *What works to ensure quality in higher education institutions*. AASCU/CSHE Policy Seminar on Student Success, Accreditation, and Quality Assurance. Washington, DC: AASC&U and Penn State University.

Volkwein, J. F., Lattuca, L. R., Harper, B. J., & Domingo, R. J. (2007). The impact of accreditation on student experiences and learning outcomes. *Research in Higher Education, 48 (March),* 129–148.

Wankat, P. C., Felder, R. M., Smith, K. A., & Oreovicz, F. S. (2002). The scholarship of teaching and learning in engineering. In M. T. Huber & S. Morreale (Eds.), *Disciplinary styles in the scholarship of teaching and learning: Exploring common ground*. Washington, DC: American Association for Higher Education.

Western Association of Schools and Colleges. (2005). *How to become accredited: Procedures manual for eligibility, candidacy, and initial accreditation*. Alameda, CA: Author. Retrieved April 26, 2008, from http://www.wascsenior.org/wasc/PDFs/HowtoBecomeAccreditedManual8.4.06.pdf

CHAPTER 3

SOUND EVALUATION OF STUDENTS—INTEGRAL TO FULFILLING ABET PROGRAM EXPECTATIONS[1]

Arlen R. Gullickson
Western Michigan University
Jean Gullickson
Brown University

Introduction

Student evaluation practices continue to vex educators in all disciplines, including those engaged in engineering education. Evaluations are expected, indeed ABET (2007) mandates it, but there continue to be major challenges in implanting evaluations to achieve the desired objectives.

To serve program accreditation purposes, ABET sets forward eight general criteria for baccalaureate level programs. Five of the criteria set precursor conditions for accrediting engineering programs: Criterion 4 (professional component) specifies three curriculum components that are required for accreditation, but does not prescribe specific courses; Criteria 5 through 8 set forward expectations for faculty, facilities, institutional support and financial resources, and program criteria**.** The first three criteria, Criteria 1 through 3, individually and collectively call for strong evaluations of engineering programs and students within those programs. These criteria neither specify how these evaluations should be conducted (i.e., planned and carried out) nor what specific topics should be taught and learned by the students. Determination of these evaluation objects and practices is left to local institutions and educators. However, many students will take the Fundamentals of Engineering (FE) exam (National Council of Examiners for Engineering and Surveying, 2007) in order to earn certification in industry, thus curricula at many institutions are designed to cover the engineering areas specifically covered by this exam.

In this chapter we introduce and encourage use of *The Student Evaluation Standards* (Joint Committee on Standards for Educational Evaluation, 2003) as a means to improve the quality of student evaluations. We also focus directly on a particular type of evaluation practice that has been demonstrated to improve student interest and student learning. We argue that engineering programs and ABET stand to gain a great deal

[1]Table 1 and the survey instruments included in this chapter are copyrighted by the chapter authors. Permission to reproduce and disseminate Table 1 is granted with the proviso the source is cited. If the surveys are reproduced and disseminated, their origins, The Student Evaluation Standards, should be cited.

by learning and applying these materials and practices. We begin by addressing three distinct but key aspects of evaluation practice for engineering education:

1. ABET's criteria, which call for strong evaluation in basic level engineering programs;
2. *The Student Evaluation Standards* (Joint Committee, 2003), which guide educator practices in the conduct of sound student evaluations; and
3. current research on the use of assessment practices to serve student learning—assessments of and for learning.

Based on those strands of work, a small self-study of classroom evaluation practices was conducted. This self-study included three small-scale surveys to examine assessment practices within an individual course. Findings from that self-study were used in conjunction with our knowledge of current research and practice to make recommendations for improving assessment and evaluation of undergraduate engineering students.

ABET Criteria

Every engineering program builds from a core of basic courses. No one advances to higher levels or engineering specialties without taking these courses. As such, they are ubiquitous and serve as gatekeepers for the profession and provide essential building blocks for professional engineering expertise. These courses are the foundation upon which engineering programs rest.

Recognizing the importance of these foundation courses, for the 2007–2008 accreditation cycle, ABET specifies eight essentials for baccalaureate level programs:

1. Students
2. Program Educational Objectives
3. Program Outcomes and Assessment
4. Professional Component
5. Faculty
6. Facilities
7. Institutional Support and Financial Resources
8. Program Criteria (ABET, 2007)

In 2008, the ABET accreditation will have nine criteria; the change being the addition of a new Criterion 4 on continuous improvement and the subsequent renumbering of the other criteria.

Within ABET's explanations of these criteria are powerful expectations for students and the use of evaluations to serve their instruction. For example, the first criterion, students, calls for programs to establish policies and procedures that ensure the quality and performance of students. In part this criterion states that, "The institution must *evaluate student performance*, advise students regarding curricular and career matters, and *monitor students' progress* to *foster their success* [italics added] in achieving program outcomes, thereby enabling them as graduates to attain program objectives" (ABET, 2007, p. 1). Those who read this criterion carefully will note that the focus is on individual students, not on students collectively.

The second criterion, program educational objectives, calls for published educational objectives consistent with the mission of the institution. Additionally, it calls for "a curriculum that prepares students to attain program outcomes and that fosters accomplishments of graduates" and "a process of ongoing evaluation of the extent to which these objectives are attained" (ABET, 2007, p. 1).

The third criterion, program outcomes and assessment, calls for strong assessment (evaluation) to ensure success in meeting programmatic objectives:

> Each program must formulate program outcomes that foster attainment of the program objectives articulated in satisfaction of Criterion 2 of these criteria. There must be processes to produce these outcomes and an assessment process, with documented results, that demonstrates that these program outcomes are being measured and indicates the degree to which the outcomes are achieved. (ABET, 2007, p. 2)

In order for the engineering program to meet these criteria, individual courses must attend to these same matters. That is, students must be monitored and fostered in each course. Each course must have clearly stated objectives aligned with the program's intentions, and each course must have stated outcomes aligned with each objective, along with processes to achieve those outcomes. Just as each criterion has evaluation expectations to serve programmatic needs, each course must engage evaluation to similarly assure compliance with the stated criteria at the course level. Meeting the criteria at the course level is a necessary condition for achieving the criteria at the program level.

As these criteria clearly show, ABET expects assessments to be conducted within engineering classrooms and for the program as a whole. The strength of ABET's criteria for assuring the quality of engineers is rooted in assessment of engineering students, which first of all must foster individual student learning.

The Student Evaluation Standards

ABET guidelines leave it to individual programs to determine the ways and means by which to achieve these evaluation expectations. That is where *The Student Evaluation Standards* come to the fore. These standards were developed by the Joint Committee on Standards for Educational Evaluation (Joint Committee). The Joint Committee is a nonprofit organization, American National Standards Institute (ANSI) Accredited Standards Developer, which comprises members from 16 national and international education and evaluation organizations. These standards are certified as American National Standards by ANSI.

Published as a book, *The Student Evaluation Standards* (Joint Committee, 2003) are guidelines for conducting sound student evaluations to serve student learning. This book contains 28 separate standards and provides general guidance for employing these standards to conduct evaluations within classrooms to serve both educational accountability and student learning purposes. The standards are organized in four categories of essential evaluation characteristics to assure that student evaluations are conducted in ways that are sound in terms of propriety, utility, feasibility, and accuracy:

- The Propriety standards help ensure that student evaluations are conducted lawfully, ethically, and with regard for the rights of students and other persons affected by student evaluation.
- The Utility standards promote the design and implementation of informative, timely, and useful student evaluations.

- The Feasibility standards help ensure that student evaluations are practical; viable; cost-effective; and culturally, socially, and politically appropriate.
- The Accuracy standards help ensure that student evaluations will provide sound, accurate, and credible information about student learning and performance (Joint Committee, 2003).

Although intended for broad application, the primary focus of these standards is to promote sound, credible, and accurate evaluations that foster student learning and development at the classroom level. That focus on fostering student learning makes these standards substantially different from other standards that focus on preparation and use of tests and other assessments to measure student learning (e.g., *Standards for Educational and Psychological Testing* [AERA, APA, & NCME, 1999]). In the current vernacular, *The Student Evaluation Standards* address both assessment *of* learning and assessment *for* learning. Assessment of learning aligns directly with the part of ABET's Criterion 1 that calls for evaluating student performance to monitor student progress. Similarly, assessment for learning aligns with the part of Criterion 1 that calls for evaluating student performance to foster student success. We note that ABET calls for evaluations to "monitor student's progress to foster their success *[sic]*." In the following paragraphs we argue that such language confuses the important differences embedded in monitoring and fostering.

Moving Beyond Assessment *of* Learning to Assessment *for* Learning

Assessment of learning focuses on gathering information to help determine whether a student has achieved educational objectives. (Assessment of learning is monitoring student progress.) Examples of this type of assessment include chapter tests, midterm exams, final exams, industry examinations for certification (e.g., the Fundamentals of Engineering exam, [NCEES, 2007]), and virtually any examination that is used for accountability purposes. Results from these tests are used regularly to grade student progress and certify individuals as capable of conducting skilled work. To a much less extent, these assessments provide professors and teachers with insights regarding the effectiveness of their instruction and where changes may be needed.

Assessments of learning practices have been strongly influenced by the measurement profession's *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999). These standards call for careful construction and validation of tests, standardization of testing practices, and the creation of norm groups for use in interpreting test scores. While these standards are best known for assuring the quality of large-scale, standardized exams (e.g., the Graduate Record Exam and the Fundamentals of Engineering exam), educators regularly depend on them in their assessment of student learning in classrooms as well.

Currently, the most dramatic and largest scale assessment of learning taking place in the United States occurs under the auspices of the No Child Left Behind Act (NCLB). NCLB requires annual measures of student outcomes in reading and mathematics. In turn, these measures provide a yardstick to gauge how well our students, teachers, and schools are performing.

One important result of NCLB testing is the growing understanding that whether used on a large or small scale, assessment of learning does not by itself serve student learning well (i.e., it does *not* foster student learning). Certainly, test results help pinpoint classrooms, teaching situations, and schools where students perform well or not well. However, because these tests are typically given summatively (i.e., after learning is to have occurred), infrequently (e.g., end of term), and are broad in scope, they provide poor guidance and support for student learning. Typically, a teacher may gain some understanding of what was or was not well learned in the course and might use that information when he or she teaches the course again. Much less use is made of these exams for immediate reteaching or correcting the understanding of students. Despite its limitations for serving learning, assessment of learning is the "elephant in the classroom." It has been and remains the major use of assessments.

In tandem with, but almost unrelated to, assessment of learning has been a growing interest in and attention to assessment for learning. Assessments for learning practices have grown out of classroom-based research, with little corporate support, and are keyed to student learning. Instead of highly standardized processes that are completed summatively, assessment for learning focuses on a variety of classroom activities that are teacher-based, geared to specific (day-to-day) learning activities and, as Wiggins (2004) noted, have three primary attributes: feedback, guidance, and evaluative judgments focusing on serving student learning needs.

Assessment for learning is rooted in work by authors such as A. R. Gullickson (1984, 1985, 1986, 1993); Haertel (1986); Stiggins and Conklin (1988, November); Stiggins, Griswold, and Wikelund (1989, Fall); and others who described teachers' assessment practices; lamented teachers' preparation for testing as well as the quality of their assessments; and noted that classroom-based tests focused not on serving student learning, but on determining what students had learned. That early work led to many studies in this target area and a concerted attempt by U.S. and Canadian educators to develop and codify professional standards to aid educators in matters of student evaluation practices in service of student learning. Those efforts resulted in two sets of standards: *Principles for Fair Student Assessment Practices for Education in Canada,* developed by the Joint Advisory Committee (1993) that represents 10 Canadian education organizations and *The Student Evaluation Standards,* developed by the Joint Committee on Standards for Educational Evaluation (2003) that represented 16 U.S. and Canadian education and evaluation organizations.

Concurrent with the U.S. and Canadian research and development efforts, a major meta-analysis was being completed by Black and Wiliam (1998) in England. Their findings mark what appear to be a watershed in educators' understanding of the importance of assessment practices and their central role in student learning. Black and Wiliam found "typical effect sizes of the formative assessment experiments increased the average score of students between 0.4 and 0.7 of a standard deviation" (p. 141). Increases of this size are large instructional effects—"larger than most of those found for educational interventions" (p. 141). To illustrate what these gains really mean, they provided these examples:

- An effect size of 0.4 would mean that the average pupil involved in an innovation would record the same achievement as a pupil in the top 35 percent of those not so involved.
- An effect size gain of 0.7 in the recent international comparative studies in mathematics would have raised the score of a nation in the middle of the pack of 41 countries (e.g., the U.S.) to one of the top five (Black & Wiliam, 1998, p. 141).

They also found that "many of these studies arrive at another important conclusion: that improved formative assessment helps low achievers more than other students and so reduces the range of achievement while raising achievement overall" (Black & Wiliam, 1998, p. 141).

Publication of these findings, in concert with an array of other studies on this same topic, led to widespread recognition of the importance of this assessment for learning. In 2001, the National Research Council (NRC) summarized these findings in its book, *Knowing What Students Know*. There the NRC stated, "Moreover, national standards in science and mathematics recognize this type of assessment as a fundamental part of teaching and learning" (p. 41).

There also has been a growing interest on the part of major testing corporations (e.g., Educational Testing Service or ETS) in developing assessment protocols and products that can serve formative assessment purposes. Yet, the continued work of researchers such as Black, Harrison, Lee, Marshall, and Wiliam (2004); Hill and Crevola (2003); and Stiggins (2001, 2004) shows that the strengths of assessment lie more in educational leadership and practices rather than in specific instruments. Work by Black et al. (2002, 2004), for example, has focused on working with teachers to improve four areas of classroom teaching and learning: questioning, feedback through marking, peer- and self-assessment by students, and the formative use of summative tests. Researchers at the Center for Applied Special Technology (www.Cast.org) have introduced the practice of sharing student lecture notes as a feedback and instructional tool to guide and promote learning. Richard Hake (2006) has addressed pre- and post-testing dilemmas in measuring student learning. Wiggins (2004), as noted above, has parsed teachers' curriculum-based assessment practices into three component pieces (feedback, guidance, and evaluation). Chapter 6 of *Knowing What Students Know* (National Research Council, 2001) illustrates a variety of classroom-based assessments for learning purposes.

This description of assessment for learning is not an argument against assessment of learning. Both types of assessment practice play important roles in education. Of course, there must be accountability just as there must be well-developed feedback and guidance to assist students in learning. What we posit is that while assessment practices have developed and flourished to address assessment of learning for matters of accountability (e.g., student grading), the research findings show that assessment for learning produces large rewards in student motivation and learning. What assessments are used, when and how they are used, why—for what purposes, and their quality are issues of substantial importance in achieving student gains. Table 1 enumerates several points of distinction between assessment of learning and assessment for learning.

Table 1

Points of Distinction Between Student Evaluation for Accountability and Development Purposes

Assessment of Learning	**Assessment of Learning**
Purposes	
To provide summative judgments regarding student work and accomplishments	To help students learn
Expectations	
1. Information is for high-stakes decisions (e.g., course grade). Such decisions have a major impact and are not readily reversed once made.	1. Decisions tend to be low stakes in nature (e.g., which strategy is likely to result in greater student improvement), and magnitude of effect for any one decision tends to be small.
2. The number of decisions made (on a per-person basis) as a result of the evaluation is small, and decisions occur infrequently (e.g., tests and other work to determine course grades).	2. The number of decisions tends to be large, and decisions occur frequently (e.g., daily or weekly homework)
3. Each decision requires substantial supporting evidence. Reversal of a decision is infrequent and may require special approvals.	3. Changes occur frequently; reversals of previous decisions are part of normal expectations and frequently require no special approvals.
4. Must focus on minimum standards.	4. Focus is on improvement, not on minimum standards.
5. Must be comprehensive.	5. Is narrowly focused, not comprehensive. Attention is given to a limited segment of the student's work, knowledge, or skills.
6. Must be overtly concerned with integrity of the data (i.e., reliability and validity issues).	6. Data tend to be snapshot-like in character. Users know the data quality is limited and rely on past experience in interpreting the data, expecting limited success based on use of the data.
7. Evaluator and evaluatee cannot be the same person or a peer where conflict of interest or bias can be argued.	7. Evaluatee and evaluator may be the same person or a peer.
8. Data security is essential. Evaluatee usually is not allowed access to raw data, but rather is provided summaries of information collected. Additionally, each category of stakeholder has strict limitations regarding access to the evaluation data, reports, and findings, etc.	8. Minimal concern for data security. In fact, sharing of data and seeking alternative explanations of data meaning together with alternative decision options based on the data is normal and expected.
9. Well-being of the students together with their rights are paramount issues.	9. Mutual trust between those being evaluated (e.g., the teacher) and those providing input to the evaluation (e.g., students and other teachers) is essential.

Regardless of an assessment's purpose, teachers must consider and attend to issues of propriety, ensuring that assessments and attendant evaluative judgments truly serve student learning and other educational purposes across all students without harm to individual students or groups of students. Further complicating matters is the fact that teachers are always beset by bounds on time and resources. Just how much assessment is viable? Under what circumstances do evaluation practices no longer serve student development but detract from it? Unfortunately, there are no easy answers to these questions. Teachers and professors currently are left to their own devices to address these matters. Fortunately, *The Student Evaluation Standards* provides guidance in these matters.

Self-Study of Classroom Evaluation Practices

The investigative exercises described in this chapter are exploratory ones. Our intention was to engage in a "self-evaluation" or analysis of assessment practices that were occurring in an engineering classroom. The engineering course is one that is typical for basic mechanical engineering at the undergraduate level. *The Student Evaluation Standards* chapter, "Applying the Standards," recommends five steps for this activity. In a general sense, we followed these steps, as outlined below:

1. Become very familiar with the standards.
2. Clarify your purpose for applying the standards.
3. Review and select one or more appropriate standards.
4. Apply the standards that you have selected.
5. Based on your application of the standards, decide on and implement a course of action (Joint Committee, 2003, p. 11).

Step 1 had been addressed by both authors prior to starting this work. We participated in development of these standards and were quite familiar with them.

Step 2, clarification of the study's purpose, emerged as we considered options and opportunities to serve both personal instruction interests and increase our understanding of evaluation issues common in basic engineering courses. We used ABET criteria as base points and selected individual student evaluation standards that would help us implement sound practices to meet the ABET criteria.

Simply put, ABET requires each program to set forth objectives; these objectives are to be realized in demonstrable outcomes that are achieved through well-specified processes. As discussed above, meeting these criteria requires parallel objectives, outcomes, and processes for students at the course level. Similarly, student evaluation practices must be engaged in and for student learning within each course.

Our study's initial, explicit purpose was to confirm that the course evaluation practices were being conducted in sound ways that meet ABET's criteria. We wanted evidence that the course monitored student learning and fostered it as well. The study was extended to look more fully at course evaluation practices that foster student learning to achieve the desired outcomes (i.e., serve assessment for learning purposes).

We chose to address these criteria and assessment issues within a basic engineering course for which the coauthor was serving as a section teaching assistant (TA). She had worked with the course professor over several years and assisted in the course's instruction for several semesters in differing capacities. The mechanical engineering course we chose to study was entitled EN31: Mechanics of Solids and Structures and is described as follows:

> "This course covers the basic mechanical behavior of materials and analysis of stress and deformation in engineering structures and continuous media. Topics include concepts of stress and strain; the elastic, plastic, and time-dependent response of materials; principles of structural analysis and application to simple bar structures, beam theory, instability and buckling, torsion of shafts; general three-dimensional states of stress; Mohr's circle; stress concentrations. Lectures, recitations, and laboratory". (Brown University, 2007).

At the time of this study, the course was one semester in length and met for professor-led, 50-minute lectures three times a week. During these lecture periods, the concepts of the course were covered and some example problems were solved.

Supporting the lectures were regular sections, homework, and laboratory exercises, each with its own graduate teaching assistant (TA). Students met weekly with the section TA for an optional 80-minute section. This section included a short review of the material covered in lecture that week. After the review, sample problems resembling the homework assignments were solved and questions on general material from the course were answered. The laboratory TA met the students twice during the semester for course laboratory exercises and the homework TA had no scheduled meeting with them. Students rarely had interactions with the two TAs who were not the section TA; this was especially true of the homework TA. While all three TAs had two hours of office hours during the week, only the office hours of the section TA were well attended.

Over the semester, student evaluations were based on 11 homework assignments, two laboratory exercises, a design project, a midterm exam, and a final exam. The grade for the course was based on the following weight scheme: homework (20%), laboratory (15%), design project (15%), midterm exam (20%), and final exam (30%) (www.engin.brown.edu/courses/en31, 2007).

The two laboratory exercises were developed by the professor and led by a graduate student TA; they occurred at the beginning and end of the semester. Students were required to write comprehensive laboratory reports detailing the procedures and results of the lab. These reports were graded by TAs and the professor. Homework was assigned weekly. Homework solutions were prepared by a TA and graded by undergraduate students who were selected for the job based on their performance in the class in previous years. The midterm and final exams were created by the professor. The TAs took the exam prior to the students to test the validity and feasibility of the exam. The midterm exam covered approximately half of the course material and was given midsemester. The final exam was comprehensive but weighted toward the material not covered by the midterm exam. The final exam was given outside of lecture time; students were allowed 4 hours and 50 minutes to complete the exam. All exams were graded by the professor.

Students were warned about plagiarism and informed of course policies via the course web site. They were allowed to collaborate with other students on homework assignments and laboratory reports, but required to turn in individual assignments and reports. Both the midterm and final exam were completed individually by the students; collaboration was not allowed. However, students were allowed to use their textbooks, homework assignments and solutions, notes, and other handwritten material as references during the exams. The final project was completed in groups, and each group turned in a single report detailing its design and findings.

In applying the standards, step 3 calls for selection of specific standards to address. For purposes of this study, the coauthor identified five student evaluation standards that were of personal interest and consistent with the noted ABET criteria:

P1–Service to Students: Evaluations of students should promote sound education principles, fulfillment of institutional missions, and effective student work, so that educational needs of students are served.

P2–Appropriate Policies and Procedures: Written policies and procedures should be developed, implemented, and made available, so that evaluations are consistent, equitable, and fair.

P5–Rights of Students: Evaluations of students should be consistent with applicable laws and with basic principles of fairness and human rights, so that students' rights and welfare are protected.

P7–Conflict of Interest: Conflicts of interest should be avoided, but if present should be dealt with openly and honestly, so that they do not compromise evaluation processes and results.

A1–Validity Orientation: Student evaluations should be developed and implemented, so that the interpretations made about the performance of a student are valid and not open to misinterpretation.

Each of these standards has associated guidelines and common errors to provide explicit guidance and support for those applying the standards. We consulted those guidelines and common errors in developing the survey questions. Ultimately, we selected individual guidelines or common errors from each standard as a point of emphasis for individual questions. For example, Standard P1–Service to Students (Joint Committee, 2003, pp. 29–32) includes eight guidelines and four common errors. Its Guideline B (p. 30), which calls for clarity of the purposes and uses of the evaluation, provided the foundation for Survey 1, item 2 (see Appendix A), "Has the instructor informed you what the purposes and uses of the evaluations in this class will be?" In all, 15 individual guidelines and common errors were directly used to construct survey items; these 15 in turn link to other guidelines and standards. To assist those who want to use our surveys or develop a comparable instrument, the individual guidelines have been referenced on Surveys 1 and 2 (see Appendices A and B).

To apply these standards (step 4), we sought the opinions of students in the course via three surveys. The first survey (see Appendix A) was keyed to ABET Criterion 2 (program educational objectives) and Criterion 3 (program outcomes and assessment) and to *Student Evaluation Standards* P1–Service to Students and P2–Appropriate Policies and Procedures. The second survey (see Appendix B) also addressed Standard P1, as well as A1–Validity Orientation, P5–Rights of Students, and P7–Conflict of Interest.

The third survey (see Appendix C) probed evaluation's use to foster student learning as called for in ABET Criterion 1 and is inherent in Criterion 2. Standard P1–Service to

Students provided the overall basis for this survey. The survey focused on ways in which the course meets the standard's Guidelines C, D, and E (Joint Committee, 2003, p. 30). Guideline C calls for promoting student evaluations as an integral part of instruction for student improvement. Guideline D seeks to "Ensure that evaluations are aligned with expected learning outcomes and take into account the instruction provided to students" (Joint Committee, 2003, p. 30). Guideline E states, "Show students the relationship between evaluation and student learning. Where possible, directly involve them in the process as student evaluators." We used the assessment ideas promoted by Wiggins (2004) as operational definitions for ways in which Guidelines C, D, and E could be met.

The first two surveys were administered during the student sections. For each, the students were informed that their participation in the survey was voluntary. About one-third of the course's 48 students attended the section on the respective days that Surveys 1 and 2 were administered. In each case all students in attendance (18 and 16 respectively) anonymously completed and returned the survey to the section TA. Because the third survey was conducted by e-mail where anonymity of response was not possible, all 48 students were requested to complete the survey and assured confidentiality of their responses.

Survey 1

This survey was administered to students near the beginning of the semester during one of the 80-minute sections. First, it asked students if the course goals were communicated to them and if the instructor informed them of the purpose and uses of evaluations in the class—students who answered affirmatively were asked to indicate the source(s) of the information (i.e., syllabus, web site, lecture, section, professor, TA). Next, students were asked to indicate if they thought the homework thus far was based on information presented in lectures and if they were informed of alternate evaluation procedures for students with disabilities.

The next set of questions asked students if policies were communicated to them regarding resolving disagreements with instructors, determination of final grades, reporting of evaluation results, plagiarism, use of other students' work, working with other students, and cheating. Students were instructed to mark all the response options that applied to them. Choices included these: *yes, I know the information; no, I do not know the information; yes, I know where to find the information; no, I do know not where to find the information;* and *not applicable*. The final question asked if any of those policies were established following the first homework assignment. While answers to some of these questions were already known to the coauthor, their inclusion provided information regarding the extent to which students attend to various information sources and corresponding needs for redundancy.

As shown in Table 2, the large majority of students (15 of 18 or 83%) confirmed that the educational purposes of the course had been communicated to them. The syllabus and course web site were most frequently identified as the sources of this information.

Also shown in Table 2, noticeably fewer students (12 of 18) reported that they were informed about the purposes and uses of evaluations in the class. Additionally, here the

TA was cited nearly as often as the professor as the source of information regarding how evaluation was to be used. Our own experience with classroom instruction and evaluation practice suggests this is an indicator of a tacit understanding on the part of students that evaluation is just used for summative purposes (assessment of learning). Importantly, these responses provide substantial indications that more must be done to embed evaluation into the fabric of instruction if students are to view evaluation efforts as serving student learning. As we discuss later, when students view evaluation as serving their learning needs, they attend to it differently and use it better.

All of the students indicated that they thought the homework was based on information presented in lectures (see Appendix A, Survey 1, item 3). These responses, when coupled with responses for Survey 1, item 1 above, suggest compliance with ABET's criteria for clarity in instructional purposes and shows consistency of learning opportunities in line with stated purposes.

Table 2
Responses to Items 1 and 2 on Survey 1 (n=18)

Question			Source*					
	Yes	No	Sy	W	L	Se	P	T
1. Was the educational purpose/goals of this class (as outlined in BOCA[1] or according to the ABET standards) communicated to you?	15	3	7	7	5	0	6	0
2. Has the instructor informed you what the purposes and uses of the evaluations in this class will be?	12	6	1	3	3	1	4	4

* Sy=syllabus W=Web site L=lecture Se=section P=professor T=TA
[1] The Brown University Online Course Announcement (Brown University, 2007)

Just one student reported being informed of alternative procedures of evaluation for students with disabilities (Survey 1, item 4). This lone positive response suggests that this person may have approached the professor individually to request "special consideration." We do not know that such was the case. However, it is clear that such information was not conveyed to the class as a whole. As a result, if students did have a need for alternative procedures, it would have been incumbent on them to request such information—something many students, for personal reasons, would not request.

Responses to questions regarding students' knowledge of specific evaluation-related policies indicate that they generally knew how evaluation information would be handled (i.e., they knew the policy and/or knew where to locate the information). As shown in Table 3, they reported being best informed about plagiarism and matters related to cheating. Fewer reported being informed about how information would be compiled, analyzed, and shared with them. Fewest were informed in matters of how to deal with disagreements with instructors. This last point suggests as tacit belief that the professor has the last say or that students did not expect problems of that nature so they had not bothered to "study up" on the matter. None of the students reported that any of these policies were established following the first homework assignment.

Table 3
Positive Responses Regarding Knowledge of Evaluation-Related Policies (n=18)

Question	Positive Responses
1) Do you know how disagreements between a student and the professor/TA on evaluation matters will be resolved?	5
2) Do you know how scores will be aggregated and final grades determined?	12
3) Do you know how the results of evaluations will be reported to you?	9
4) Do you know the policy on plagiarism?	17
5) Do you know the policy on use of other students' work?	14
6) Do you know the policy on sharing of information (i.e., working together on homeworks, labs and/or exams)?	14
7) Do you know what constitutes cheating in this class?	15

Responses to these questions point to some possibilities for improvement in procedures, but certainly do not suggest the need for major changes in the current procedures used to communicate policies.

Survey 2

The second survey was administered to students during one of the 80-minute sections, after the midterm exam and one of the laboratory experiments had been completed. This survey addressed a range of evaluation matters related to students' awareness of their rights, perceptions of security, and appropriateness and fairness of evaluations.

Results from this second survey, which are presented in Table 4, show a consensus in student opinion on seven of the eight items (response options were *yes* or *no*):

- Students indicated that assessments were consistent with the professor's in-class emphases (item 8).
- The instruction (by professor and section TA) was done in a way that promoted validity of the examinations—did not focus on teaching the test (items 3 and 4).
- The instructors (professor and TAs) maintained appropriate evaluation security (item 2).
- Students were not aware of any conflicts of interest that could cause bias on the part of the instructional staff (item 7).

- Students were <u>not</u> involved in the evaluation process. They had not graded their own work or the work of other students (item 5).
- Students tended <u>not</u> to know their rights in evaluation situations (item 1).

On just one item (item 6) did responses indicate a lack of consensus among students. That item asked whether the evaluations properly represented their achievement and understanding of the core concepts in the course. Slightly less than half answered affirmatively.

Table 4
Responses to Survey 2

Question	Positive ("Yes") Responses
1. Do you know what your rights are in an evaluation (i.e., do you know if you have the right to see your graded homework, the conditions under which you are entitled to an alternative evaluation, how to appeal the results of the evaluation, etc.)?	5
2. Do you believe that the evaluators (professor or TA) have kept the results of your evaluations private?	12
3. In class lectures do you believe that the professor(s) teach specifically to the evaluation (i.e., teach such that the questions posed in a given homework or exam are readily answered but general concepts are lost)?	1
4. In sections do you believe that the professor(s)/TA(s) teach specifically to the evaluation (i.e., teach such that the questions posed in a given evaluation are readily answered but general concepts are lost)?	2
5. Have you been involved in the evaluation process (i.e., grading your own paper, grading a classmate's paper, etc.)?	0
6. Overall, do you believe that the evaluation results (so far, the results of your graded homework) properly represent your achievement and understanding of the core concepts in this course?	8
7. Are you aware of any conflict of interest on the part of an evaluator (professor or TA) that might cause a bias in the scoring or grading of your work?	0
8. Do you believe that the content of the assessments (homework assignments) are balanced; that is, that they are consistent with what the professor has emphasized in class?	15

We also note that although the student consensus was that appropriate evaluation security was maintained, in at least one respect this was not the case. Graded assessments were left out in an unmonitored box in the hallway near the professor's office and were therefore not "secured" in any real manner. However, the students may have felt that this was an appropriate amount of security, because they were not concerned about their graded assignments being seen by others.

In summary, responses to Surveys 1 and 2 suggest that students were reasonably informed regarding evaluation policies and procedures and that they generally perceived them to be fair, appropriate, and consistent with ABET criteria. These students did not view themselves as integrally involved in evaluations, except as the persons being evaluated. We would argue that these practices are consistent with evaluation practices in most engineering classrooms and higher education classrooms generally. Responses such as those received here are indicative of sound assessment of learning efforts, but indicate that there are no assessment *for* learning practices currently being implemented.

Where questions could also be viewed as informative about assessment for learning, the results suggest evaluation activities were not aligned with those purposes. For example, all students answered negatively when asked if they had been involved in the evaluation process. This raised questions about the viability of the course assessments in terms of truly serving student learning.

To investigate this issue further, we developed a third survey, described below, to query the students about assessment for learning matters.

Survey 3

In the final survey, we sought to gain understanding of whether evaluations of student work were aligned with assessment for learning practices. We structured this survey based on ideas proffered by Wiggins (2004) for enhancing student learning. He focused on four points: (a) learning objectives and expectations, (b) feedback, (c) guidance, and (d) evaluation. These central concepts are explained below:

1. Standards: Learning objectives and expectations. These typically are conveyed to students in three general forms:
    a. Specifications (e.g., 80 words per minute with 0 mistakes)
    b. Models: exemplars of each point on the grading scale (e.g., anchor papers)
    c. Criteria: conditions to be met to achieve goals (e.g., "persuasive and clear" writing)
2. Feedback: Information conveyed from an instructor to a student about his or her progress toward meeting the intended learning goals (may be written or verbal, directed to a student individually or to the class as a whole). Feedback may include some or all of the following elements:
    a. Facts: what events/behavior happened related to goal, communicated in specific, concrete, nonjudgmental language
    b. Impact: a description of the immediate effects of the facts (results and/or reactions)
    c. Commentary: facts and impact explained in the context of the explicit or implicit goal/intent/standard/model; confirms what was on target, where effect matched intent, to reinforce it; and notes where actions were off target, where effect did not match intent, to underscore the need for redirection)
3. Evaluation:[2] Value judgments about the meaning of the results; may be conveyed in the following forms:
    a. Value judgments made about the facts and their impact

[2]Wiggins (2004). This definition is much more limited than the definition employed by the Joint Committee on Standards for Educational Evaluation, but serves nicely in this particular context.

   b. Praise/Blame: appraisal of individual's performance in light of expectations for that performer
   c. Praise/Blame with feedback
4. Guidance: Information intended to guide future performance; may take either of the following forms:
   a. Advice about what to do in light of the feedback
   b. Redirection of current practice in light of results

From those four key elements, 11 survey items were developed, as shown in Table 5. Students were queried to determine the extent of their agreement (on a 5-point scale from *strongly disagree* to *strongly agree*) with these statements with respect to homework, laboratory exercises, and a course project. In total, the students were asked to provide 33 responses—three responses (one response each for Homework, Laboratory and Exams) for each of the 11 item stems.

Table 5
Item Stems and Their Focus for the Third Survey

Item	Assessment Focus	Item Stems
1	Standards	Learning objectives were clearly specified.
2	Standards	The instructor provided examples for what was expected in terms of learning outcomes.
3	Standards	When I began to work, I understood the conditions that had to be met in order to achieve learning objectives.
4	Feedback (facts)	The instructor's feedback on my completed work concretely described it in relation to the expected learning goals.
5	Feedback (impact)	Instructional feedback provided to me described the results of my work in real terms of my progress toward the learning objectives.
6	Feedback (commentary)	The instructional feedback on my work described my performance in terms of the explicit or implicit learning objectives of the assignment.
7	Evaluation (value judgments)	In evaluating my work, the instructor made value judgments about the facts of my work and their effects in terms of reaching the intended learning objectives.
8	Evaluation (praise/blame)	The instructor did not praise or blame my work in light of the instructor's expectations for me.
9	Evaluation (praise/blame orientation)	When the instructor did provide praise or blame regarding my work, it was not preceded or followed by feedback.
10	Guidance (advice)	Information provided to me regarding this assignment did not include advice about what to do based on the feedback.
11	Guidance (redirection)	Information provided to me regarding my work did show me how to redirect my learning in light of results.

Note that all but three items were worded positively. That is, student agreement with the statements would indicate that the instructors' actions corresponded to practices known to serve student learning. Three items—8, 9, and 10—were worded negatively. Student agreement with those items would be interpreted as indicative that instruction was not being conducted in ways to serve student learning. In item 9, for example, praise/blame is viewed as productive in a learning sense only if it was based on clear criteria and expectations for the student and concretely coupled to feedback.

Given the timing of this survey, the students were no longer enrolled in the course and may have forgotten or misremembered some aspects of it. Also, because the students could not be accessed via a student-help session and the surveys had to be delivered by e-mail, the response rate to the survey was lower (13 students responded). Despite these limitations, we think the findings are helpful in terms of understanding the extent to which assessment for learning occurred in this course.

Because this survey was conducted over e-mail, and the coauthor was not available in person at the survey site to answer questions regarding the survey items and their meanings, we included more contextual information in this survey than the previous two. This contextual information included a glossary meant to assist students who might not know the precise meaning of various terms used in the survey.

For each item, we tallied the number of positive responses. Positive responses were *agree* or *strongly agree* for items 1–7 and 11 and *disagree* or *strongly disagree* for the negatively worded items 8–10. These summary counts are presented in Table 6.

Table 6
Summary of Positive Student Responses for the Homework, Laboratory, and Project Activities (n=13)

Assessment for Learning Element	Item	Homework	Laboratory	Project
Standards	1. Specifications	13	12	13
	2. Models	12	11	13
	3. Criteria	10	10	11
Feedback	4. Facts	7	7	8
	5. Impact	9	5	8
	6. Commentary	12	7	11
Evaluation	7. Value Judgments	4	7	5
	8. Praise/Blame	2	2	3
	9. Praise/Blame Orientation	6	7	6
Guidance	10. Advice	4	4	6
	11. Redirection	2	8	7

As this table shows, the students generally agreed that they were provided with standards for their performance (specifications, models, and criteria) for all three types of assignments. They were much less likely to agree that matters regarding feedback, evaluation, and guidance were handled in recommended ways. These responses are fairly consistent across assessment types. Ratings were more favorable for feedback than for guidance and evaluation. However, we are cautious about the findings for the evaluation items because one student raised questions regarding the meaning of these questions and two of the three items were negatively worded. The result is that students may have responded incorrectly to them.

Step 5 in applying *The Student Evaluation Standards* calls for deciding and implementing a course of action based on the results of the previous step (Joint Committee, 2003, p. 11). Therefore, in the remainder of this chapter, we put forth a number of suggestions for improving evaluation practices for this course, as well as transforming the evaluation of undergraduate engineering students in general. These ideas are grounded in our findings, but also emerged from our discussions about the findings and in consideration of our findings in light of other studies.

Recommendations for Improving and Transforming

Student Evaluation in Engineering

Based on the survey findings, our knowledge of *The Student Evaluation Standards*, and the research on assessment for learning, we have a number of recommendations for (a) improving the student evaluation practices in the engineering course that served as the context of our investigation and (b) transforming the evaluation of undergraduate engineering students to enhance student learning.

Improving Student Evaluations in EN31 and Similar Courses

Our survey findings point to specific changes that could be made in the evaluation practices for the particular course we examined. However, we believe this course is fairly representative of many engineering courses both at the institution where the survey was conducted and at other institutions where engineering courses are taught. This assessment is made based on the coauthor's participation as both a student and a teaching assistant at two Tier I research institutions over the last ten years and conversations with colleagues who have also graduated from other Tier I research institutions where many of the policies are the same.

Articulate and reinforce course purposes. *Knowing What Students Know* (National Research Council, 2001) describes in some detail ideas and issues embedded in the nature and development of subject-matter expertise (pp. 72–92). That discussion draws upon and emphasizes that metacognition is one of the most important aspects of cognition and is essential to building expertise. Here *metacognition* is described as "the process of reflecting on and directing one's own thinking" (p. 79). Providing clarity about course purposes and reinforcing them in the context of daily instruction directly serves these metacognitive objectives.

The fact that fewer than half of the student respondents identified the professor and/or lectures as sources of information about the purpose of the course indicates that it would be beneficial to more overtly identify learning objectives during instruction. Also, no student cited the TA as conveying course purposes to them, suggesting that TAs should provide assistance to students around specified course purposes *and* to refer overtly to those purposes in day-to-day work with students.

Provide detailed feedback. In many engineering classes, feedback and guidance are limited. Feedback often has little detail and is restricted to reporting the grades on homework, exams, laboratories, and projects. For instance, it would be typical for a student in an undergraduate engineering course to receive a corrected homework assignment that has only point tallies for each problem and no comments as to what has been done correctly or incorrectly. There may be praise/blame on these corrected homework assignments; but more often than not, there is no constructive guidance following the praise/blame. Feedback also comes in the form of homework solutions, which are usually posted to the course web site or handed out in class. These will show the students one correct way to solve a problem, but may not help students determine what went wrong with their solution methods.

In our survey, ratings for feedback were more favorable than for guidance and evaluation. However feedback appears to be much more important in terms of student learning, so attention focusing on improving feedback to students can produce greater rewards (cf., Huba & Freed, 1999). Additionally, because value judgments and praise/blame play a quite limited role in terms of supporting student learning, one first step could be to eliminate all evaluative judgments and work solely at increasing the quality of feedback (e.g., presentation of facts) and associated guidance.

Train graders to provide feedback and guidance to student. In most undergraduate engineering courses, grading is done either by undergraduate graders (students who have previously taken the course and are paid to grade homework assignments for the current semester) or graduate teaching assistants. Often these graders are given an answer key or grading rubric and left to grade as they see fit. One improvement to this method would be to give the graders specific instructions on grading. Graders could be instructed to not only give point totals for a given problem, but also explain specifically where points were deducted and why they were deducted. The graders should also be instructed to comment on any solution method that was used incorrectly and suggest alternate methods for solving problems. These steps would likely increase the time required for scoring papers, but would serve in substantial ways to move the evaluation of student work much more toward assessment for learning.

Transforming student evaluation in undergraduate engineering programs

Benjamin Bloom (1984) is well known for his taxonomy, a hierarchical ordering of cognitive skills. Bloom's ranking, ordered from lowest to highest, includes knowledge, comprehension, application, analysis, synthesis, and evaluation. He logically showed that evaluation requires all the other cognitive skills. That well-established understanding of cognition presents a potent argument for teaching in ways that directly serve development of evaluation skills. We believe that direct instruction to enhance students' evaluation skills should be pursued in every core engineering course.

Consider the content of ABET's third criterion:

Engineering programs must demonstrate that their students attain:

a) an ability to apply knowledge of mathematics, science, and engineering;
b) an ability to design and conduct experiments, as well as to analyze and interpret data;
c) an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability
d) an ability to function on multi-disciplinary teams;
e) an ability to identify, formulate, and solve engineering problems;
f) an understanding of professional and ethical responsibility;
g) an ability to communicate effectively;
h) the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context;
i) a recognition of the need for, and an ability to engage in life-long learning;
j) a knowledge of contemporary issues; and
k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice (ABET, 2007, p. 2).

Attainment of those knowledge and ability elements depends heavily on students' evaluative skills. If engineering professors agree that students learning how to evaluate is integral to learning progress, then evaluation, feedback, and guidance practices in typical engineering classrooms require transformation. Standards F1–Practical Orientation and F3–Evaluation Support both strongly urge educators to use practical, feasible ways to serve evaluation practices. Can the content objectives of courses be reached while incorporating evaluation instruction into the mix? We think so. A wide array of assessment methods can be used for learning purposes. Some just require small modifications of current assessment tools. Below, we offer ways in which to develop students' evaluation skills and engage them more directly in the evaluation process in order to enhance their learning of the subject matter.

<u>Peer correction of homework assignment</u>. Students could benefit from correcting one another's homework assignments. In many engineering classes, there is already non-lecture classroom time that could be used for this activity in the form of the weekly section or recitation periods. The section leaders could solve the homework problems from the previous assignment in class (possibly showing multiple solution methods to each problem). Students would benefit by seeing how their peers attack problems and from being able to ask real-time questions of the section leader about the homework problems instead of seeing only one version of a solved problem in the form of the homework solutions. An alternative method is to form groups in which students are responsible for correcting their peers' homework assignments and giving feedback regarding the solution methods to each problem.

Note that the term "correcting" rather than "grading" student work was used in the above paragraph. This wording was used for two reasons. First, students grading their own or others work has long been a contentious issue. It is a matter that was heard and judged by the U.S. Supreme Court in 2002. If the purpose is assessment of learning, the student evaluation standard P7 (Conflict of Interest) provides strong arguments against it. The student's personal interest in getting a good grade or helping a peer obtain a good grade creates a substantial incentive for bias and outright cheating. But, if the intended purpose is to correct and provide feedback and guidance, then the prohibitions against student engagement diminish. In 2002, the Supreme Court found in favor of students' correcting one another's work principally to maintain this learning option—see Figure 1 for a "sidebar" discussion on this topic. Second, other research has shown that grading student papers works against the positive effects of feedback (Black et al., 2004). So, there are good reasons to engage students in *correcting* papers as part of the learning process and equally good reasons for *grading* to be excluded from this process.

<u>Engagement of students in development of grading criteria.</u> A key part of student learning must necessarily focus on setting criteria for what is a correct answer. Coupled with those criteria are process matters such as what/how much is required in order that a student's answer sufficiently demonstrates understanding of the concept and procedural steps needed to provide the correct answer. If the student cannot "justify" the answer, it doesn't matter what answer he or she has given—it is not sufficient. They must learn to evaluate so that they don't have to depend on the professor to tell them whether they are right or wrong.

This is not a trivial issue. Teaching students to effectively evaluate their own and their peers' work lies at the heart of sound education. Accurate evaluation is, in fact, the final determiner of whether you really know something. The abilities of students to observe, learn, and construct sound criteria often mean the difference between success and failure. Knowledge of facts is not enough. In an article titled, *The High Stakes in Science Education*, King (2007), who is a professor of microbiology at MIT, argues convincingly that some students are so fixated on meeting the professor's criteria that they cannot rely on their own judgments. His comments are made in the context of a biology laboratory course where "one of the modules involved students observing with a high-powered light microscope the development of a newly fertilized zebra fish egg into a baby fish" (p. 34). He notes that some students were so dependent on external approval that they could not draw the required figures: "They were unwilling to draw images that didn't correspond to 'the right answer'" (p. 34). Note the consistency of the issue King raised with our starting discussion of the elements of sound evaluation presented by Wiggins (2004). Black and Wiliam (1998) argue the same point:

> Thus self-assessment by pupils, far from being a luxury, is in fact *an essential component of formative assessment*. When anyone is trying to learn, feedback about the effort has three elements: recognition of the *desired goal*, evidence about *present position*, and some understanding of a *way to close the gap* between the two. All three must be understood to some degree by anyone before he or she can take action to improve learning (p. 6).

The U.S. Supreme Court ruled that students correcting other students' papers does not violate the Family Educational Rights and Privacy Act (FERPA) of 1974 (Owasso Indep. Sch. Dist. No. I011 V. Falvo). In declaring that the practice is legal, Justice Kennedy presented the Court's opinion and began his statement as follows: "Teachers sometimes ask students, including respondent's children, to score each other's tests, papers, and assignments as the teachers explain the correct answers to the entire class." Regarding the plaintiff's claims he stated, "Respondent claimed the peer grading embarrassed her children. She asked the school district to adopt a uniform policy banning peer grading and requiring teachers either to grade assignments themselves or at least to forbid students from grading papers other than their own." In finding for the school, the court addressed several points but ultimately based its ruling on a narrow point:

> For these reasons, even assuming a teacher's grade book is an education record, the Court of Appeals erred, for in all events the grades on students' papers would not be covered under FERPA at least until the teacher has collected them and recorded them in his or her grade book. We limit our holding to this narrow point, and do not decide the broader question whether the grades on individual student assignments, once they are turned in to teachers, are protected by the Act.

Before arriving at that conclusion, Justice Kennedy seems to border on addressing issues of appropriate instruction and evaluation at a couple of points. For example, he makes a preliminary statement:

> Correcting a classmate's work can be as much a part of the assignment as taking the test itself. It is a way to teach material again in a new context, and it helps show students how to assist and respect fellow pupils.

That statement might be interpreted to suggest the technique is educationally sound. Also, he stated,

> It would force all instructors to take time, which otherwise could be spent teaching and in preparation, to correct an assortment of daily student assignments. Respondent's view would make it much more difficult for teachers to give students immediate guidance. The interpretation respondent urges would force teachers to abandon other customary practices, such as group grading of team assignments.

This latter statement both alludes to this practice as common and suggests that employing students to score other students' work reduces the evaluation burden with potential attendant instructional benefits. Those two points touch issues related to propriety and feasibility of evaluations.

Figure 1. Supreme Court ruling on the legality of students correcting other students' work.

Suppose instructional faculty chose to focus student attention on developing criteria for scoring/evaluating student work. This could be achieved through a process in which students are grouped and work collectively to design scoring keys. (Black et al. [2003, p. 60] describe a variation of this idea at the secondary school level using student pairs to answer and defend answers to problems.) The professor or TA could then review those criteria (perhaps in concert with students) and choose either one set of criteria or an amalgamation of criteria from different groups (student groups would then be awarded merit on the quality of their criteria). Ostensibly, the resulting criteria would be identical to or consistent with what the professor had previously determined. Students, in turn, would be charged with checking answers and providing feedback to others. Work could be exchanged and rescored by another group. This multiple scoring could confirm quality of scoring and evaluative feedback; and attention could be given to identification of novel approaches, most succinct answers, most common errors, and the like. Examples of all these could easily be posted electronically to serve student learning. The professor's and TA's roles then could be turned to evaluating the evaluations, catching errors of the student-peer evaluators, searching for common feedback to improve instruction, providing guidance on how to evaluate, and deciding next steps to best serve student learning. Implications from the studies by Black et al. (2003) suggest these methods of forcing students to make their thinking explicit directly supports student learning and improves the professor's scaffolding for instruction (pedagogy) as well.

In this scenario, students engage directly and extensively in the evaluation process; teachers (professor and TAs) are also heavily engaged, but engage differently than in current practice. We argue that, in time, this suggested approach would require equal or less total time than current practices. Initially at least, we anticipate that professors and TAs also will spend more time developing and implementing these new processes. So, there are substantial trade-offs in taking such steps, but the work by Black et al. (2004) and others shows that it can lead to major improvements in student interest and learning. Studies reported in Black et al. (2003) suggest that when students perform these types of evaluative tasks, students' ability to help one another far outweighs the loss in accuracy of correcting papers. These efforts are not likely to be productive if done on an ad-hoc basis; they require careful planning and implementation to reap learning rewards.

Yet, it is not just professors who need to be convinced of the merits of such changes. One reviewer for this chapter stated, "I am not sure that turning the students into evaluators at the college level is feasible. It will require a lot of in and out of classroom time that they would rather spend on doing homework and studying instead of developing grading policies and grading each other's homework (especially multiple times)." As that response suggests, the issue quickly boils down to concerns for grades versus learning and for feasibility in terms of student time and interest.

<u>Role of the student evaluation standards.</u> Taking steps to transform evaluation in undergraduate engineering programs requires a shift in thinking for both professors and students. When evaluative information is to be used for accountability purposes

(assessment of learning), there is a need for confidentiality, security, and strong rules surrounding the evaluation process. Standards such as P4–Treatment of Students; P5–Rights of Students; U4–Evaluator Qualifications; A1–Validity Orientation; A4–Documented Procedures; A6–Reliable Information; A7–Bias Information and Quality Control; and A10–Justified Conclusions must be addressed with care. If professors are willing to lessen the grip of accountability to serve student learning, then attention shifts to some extent to other standards such as U3–Information Scope and A5–Defensible Information. Other standards important in grading practices, such as A1–Validity Orientation, would serve different purposes and lead to different actions. To this point, we have emphasized the importance of assessment for learning.

Conclusion

We believe that making the changes proposed here represents a huge paradigm shift for engineering students and faculty. It changes the status quo between student and instructors in profound ways. Changing the rules in these ways is more than most want to consider. Yet, Heifetz (2004) argues that the one way to make significant sustainable change is to get students to do more of the thinking. Similarly, Persaud and Freeman (2005) found that actively engaging students in shaping their learning is important to student retention and success in an engineering bridge program. We argue that engaging them in the evaluation process is at the heart of that thinking. We also agree with Fullan (2005) who has argued that these steps are neither easy nor a quick fix. Importantly, all these education scholars argue that this is not something one teacher or professor can do alone. It requires buy-in by the larger group and a change in school culture.

We think that producing impetus for the kinds of change described above is what ABET is about. Ultimately, we believe that ABET must be more explicit and provide more guidance on matters of assessment for learning. Professors and students need help in their evaluative efforts to foster student success in achieving program outcomes. Attention to evaluation in terms of assessment of learning is tightly woven into all the ABET criteria. We think the reader must look very closely to see ABET's interests in assessment *for* learning. We have demonstrated that *The Student Evaluation Standards* provide a sound basis for exploring evaluation practices and improving them. Through these explorations we have shown how the use of these standards provides insights into classroom evaluation practices and can lead to substantial improvements. Finally, we have argued that assessment *for* learning provides the necessary scaffolding for assessment *of* learning and will produce large dividends in student outcomes.

Acknowledgements

The authors are indebted to numerous individuals who have assisted us in preparation of this chapter. We especially note our appreciation for assistance provided by the following persons:

- the students at Brown University in EN31 who participated in our surveys;

- Professor Pradeep Guduru, EN31 instructor, and Brown University in general for allowing us to survey students;
- Ms. Lori Wingate for her extensive editing and organizational work to improve readability of this material; and
- Ms. Sally Veeder for her edits and corrections of our final drafts.

References

ABET. (2007). 2007-2008 *Criteria for accrediting engineering programs*. Baltimore: Author. Retrieved April 2, 2008, from http://www.abet.org/forms.shtml#For_Engineering_Programs_Only

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Black, P., Harrison, C., Lee, C., Marshall B., & Wiliam, D. (2002). *Working inside the black box: Assessment for learning in the classroom.* London: School of Education, King's College.

Black, P., Harrison, C., Lee, C., Marshall B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice.* Berkshire, England: Open University Press.

Black, P., Harrison, C., Lee, C., Marshall B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 9–21.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Bloom, B. S. (1984) *Taxonomy of educational objectives*. Boston: Allyn and Bacon.

Brown University. (2007). *Brown University online course announcement and Web syllabi* (BOCA). Retrieved April 2, 2008, from http://boca.brown.edu/

Fullan, M. (2005). *Leadership & sustainability.* Thousand Oaks, CA: Corwin Press.

Gullickson, A. R. (1984). Teacher perspectives of their instructional use of tests. *Journal of Educational Research, 77*(4), 244–248.

Gullickson, A. R. (1985). Student evaluation techniques and their relationship to grade and curriculum. *Journal of Educational Research, 79*(2), 15–18.

Gullickson, A. R. (1986). Teacher education and teacher-perceived needs in educational measurement and evaluation. *Journal of Educational Measurement, 23*(4), 347–354.

Gullickson, A. R. (1993). Matching measurement instruction to classroom-based evaluation. *Educational Assessment*. Lincoln, NE: Buros Institute.

Haertel, E. (1986, April). *Choosing and using classroom tests: Teachers' perspectives on assessment.* Paper presented at the annual meeting of the American Educational Research Association.

Hake, R. (2006). Possible palliatives for the paralyzing pre/post paranoia that plagues some PEP's. *Journal of MultiDisciplinary Education, 6,* 59–71. Retrieved on April 20, 2007, from http://evaluation.wmich.edu/jmde/JMDE_Num006.htm.

Heifetz, R. (2004, February). *The adaptive challenge*. Quoted in M. Fullan, (2005). *Leadership & sustainability,* p. 56. Thousand Oaks, CA: Corwin Press.

Hill, P., & Crevola, C. (2003). Organizational learning. In B. Davies & J. West-Burnham (Eds.). *Handbook of educational leadership and management* (pp. 394–403). Harlow, UK: Pearson Education.

Huba, M. E. & Freed, J. E. (1999). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning.* _________: Allyn & Bacon. .

Joint Advisory Committee. (1993). *Principles for fair student assessment practices for education in Canada.* Edmonton, Alberta: Author. Retrieved on April 2, 2008, from http://www.bced.gov.bc.ca/classroom_assessment/fairstudent.pdf

Joint Committee on Standards for Educational Evaluation. (2003). *The student evaluation standards.* Thousand Oaks, CA: Corwin Press & Educational Policy Leadership Institute.

King, J. (2007, May). The high stakes in science education: Risking the roots of American productivity. *Education Week,* 26(36), 34, 44.

National Council of Examiners for Engineering and Surveying (NCEES). (2007). *Fundamentals of engineering exam* available at http://www.ncees.org/exams/fundamentals/

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

Persaud, A., & Freeman, A. (2005). Creating a successful model for minority students' success in engineering: The PREF summer bridge program. Proceedings of the 2005 WEPAN/NAMEPA Joint Conference, ___________..

Stiggins, R., & Conklin, N. (1988, November). *Teacher training in assessment.* Portland, OR: Northwest Regional Educational Laboratory.

Stiggins, R., Griswold, M., & Wikelund, K. (1989, Fall). Measuring thinking skills through classroom assessment. *Journal of Educational Measurement,* 26(3), 233–246.

Stiggins, R. J. (2001). *Student-Involved Classroom Assessment* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan,* 86(1), 22-27.

Wiggins, G. P. (2004). *Assessment as feedback.* Retrieved on April 2, 2008, from http://www.newhorizons.org/strategies/assess/wiggins.htm

Appendix A

Survey 1 Items. The *Student Evaluation Standards* and associated Guidelines and Common Errors used to create the items are identified in brackets for each item.

First Survey Administered in __________

Survey 1: Teaching Evaluation in Engineering

Please answer the questions below to the best of your ability. For the purpose of this survey, evaluations refer to homeworks, labs and examinations. Answer the questions based on information given to you in class, during section, by a TA or a Professor, via e-mail or the course website.

1. ____Yes ____No Was the educational purpose/goals of this class (as outlined in BOCA or according to the ABET standards) communicated to you? [ABET Criterion 1]

1a. If the answer to (1) was yes, please note where you obtained this information. (Circle all that apply)

syllabus website lecture section Professor TA other (specify)

2. ____Yes ____No Has the instructor informed you what the purposes and uses of the evaluations in this class will be? [P1, B]

2a. If the answer to (3) was yes, please note where you obtained this information and what you understand the purpose of the evaluations (homework, labs, exams) to be. (Circle all that apply)

syllabus website lecture section Professor TA other (specify)

Purpose of the evaluations:

3. ____Yes ____No Do you feel that the homeworks so far have been based upon the information presented to you in lectures? [P1, D]

4. ____Yes ____No Were you informed of alternate procedures of evaluation for students with disabilities (including learning disabilities)? [P1, G]

[Section 2 of the Survey]
Were the following policies made clear to you or do you know where you could find the information related to those policies? (N/A means not applicable to you).

1. Do you know how disagreements between a student and the professor/TA on evaluation matters will be resolved? (Please check all that apply) [P2, B]
    - ______ yes, I know the information
    - ______ no, I do not know the information
    - ______ yes, I know where to find the information
    - ______ no, I do not know where to find the information
    - ______ N/A

2. Do you know how scores will be aggregated and final grades determined? (Please check all that apply) [P2, B]

_____ yes, I know the information
_____ no, I do not know the information
_____ yes, I know where to find the information
_____ no, I do not know where to find the information
_____ N/A

3. Do you know how the results of evaluations will be reported to you? (Please check all that apply) [P2, B]

_____ yes, I know the information
_____ no, I do not know the information
_____ yes, I know where to find the information
_____ no, I do not know where to find the information
_____ N/A

4. Do you know the policy on plagiarism? (Please check all that apply) [P2, H]

_____ yes, I know the information
_____ no, I do not know the information
_____ yes, I know where to find the information
_____ no, I do not know where to find the information
_____ N/A

5. Do you know the policy on use of other students' work? (Please check all that apply) [P2, H]

_____ yes, I know the information
_____ no, I do not know the information
_____ yes, I know where to find the information
_____ no, I do not know where to find the information
_____ N/A

6. Do you know the policy on sharing of information (i.e., working together on homeworks, labs and/or exams)? (Please check all that apply) [P2, H]

_____ yes, I know the information
_____ no, I do not know the information
_____ yes, I know where to find the information
_____ no, I do not know where to find the information
_____ N/A

7. Do you know what constitutes cheating in this class? (Please check all that apply) [P2, H]

_____ yes, I know the information
_____ no, I do not know the information
_____ yes, I know where to find the information
_____ no, I do not know where to find the information
_____ N/A

8. _____Yes _____No
Do you know if any of these policies were established after your first homework? [P2, Common Errors]

9. If the answer to (8) is yes, please give an example of a policy that was established after your first evaluation, or list the number of that policy from the question above. [P2, Common Errors]

Appendix B

Survey 1 Items. The *Student Evaluation Standards* and associated Guidelines and Common Errors used to create the items are identified in brackets for each item.

Second Survey Administered in Fall 2006

Survey 2: Teaching Evaluation in Engineering

Please answer the questions below to the best of your ability. For the purpose of this survey, evaluations refer to homework assignments, labs and examinations given in EN31. Your answers should be based on information given to you in class, during section, by a TA or a Professor, via e-mail or the course website. For each item note your answer as Yes or No by placing an X at the appropriate location. Please answer all items.

1. ____Yes ____No Do you know what your rights are in an evaluation (i.e., do you know if you have the right to see your graded homework, the conditions under which you are entitled to an alternate evaluation, how to appeal the results of the evaluation, etc.)? [P5, B]

2. ____Yes ____No Do you believe that the evaluators (Professor or TA) have kept the results of your evaluations private? [P5, Common Errors]

3. ____Yes ____No In class lectures do you believe that the Professor(s) teach specifically to the evaluation (i.e., teach such that the questions posed in a given homework or exam are readily answered but general concepts are lost)? [A1, A]

4. ____Yes ____No In sections do you believe that the Professor(s)/TA(s) teach specifically to the evaluation (i.e., teach such that the questions posed in a given evaluation are readily answered but general concepts are lost)? [A1, A]

5. ____Yes ____No Have you been involved in the evaluation process (i.e., grading your own paper, grading a classmate's paper, etc.)? [P1, E]

6. ____Yes ____No Overall, do you believe that the evaluation results (so far, the results of your graded homeworks) properly represent your achievement and understanding of the core concepts in this course? [P1, D]

7. ____Yes ____No Are you aware of any conflict of interest on the part of an evaluator (Professor or TA) that might cause a bias in the scoring or grading of your work? [P7]

8. ____Yes ____No Do you believe that the content of the assessments (homework assignments) are balanced, that is that they are consistent with what the Professor has emphasized in class? [P1, D; A1, A]

Note: The survey instruments included in this chapter are copyrighted by the chapter authors. If the surveys are reproduced and disseminated, their origins, *The Student Evaluation Standards*, should be cited.

Appendix C

Third Survey Administered in Spring 2007

Survey 3: Teaching Evaluation in Engineering

For each of the following descriptors please rate your instructor for the course:

Spring Semester 2006

Your responses will be treated confidentially. All information provided will be reported only in aggregate form. Individual names of students, instructors, or courses will not be reported in connection to any of the information obtained or reported on this form.

This information is requested to help us learn more about teaching practices and will be used as examples of student perceptions of teaching practice.

The survey consists of a series of statements about assessment practices. For each statement you are asked to respond for three types of student assignments: homework, laboratory exercises, and project. Your job is to choose your level of agreement with each statement, for each type of assignment, based on your perception of the instructor's (professor's) performance on that assessment practice from the drop down menu. Your response for each type of work should be provided in the designated column.

A glossary of terminology with examples is located at the end of the survey in case some of the terms used are unfamiliar to you.

For each type of work respond to each statement with your judgment where:

SD = Strongly disagree
D = Disagree
? = Neither Disagree nor Agree
A = Agree
SA = Strongly Agree

You should select one of these choices from the pull down menu in each of the three columns.

Learning outcomes/objectives refer to the objectives set forth in the course description, which is quoted below:

"Mechanical behavior of materials and analysis of stress and deformation in engineering structures and continuous media. Topics include concepts of stress and strain; the elastic, plastic, and time-dependent response of materials; principles of structural analysis and application to simple bar structures, beam theory, instability and buckling, torsion of shafts; general three-dimensional states of stress; Mohr's circle; stress concentrations."

For example, a learning objective would be an understanding of the concepts outlined in the course description and covered in class, i.e. in the case of beam theory the learning objective would be that the student understands how to create a shear/bending moment diagram and how the moments are related to stresses and deflections. A course outcome would be the ability to complete a problem which involved calculating a 3D stress state on a beam under a given loading (i.e., find shear and normal stresses caused by axial loads, bending and torques), or the ability to calculate the critical buckling load for a beam under a set of given constraints.

Descriptors of the Instructor's (Faculty Member or TA) Assessment Practices	Homework SD D ? A SA	Laboratory SD D ? A SA	Project SD D ? A SA
Learning objectives were clearly specified.			
The instructor provided examples for what was expected in terms of learning outcomes.			
When I began to work I understood the conditions which had to be met in order to achieve learning objectives.			
The instructor's feedback on my completed work concretely described it in relation to the expected learning goals.			
Instructional feedback provided to me described the results of my work in real terms of my progress toward the learning objectives.			
The instructional feedback on my work described my performance in terms of the explicit or implicit learning objectives of the assignment.			
In evaluating my work, the instructor made value judgments about the facts of my work and their effects in terms of reaching the intended learning objectives.			
The instructor did not praise or blame my work in light of the instructor's expectations for me.			
When the instructor did provide praise or blame regarding my work it was not preceded or followed by feedback.			
Information provided to me regarding this assignment did not include advice about what to do based on the feedback.			
Information provided to me regarding my work did show me how to re-direct my learning in light of results.			

Here is a glossary of terms with examples which may be useful in answering the questions posed in this survey. Refer to this as you are taking the survey if you are unclear about the meaning of any terms or the meaning of a question. This example is given in the context of an introductory dynamics course, but you should be able to apply the concepts to your course. The course description of the dynamics class is given as the following:

Study of the kinematics and dynamics of particles and rigid bodies. Principles of motion of mechanical systems. Concepts of inertia, work, kinetic energy, linear momentum, angular momentum, and impact. Applications to engineering systems, satellite orbits, harmonic vibrations of one and two degree of freedom systems. Lectures, recitations, and laboratory.

Glossary

learning goals—a statement of what students are expected to learn in a given lesson, unit, course, program, or across educational and training programs.
Example: Principles of motion of mechanical systems.

explicit—when the information or criterion is directly stated verbally in writing

implicit—when the information or criterion is expected to be understood from general comments or conditions of the learning situation.

assignment—a task or amount of work assigned or undertaken by the student at the direction of the teacher.
Example: Weekly homework or laboratory reports.

value judgments—an estimation or a measure of the merit and/or worth of a student's work in terms of the intrinsic qualities of the individual student's progress toward reaching a stated objective.
Example: Grade received on a given homework as it relates to the total grade in the course.

learning objectives—the learning planned to be developed within the lecture lesson, or assignment.[1]
Example: Using work energy with non-conservative forces (friction) to determine the velocity or acceleration of a body in motion.

learning outcomes—how one will recognize the accomplishment of the planned learning.[1]
Example: Given the following problem, one is able to draw the free body diagram, and determine the solution of the problem:

> A cylinder at the top of a hill of height H and length L is released, at the bottom of the hill is a flat path. The coefficient of rolling friction is μ_r. Find the velocity of the cylinder at the bottom of the hill, and the distance traveled by the cylinder before it comes to rest in terms of H and L.

feedback—information the learner receives about current work that describes the extent to which the work meets explicit or implicit specifications of the learning objective.
Example: Michael's returned homework had feedback in the form of comments made by the grader. The grader's comments included an explanation of what Michael had done right and wrong in the problems he had not gotten full credit on.

modeling—using exemplars to illustrate assessment practices so that pupils are able to judge the standard of their work.[1]
Example: During section, the instructor did sample problems like those given in the homework assignments.

Note: The survey instruments included in this chapter are copyrighted by the chapter authors. If the surveys are reproduced and disseminated, their origins, *The Student Evaluation Standards*, should be cited.

CHAPTER 4

ASSESSING THE EDUCATIONAL BENEFITS OF EPA'S P3 AWARD—A NATIONAL STUDENT DESIGN COMPETITION FOR SUSTAINABILITY[1]

Estella Waldman and Cynthia Nolt-Helms
U.S. EPA, Office of Research and Development

Julie Zimmerman and Elizabeth Dunford
Yale University

Robert Yackee
Cornell University

Linda Vanasupa
California Polytechnic State University – San Luis Obispo

Introduction

There are 1.4 million engineers working in the United States today (Bureau of Labor Statistics, 2006) in a variety of fields including health care, agriculture, and the environment. Over the past century, engineers have made breakthroughs in supplying clean drinking water, building highways to link the country, and developing resources to meet society's expanding energy needs. Today's engineering students face unique challenges that require them to blend new technologies with current environmental concerns in both the developed and developing worlds. As global population increases and the availability of natural resources decreases, engineers will have to design more sustainable systems and develop technologies that consider the social and economic impacts of engineering choices in both a national and global setting.

To prepare for their careers, today's students complete a rigorous course of study that includes classes in their engineering field of choice as well as classes in mathematics, physics, and computer sciences. Key to assuring that colleges and universities are setting high academic standards—and that students are meeting these standards—is

[1]DISCLAIMER: The views expressed in this chapter are those of the individual authors and may not necessarily reflect the views and policies of the United States Environmental Protection Agency (EPA). EPA, through its Office of Research and Development, funded and managed the program described here. EPA program managers have prepared the EPA sections, and those sections have been reviewed in accordance with EPA's review policies and approved for publication.

ABET, a federation of 28 societies that accredits colleges and universities in applied science, computing, engineering, and technology. Founded in 1932, ABET is celebrating 75 years as the nationally-recognized accreditation agency in these fields. The organization monitors 2,700 programs in over 550 institutions and considers a diverse set of requirements that include quality of students and graduates, educational objectives of the program, curricula, and faculty qualifications (ABET, 2007).

Even a cursory reading of the "Criteria for Accrediting Engineering Programs Effective for the 2007–2008 Accreditation Cycle" (ABET, 2007) reveals the emphasis that ABET places on experience. Rather than just learning math and science, students must be able to apply their lessons, conduct experiments, and use their talents to solve problems.

ABET's Criterion 4 requires a culminating design experience where students incorporate appropriate engineering standards and multiple realistic constraints. Criterion 3(c) includes *environmental, social, political, ethical,* and *sustainability* among the list of realistic constraints to be considered (ABET, 2007).

ABET's emphasis on practical, real-world, trial-and-error experimentation makes it clear why student design competitions are held in high regard in the engineering world. Each year, thousands of students participate in competitions, performing tasks such as building sea-worthy concrete canoes, constructing structurally-sound steel bridges, or creating robots to help wheelchair-bound individuals with everyday chores. After more than 20 years, the concrete canoe competition is a perennial favorite with students who spend a year building their canoes, often with the help of faculty advisors. Undergraduate students experiment with different combinations of cement, water, and various aggregates to build vessels that will allow them to stay afloat while racing and take them from regional competitions to the national race. Judged on four factors—their design paper, presentation, the canoe itself, and their standing in the race—students gain an understanding of what it means to follow a proposal from nascent idea to completed project. Sometimes called "America's Cup of Civil Engineering," the competition rewards winning teams with scholarships, trophies, and bragging rights.

As the fine tradition of learning-through-doing continues in engineering in the 21st century, members of the community have become increasingly interested in developing critical thinking skills in the context of broader social and environmental issues. In fact, ABET called for a focus on environmental sustainability when revising its accreditation requirements in *Engineering Criteria 2000* (as cited in Lattuca, Terenzini, & Volkwein, 2006). As part of the professional component, ABET demands that students be taught economics, ethics, and the value of sustainability in addition to traditional engineering skills.

A New Approach: EPA's People, Prosperity and the Planet (P3) Award

The concept of sustainable development became widely promoted following publication of G. Bruntland's *Our Common Future* in 1987 by the World Commission on Environment and Development. Although the definitions of sustainability have varied during the past 20 years, a useful definition for engineers was set forth by Mihelcic et al. (2003): "Design of human and industrial systems to ensure that humankind's use of natural resources and cycles do not lead to diminished quality of life due either to losses in future economic opportunities or to adverse impacts on social conditions, human health and the environment." Fundamentally, sustainability requires the balancing of economic prosperity, environmental responsibility, and social fairness. Coincidently, Criterion 3(c) requires engineering students to demonstrate an ability to design within a wide variety of realistic economic, environmental, and social constraints—the three pillars of sustainable development. Further, Criterion 3(h) requires that engineering students attain the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context. These criteria fit well with the overall goals of P3 and what students gain from participation in the program.

The U.S. Environmental Protection Agency (EPA) has incorporated these principles of sustainable design into its P3 Award competition, launched in 2004. The program fosters future scientists, engineers, architects, business and marketing professionals, and others in advancing the principles of sustainability through technology innovation in the areas of water, energy, agriculture, information technology, materials and chemistry, and the built environment. Through grant awards to teams of undergraduate and graduate students and their faculty advisors, P3 teams design and develop sustainability projects in the above areas. The competition also includes a requirement for the integration of sustainability into higher education curricula.

> The EPA's P3 Award program has the goal and potential to reach out to thousands of colleges and universities across the country and transform the way we prepare today's workforce to meet tomorrow's challenges. To learn more about the P3 Award competition and the National Sustainable Design Expo, visit www.epa.gov/p3 . EPA requests applications for the P3 Award program each year from August to December. If you would like to become involved in the program, please contact Cynthia Nolt-Helms, P3 Award program coordinator, at nolt-helms.cynthia@epa.gov.

Planning for the future is a critical aspect of sustainability. For the body of creative technology to advance, it is important to train future generations of engineers and scientists to value sustainable solutions. However, academic curricula in science and engineering are primarily structured along traditional lines and offer only a small number of courses that discuss sustainability. Within the broader field of engineering education, there is growing recognition that those who seek to develop solutions to sustainability issues must have the ability to collaborate with people who have different expertise, perspectives, and levels of schooling. The successful application of sustainability ideas and principles in the engineering profession will also require engineers who are capable of critically assessing the social and economic implications of their projects. Huntzinger, Hutchins, Gierke, and Sutherland (2007) argue that students will need these skills to supply creative solutions to complex problems in the face of ambiguity and potentially conflicting goals. The P3 program provides a mechanism for students who are rigorously trained in the fundamentals of science and engineering to undergo more problem-based learning (Huntzinger et al., 2007) that will help develop these skills. The problems addressed by the P3 design projects are real, and the structure of the program requires them to consider the impacts they will have on the economy, society, and the environment.

To launch the P3 program, EPA brought together more than 40 partners from the federal government, industry, and scientific and professional societies. Designed as a two-phase program, P3 encourages student teams to apply initially for $10,000 grants to research and develop their projects during the academic year. In Phase II, P3 teams are invited to Washington, DC, to compete for the P3 Award at the National Sustainable Design Expo (the Expo) (EPA, 2007). An expert panel of engineers judges the designs, and EPA chooses six P3 Award winners who are eligible for up to $75,000 to implement their designs or move them to the marketplace. EPA also recognizes other projects through honorable mentions that do not receive additional funding.

The Expo, held each spring around Earth Day on the National Mall in Washington, DC, was created to bring together scientists, engineers, and business leaders focused on sustainability issues and solutions. The design projects exhibited by the P3 student teams are the essential core of the Expo. At this event, students are given an opportunity to showcase and explain their projects to the P3 judges, other teams and exhibitors, federal agency officials, and the general public.

In the short history of the P3 program, 99 universities from 40 states and Puerto Rico have participated in the competition (see Figure 1). The majority of participating students are from engineering schools, although nine liberal arts colleges have been represented to date. In 2004, the first year of the design competition, 65 P3 teams were chosen to compete for the P3 Award out of a total of 116 applicants. In 2005 and 2006, 64 and 87 teams applied, respectively, and 41 teams competed each year.

Albion College	New Jersey Institute of Technology	University of Cincinnati
Appalachian State University	New Mexico Institute of Mining and Technology	University of Colorado at Boulder
Arizona State University - Main Campus	New Mexico State University - Main Campus	University of Colorado at Denver
Ball State University	North Carolina State University	University of Connecticut
Brown University	Northern Arizona University	University of Florida
California Polytechnic State University - San Luis Obispo	Northern Illinois University Urbana-Champaign	University of Illinois at
California State Polytechnic University - Pomona	Northwestern University	University of Iowa
California State University - Chico	Oberlin College	University of Kentucky
Carnegie Mellon University	Ohio Northern University	University of Maine - Machias
Clarkson University	Ohio State University	University of Maryland
Colorado State University	Oklahoma State University Baltimore County	University of Maryland -

Columbia University	Oregon State University	University of Massachusetts - Amherst
Cornell University	Pennsylvania State University	University of Massachusetts - Boston
Drexel University	Portland State University	University of Massachusetts - Lowell
Duke University	Prescott College	University of Miami
Duquesne University	Rensselaer Polytechnic Institute	University of Michigan - Ann Arbor
East Carolina University	Rochester Institute of Technology	University of Missouri - Rolla
Eastern Illinois University	Rutgers, the State University of New Jersey	University of Nevada - Reno
Evergreen State College	Santa Clara University	University of New Hampshire
Federal University of Vicosa	Southern Illinois University Edwardsville	University of North Carolina at Chapel Hill
Gonzaga University	Southern University	University of Pittsburgh - Main Campus
Harvard University	Stanford University	University of Tennessee - Knoxville
Illinois Institute of Technology	Stevens Institute of Technology	University of Texas at Austin
Iowa State University	Texas A & M University	University of Texas at El Paso
Ithaca College	Trinity University	University of Tulsa
Lafayette College	Tufts University	University of Utah
Lehigh University	Tulane University of Louisiana	University of Virginia
Louisiana State University - Baton Rouge	University of Alabama at Birmingham	University of Wisconsin - Green Bay
Macalester College	University of Arizona	Villanova University
Massachusetts Institute of Technology	University of California - Berkeley	Virginia Wesleyan College
Miami University - Oxford	University of California - Davis	Washington State University
Michigan Technological University	University of California - Riverside	Wentworth Institute of Technology
Middlebury College	University of California - Santa Barbara	Western Washington University

Figure 1. P3 colleges and universities, 2004–2007.

P3 teams are encouraged to be interdisciplinary, and participation in the P3 program provides universities with a tool that encourages students to collaborate with other undergraduates and graduate students, working professionals, and sometimes high school students. At many schools, not only did undergraduate P3 team members collaborate with students from a variety of engineering majors, but also across different schools within the universities. This aspect of the program corresponds with ABET Criterion 3(d), which encourages students to develop an ability to function on a multidisciplinary team. For example, the team at Gonzaga University included students from the schools of broadcasting, computer science, civil engineering, mechanical engineering, and nursing. At the University of Colorado, the team consisted of a group of students and faculty from multiple engineering and science disciplines with B.S., M.S., and Ph.D. degrees. Stanford University's green dorm project team included students in public policy, civil and mechanical engineering, and music. This form of collaboration benefits all team members, giving them experience working with others who have different expertise and perspectives. Finally, in many cases, the P3 project offers students an opportunity to collaborate with working professionals. For example, the team from the California Polytechnic Institute closely collaborated with professionals in the computer and electrical engineering fields. As stated by a Duke University faculty advisor (D. Schaad, personal communication, April 6, 2007), "By coupling together multiple classes to examine and attack this project, the educational value and impact were multiplied exponentially."

Assessing Student Performance in the P3 Award Competition

All Phase I applications for the P3 Award competition are received, processed, and initially reviewed by the EPA. Applications are then evaluated by an independent, technical peer-review panel composed of scientific and engineering experts assembled by the EPA who are not EPA employees. The review panel is chosen for their skills and expertise, and all members of the panel undergo a rigorous review to eliminate issues related to conflict of interest.

Each proposal for a P3 design project is required to follow the following criteria:

- Articulate the challenge it proposes to address and detail its relationship to sustainability (people, prosperity and the planet);
- Define the innovation and technical merit associated with the project,
- Demonstrate an ability to measure outcomes through an effective evaluation method and implementation strategy; and
- Discuss the use of the P3 competition as an educational tool.

Phase I funding of up to $10,000 is provided to all projects that are recommended by the technical peer review panel and pass a review for relevance conducted by EPA staff. To determine which projects will win a Phase II P3 Award, all teams are required to submit a written report prior to the Expo that addresses the above criteria for their completed Phase I activities. They must also submit a proposal detailing how they would use the additional funding provided by the Phase II award. All P3 teams must attend the National Sustainable Design Expo in Washington, DC, where they are given an opportunity to exhibit and demonstrate their technologies. During the Expo, each team is interviewed by judges who use the teams' written reports and interview results centered around the above criteria to evaluate and rank each project. Judges are chosen for their engineering expertise and, traditionally, they have been members of the National Academy of Engineering.

Success Stories

The P3 program has proven to be more than academic with several projects evolving into small businesses that are providing environmental solutions around the world. Some examples follow.

- A P3 team from Oberlin College designed a system for colleges to monitor energy and water use with easy-to-read, real-time data on energy at the dormitory level, allowing the school to reduce energy costs by pinpointing areas of overuse. This project is now a small business called Lucid Design Group (2007).

- P3 Award winners from the University of California at Berkeley developed a technology to disinfect contaminated water in homes. The team has been working with the Mexican Institute of Water Technology, the National Council for the Promotion of Education in Mexico, and Haiti Outreach: Pwoje Espwa (HOPE) in Haiti, to bring clean water to rural communities. The project has won several additional awards, including the Massachusetts Institute of Technology IDEAS International Technology Prize, for its innovative design (University of California at Berkeley, 2007).
- A Rochester Institute of Technology team looked at how solar ovens could be mass-produced at low cost in Latin America using local resources. These ovens reduce wood consumption and deforestation while providing local jobs. This project has been successfully implemented and evaluated in Venezuela, with plans to expand to other communities in South America (Rochester Institute of Technology, 2007).
- University of Michigan students have developed and tested a new product design concept called SITumbra using bio-composite materials to form passive, low-energy, load-bearing façades in buildings (SITumbra, 2007). Their innovative assembly concepts optimize on the unique, environmentally beneficial properties of these materials, which make them both strong and durable. The product is being prototyped for the construction industry market and is set to revolutionize design and construction methods towards more sustainable buildings.

Mechanisms to Integrate P3 into College and University Curricula

Implementation of the P3 design project as an educational tool is one of the four key criteria used to evaluate all proposals for funding. Therefore, colleges and universities have integrated the P3 Award into their curricula in a variety of unique ways. Some have sent students to developing countries to educate and serve local communities. Other students have presented their projects at local conferences and professional seminars to educate members of the community. In some cases, P3 students worked with economically disadvantaged communities to identify sustainable solutions for specific problems. The Colorado State University team, working with a new alliance of organizations, universities, and Pine Ridge Reservation residents, committed to improving the living conditions on the reservation by designing culturally appropriate, sustainable housing.

Many educators have used the P3 design project as the focus of a core class or an alternative to a required course, such as the focus of a senior capstone design course. This has been very effective at universities such as at the University of Tennessee, where P. Frymier, the team's faculty advisor, "found it highly motivating for students to receive credit for the work as an alternative to a required class" (personal communication, April 6, 2007). The biochemistry and chemical engineering students enrolled in Frymier's class focused on using photosynthetic algae to produce molecular hydrogen and designed a biohydrogen facility for their P3 design project. In addition, they constructed a small-scale photobioreactor system to demonstrate the concept. By participating in this project, all of the students received credit for a required course in their respective majors.

Other universities have found it effective to offer an elective class centered on the P3 project. The University of Virginia offered a 6-credit studio in 2006 focused on the P3 project, and one student found it to be "the best studio I have had at UVA. It is exciting and sometimes frustrating to work on a 'real life' project, but always rewarding" (P. Richbourg, personal communication, April 6, 2007). In this course, an interdisciplinary student team designed and fabricated, for kindergarten to 12th graders, a floating environmental education field station or "barge" powered solely by site-based solar and wind energy systems.

Oklahoma State University also offered an elective in 2006 where students worked on the P3 project as a requirement for a course entitled Sustainability Issues in Architecture. The course introduced the concept of environmental sustainability and calculation methods for energy consumption and culminated in the P3 design project. The professor found this to be a very effective learning tool because "the objective of the P3 Award project fits perfectly within the scope of the course" (J. Clausing, personal communication, April 6, 2007).

At Cornell University, students who participated in the P3 project in 2006 were enrolled in an independent study course for engineering credit. The students who participated in the P3 project did not receive any classroom instruction, but they all received engineering credit.

While many universities included the P3 project as a component of an existing course, or designed courses specifically around the P3 project, the results of a P3 project occasionally produced new courses or significantly affected the syllabi of existing courses. At the University of Colorado Health Sciences Center in Denver, there is now a course entitled Advanced Concrete Materials which provides "students the opportunity to design, batch, and test pervious concrete filters" as a direct result of a P3 project completed in 2007. As a result of a P3 project at the California Polytechnic State University, A. Kean added information about biodiesel to his mechanical engineering course entitled Energy Conversion. Because the P3 competition offers an opportunity for creativity at many levels and in many forms, universities are able to use the P3 project as more than an extracurricular activity by incorporating it into various types of course work.

Using P3 to Educate the Community

Through a wide variety of methods, many universities have turned the P3 project into much more than a design competition, further benefiting the participating students, those around them, those with whom they work, and those in need. Through demonstrations of P3 design projects on university campuses, many P3 teams found ways to interact with students who did not directly participate in the projects. At Albion College, the P3 team designed and installed in a residence hall several types of exercise equipment that, when used, generated electricity. This exposed other students to environmentally sensitive energy habits. The team made available environmental literature and educational posters that described energy use and Energy Star® equipment. Monthly competitions were held to see who could generate the most energy; and students pledged to only use as much energy on the last Saturday of the month as they had generated on the equipment in that month. Other teams influenced their peers through on-campus demonstrations and displays, guest lectures, and publications in campus periodicals.

In addition to educating other students, many P3 teams found ways to educate the community around them. After the University of Connecticut team completed their P3 project in 2007, data and economic feasibility assessments were presented at the Biofuels Consortium at the University of Connecticut. Over 200 farmers, scientists and politicians participated in the seminars. Similarly, members of the University of New Hampshire team, whose P3 project resulted in a trip to West Africa to implement a sustainable water extraction system, also visited a local classroom at Oyster River Middle School to talk about the implementation trip and brainstorm ideas to involve middle school students in the project. J. Jambeck, faculty advisor, said,

> Additionally, the P3 team gave presentations to the New Hampshire chapters of the American Society of Civil Engineers and the Society of American Military Engineers. The team also plans to speak at the New Hampshire Professional Engineers Association and a New Hampshire Department of Transportation conference in 2007. Through these speaking engagements and interactions, the students have also educated and enriched the lives of the practicing engineers in New Hampshire (personal communication, April 6, 2007).

Many P3 teams have gone far beyond educating their local communities and used their P3 grants to help those in need of new technologies. Students gain fieldwork experience through the implementation process while also providing valuable community service. In 2006, University of Illinois students participating in the P3 competition traveled to Orissa, India, where they began selling—at an affordable price—the solar lanterns they developed for their P3 design project. This successful project is ongoing, and even the students themselves have been surprised at how valuable the lanterns have been for the local community. Students from Lehigh University traveled to West Bengal, India, to install a system that removes and safely disposes of arsenic from drinking water. Since the system is very effective, they expect it to have significant impacts on local health around the region. Students from the University of New Hampshire traveled to Niger, where they installed a sustainable water extraction system that yields a higher output of water than traditional pumps without relying on gasoline. Many other P3 teams—including those from Columbia University and the University of Florida—have traveled to third world countries to improve local health or quality of life by implementing their designs.

Assessing the EPA's P3 Program

The mission of the EPA is to protect human health and the environment. To that end, two important agency goals are sustainability and educating the next generation of engineers and scientists. The EPA's P3 program could be examined at many levels, including use as a potential model for those who seek to expand and/or improve sustainability education in engineering curricula. The discussion below describes the possible utility of a program such as P3 as learning experiences for the complex cognitive, social, and affective development of both engineering and non-engineering students.

There has been increasing articulation of the need for future engineers to have the type of training promoted and enabled by the P3 Award competition. For example, the American Society of Civil Engineers' Code of Ethics (2004) states that "The Code of Ethics of ASCE requires civil engineers to strive to comply with the principles of sustainable development in the performance of their professional duties . . . [including] global leadership in the promotion of responsible, economically sound, and environmentally sustainable solutions that enhance the quality of life, protect and efficiently use natural resources." ABET is also acting as a driving force to move beyond traditional engineering assessments. In the new ABET EC 2000 there are criteria based on a school's ability "to demonstrate that their students attain an ability to design . . . to meet desired needs within realistic constraints such as *economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability* . . .to understand the impact of engineering solutions in a *global, economic, environmental, and societal context*" (ABET, 2007). This evolution in the engineering community towards sustainability also suggests the need for current assessment methodologies to evolve.

Traditional assessments of student learning and aptitude have been based on measures such as grades on homework and exams, ABET standards, and peer evaluation. However, these traditional assessments may not easily or readily capture the lessons of sustainability in terms of values, ethics, and behavior. Moving beyond these traditional assessments has prompted a search for new approaches that assess the students' broader development that occurs in "significant learning" as L. D. Fink defines it (2003). Fink defines a "significant learning" experience as one that has both a "process" and an "outcome" dimension and each of these dimensions has two features. The process by which significant learning occurs is one in which the students are engaged in their learning and the class/student exhibits a high energy level. The outcome dimension is a learning experience (a) that results in a significant and lasting change in the students—changes that continue beyond the duration of the course—and (b) where what is learned has a high potential value in their lives after the course is over by enhancing their individual lives, preparing them to participate in multiple communities, or preparing them for the world of work.

To better capture what is necessary for significant learning to occur, Fink built upon previous and highly respected work on taxonomies of educational objectives to develop Fink's Taxonomy. He defined learning in terms of change, purporting that for learning to occur, there must be change in the learner. He identifies six categories of significant learning: foundational knowledge, application, integration, human dimension, caring, and learning how to learn (see Figure 2).

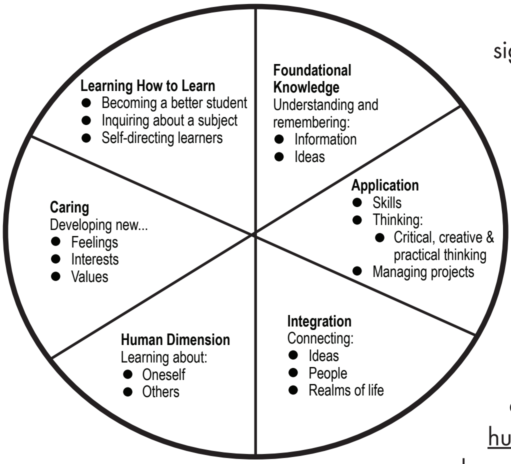

Figure 2. Fink's taxonomy of significant learning.

The ideas embedded within Fink's significant learning taxonomy are highly relevant to the P3 Award concepts and goals of sustainability. The P3 Award competition asks the learner to take foundational knowledge and critically apply it to creative problem-solving while integrating it with ideas of social systems, which can include economic markets. Engineers engaged in the P3 sustainable design of appropriate engineering systems also consider the human dimension and develop empathy because the conditions and definitions of success change with location and culture. This requires continuous learning as student teams and the engineering community adapt their designs to new information and generate new, innovative solutions for a particular situation.

Unlike the more traditional Bloom's taxonomy of cognitive development, Fink's taxonomy integrates cognitive, social, and affective development and is based on an idea of spiraled development. Fink views the developing learner in a continuous feedback mode spiraling through the categories shown in Figure 2. While Bloom's taxonomy has been expanded to create the social and affective domains of development, it considers development a linear process, necessarily progressing chronologically from the lowest levels of development to the higher ones. The benefit of using a taxonomy such as Fink's significant learning approach is that it more accurately reflects the interactive learning process, a goal that suits a program such as P3 particularly well. Therefore, Fink's taxonomy could guide the assessment of the educational success of the P3 program. For example, comparing a group of P3 participants against a control group that did not participate in P3 could highlight the experimental learning differences related to participating in the P3 Award and could determine whether the program contributed to significant learning (as defined by Fink) about sustainability. Additionally, questionnaires could be used to get a direct understanding of the value of the program to students and faculty.

Students learn about the design approaches of other P3 teams because participation in the Expo is an essential component of the P3 Award program. The P3 Award can also be used for program evaluation to assess students' significant learning. Participants in the P3 Award program could self-evaluate their designs after competing in the Expo to determine if there is value in learning about the approaches others have applied in designing for sustainability. The students could then be asked to reexamine their P3 submissions and offer suggestions for design improvement based on learning that occurred at the Expo. These evaluation tools could be used before and after participation in the P3 Award program to gauge the value of this program in terms of significant learning for engineers contributing to a sustainable future.

Demonstrating the value of these tools and learning hierarchies, such as Fink's taxonomy, could further substantiate these approaches for more broadly evaluating engineering education. By measuring these desirable characteristics in engineering students, there would be a clear message to the community about the value of including human, social, economic, environmental, and cultural factors into engineering design. In this way, high-caliber engineering design would become synonymous with sustainable engineering design.

References

ABET. (2007). *Criteria for accrediting engineering programs: Effective for evaluations during the 2007–2008 accreditation cycle.* Retrieved February 23, 2008, from http://www.abet.org/Linked%20Documents-UPDATE/Criteria%20and%20PP/E001%2007-08%20EAC%20Criteria%2011-15-06.pdf

American Society of Civil Engineers. (2004). *The role of the civil engineer in sustainable development*. ASCE Policy Statement 418. Retrieved on February 23, 2008, from http://www.asce.org/pressroom/news/policy_details.cfm?hdlid=60

Bruntland, G. (Ed.). (1987). *Our common future: The world commission on environment and development.* Oxford: Oxford University Press.

Bureau of Labor Statistics, U.S. Department of Labor. (2007). *Occupational outlook handbook, 2008–09 ed., Engineers.* Retrieved February 23, 2008, from **http://www.bls.gov/oco/ocos027.htm**

Fink, L. D. (2003). *Creating significant learning experiences: An integrated approach to designing college courses*. New York: Jossey-Bass.

Huntzinger, D. H., Hutchins, M. J., Gierke, J. S., & Sutherland, J. W. (2007). Enabling sustainable thinking in undergraduate engineering education. *International Journal of Engineering Education, 23,* 218–230.

Lattuca, L. R., Terenzini, P. T., & Volkwein, F. J. (2006). *Engineering change: A study of the impact of EC2000.* Retrieved February 23, 2008, from http://www.abet.org/Linked%20Documents-UPDATE/White%20Papers/Engineering%20Change.pdf

Lucid Design Group. (2007). Retrieved February 23, 2008, from http://www.luciddesigngroup.com

Mihelcic, J. R., Crittendan, J. C., Small, M. J., Shonnard, D. R., Hokanson, D. R., Zhang, Q., et al. (2003). Sustainability science and engineering: The emergence of a new metadiscipline. *Environmental Science and Technology*, *37,* 5314–5324.

Rochester Institute of Technology. (2007). *EPA Solar Ovens*. Retrieved February 23, 2008, from http://www.rit.edu/~633www/EPAsolarovens/index.html

Role of the Civil Engineer in Sustainable Development, The. (2004). Retrieved July SITumbra. (2007). Retrieved February 23, 2008, from http://www.situmbra.com

University of California at Berkeley. (2007). *Big ideas @ Berkeley.* Retrieved February 23, 2008, from http://bigideas.berkeley.edu/node/50

U.S. Environmental Protection Agency. (2007). *P3: People, prosperity, and the planet student design competition for sustainability.* Retrieved February 23, 2008, from http://www.epa.gov/ncer/P3/expo

CHAPTER 5

ASSESSMENT OF COOPERATIVE EDUCATION EXPERIENCES AS PART OF AN OVERALL PROGRAM ASSESSMENT PLAN

Jacqueline R. Mozrall, Ph.D.
N. Richard Reeve, Ph.D.
Emanuel Contomanolis, Ph.D.
Rochester Institute of Technology

Introduction

The benefits of cooperative education as a component of engineering education has been well-documented in the literature with respect to increased earnings and academic success (Blair et. al., 2004; Gardner, et al.,1992). However, less attention has been given to the benefits associated with assessing and evaluating cooperative education experiences in a structured way, in conjunction with other measures, in order to create positive change in the overall educational experience (Parsons et al., 2005).

In 1999, the Kate Gleason College of Engineering (KGCOE) at Rochester Institute of Technology (RIT), in response to the newly adopted ABET-EAC (EC 2000) criteria that was first released in 1997, (ABET, 2007), formed a college-level committee to coordinate the development of assessment techniques that could be utilized within program-level assessment plans. One of the significant outcomes of this committee was the development of both student and employer appraisal forms that are used to assess cooperative education experiences. These tools have evolved dramatically over time into their current form, which includes on-line administration and the development of a rich database of information that can be utilized to assess, from the student and employer perspective, the academic preparation of the student and the quality of the cooperative education experience. A discussion of the design and development of the assessment tools, the assessment and evaluation of the data, and examples of the resulting changes are presented here.

Background

RIT is a private, coeducational university in upstate New York. RIT, founded in 1829, is internationally recognized as a leader in career-oriented education, with over 15,000 students enrolled in eight colleges. RIT was one of the first universities to begin cooperative education in 1912. Students from all eight colleges participate in this highly respected co-op program which is recognized as one of the oldest and largest in the world.

The Kate Gleason College enrolls over 2,000 undergraduate students and 400 graduate students. KGCOE offers undergraduate and graduate degree programs in Computer Engineering, Electrical Engineering, Industrial and Systems Engineering, Mechanical Engineering, and Microelectronic Engineering, in addition to specialized graduate degree programs in Microsystems Engineering, Applied Statistics, Product Development and Manufacturing Leadership. All undergraduate degree programs are five year programs wherein students are required to complete at least one year of cooperative education experience. Students typically fulfill this requirement through the completion of a series of work blocks, ranging in length from three to six months. These work periods occur during the final three years of the program and are alternated with similar periods of full-time academic study.

In order for students to receive credit for their co-op experiences, students must evaluate their work assignments, and their performance must be evaluated by their sponsoring employers. Notification, via email, is sent to both students and employers several weeks before the completion of the work term with instructions for completing the on-line documents. The completed evaluations are submitted to the university where they are reviewed by an academic advisor and discussed with each student. Because of this organizational structure, wherein co-op is a graduation requirement and completed student and employer evaluations are required for students to receive co-op credit, the response rate for both students and employers is always 100%.

RIT has a fully developed infrastructure in place to support cooperative education with coordination between the colleges and the Office of Cooperative Education and Careers Services (OCECS). This office includes an Associate Vice President and Director, an Associate Director, coordinators dedicated to support all the academic programs, and a full support staff to assist students and alumni with their career goals and building and maintaining relationships with employers.

The assessment committee that was formed at the college level in 1999 was comprised of one representative from each academic department, the Associate Dean of KGCOE, and the Associate Director of OCECS, to facilitate the development of college-level assessment techniques that could be used as part of individual program assessment plans. Given that cooperative education is a common requirement of all undergraduate engineering programs at RIT, the committee focused on developing a generic assessment technique that could be adopted for use by all undergraduate engineering programs.

Assessment and Evaluation of Cooperative Education Experiences

Design, Development, and Evolution of Cooperative Education Assessment Tools

Before the initial development efforts in 1999, the college had been utilizing an open-ended form that students used to evaluate co-op experiences (V1.0). The form had three open-ended questions related to job description and major achievements, academic preparation, and personal and professional expectations. The employer form asked employers to rate students on a five-point scale in categories such as quality and quantity of work, ability to learn, initiative, reliability, judgment, attitude, and interpersonal and communication skills, as well as open-ended questions regarding strengths, recommended areas of improvement and additional comments.

In an attempt to use cooperative education experiences to assist in assessing the achievement of required ABET outcomes, both the student and employer forms were redesigned. The forms were kept closely aligned with the required ABET outcomes, so that all programs could utilize these assessment tools. Each program could then map their program-specific outcomes to the required ABET outcomes in order to ensure and demonstrate that all outcomes were being addressed. These paper-based forms (V2.0) still retained the open-ended questions that were present in earlier versions of the co-op report forms.

In 2002-03, in conjunction with OCECS, the committee significantly revised the student and employer co-op appraisal forms and produced on-line versions of each form. OCECS enlisted the services of software engineering senior design teams to produce on-line versions of both the student and employer appraisal forms (V3.0). The student form went through significant revisions and iterations, and the final form now provides an opportunity for students to assess their level of opportunity as well as their academic preparation on a five-point scale on various abilities (e.g., applying knowledge of math, science, and engineering). Assessing both the level of opportunity and academic preparation allows the independent measurement of the extent to which the co-op experience allowed students to accomplish an outcome and if they had the opportunity, then how well prepared they were to accomplish the outcome. As described in more detail in the Assessment and Evaluation section, this type of assessment also allows for the identification of gaps between opportunity and preparation.

In addition, students are asked to determine if their knowledge and skills in each category were enhanced by this co-op experience. Additional open-ended questions allowed students to describe the various contemporary issues that they were exposed to during the experience. In an attempt to strengthen our employer relationships and obtain meaningful feedback, there was a section developed to allow students to provide information on their co-op employer regarding the communication, feedback, support and mentoring that was provided by employers. This information provided helpful insights concerning student perceptions of the supervision and mentoring offered by the employer as well as in advising and counseling both students and employers about important elements of the work assignment and work setting.

It is very instructive, for example, to be able to refer to actual student experiences and insights when advising students about potential work assignments and employers. Also, employers routinely wish to know what students think about their work settings with an eye toward improving the quality of assignments and attracting the best students. Within individual student confidentiality requirements, we can present sufficient information and insight to help employers make needed and desirable improvements.

The employer report remains aligned with the required ABET outcomes as well as provides open-ended questions regarding strengths, recommended areas of improvement and additional comments. The employer data is considered to be an important direct measure, since this information is being provided by each student's immediate supervisor. Completed student and employer forms, obtained from the on-line system, are shown in Figures 1 and 2.

Please assess the extent to which your work experience provided you with the opportunity to accomplish the items below. Also, please assess the extent to which your academic program prepared you for the items below, and check the box if you feel your abilities were enhanced as a result of this co-op:

	5=Extensive, 3=Moderate, 1=Minimal **Level of Opportunity**	5=Excellent, 3=Average, 1=Poor **Academic Preparation**	**Enhanced by Co-op?**
Apply your knowledge of mathematics, science, and engineering. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	☐
Design and conduct experiments as well as analyze and interpret data. Comment: I analyzed and interpreted data but did not conduct any experiments on my how.	5 4 3 2 1 None	5 4 3 2 1 None	☐
Design a system, component, or process to meet desired needs. Comment: Developed a system of creating, documenting, reviewing, and updating preventative maintenance procedures for a functional area.	5 4 3 2 1 None	5 4 3 2 1 None	☐
Work in a multi-disciplinary team environment. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	✓
Work effectively with diverse individuals and team members. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	✓
Exhibit professional and ethical responsibility. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	✓
Communicate effectively through writing. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	✓
Communicate effectively through oral presentations. Comment: Presented my work to my supervisor on a weekly basis. Gave presentations to the department overviewing the process I developed for preventative maintenance procedures.	5 4 3 2 1 None	5 4 3 2 1 None	☐
Learn new information and skills and develop new abilities. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	✓
Use modern techniques and engineering tools. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	✓
Integrate the knowledge gained from two or more courses. Comment: Computer class on Excel, and statistics class.	5 4 3 2 1 None	5 4 3 2 1 None	☐

Gain awareness of the global and societal impact of your work. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	☐
Understand the need for continuous learning of engineering, scientific, mathematical, technical, and managerial concepts and solutions throughout your career. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	✓
Become more aware of contemporary issues pertaining to engineering, science, mathematics, and management. Comment:	5 4 3 2 1 None	5 4 3 2 1 None	☐

My employer/supervisor:	5=Strongly Agree, 3=Neutral, 1=Strongly Disagree
Provided and explained the priorities and objectives for my work term assignment. Comment:	5 4 3 2 1 N/A
Provided a work experience that was educationally meaningful and challenged my abilities. Comment:	5 4 3 2 1 N/A
Assigned additional responsibilities consistent with my growth on the job. Comment: After the first half of the co-op, I asked for some new projects. I performed time studies, and updated speed models for several tools in the Thin Films area, coordinated several tactical projects, and continued the preventative maintenance procedures project.	5 4 3 2 1 N/A
Assisted me in developing effective relationships with co-workers. Comment:	5 4 3 2 1 N/A
Provided me with appropriate feedback relative to my performance and professional development. Comment:	5 4 3 2 1 N/A
Was reasonable, sincere, and fair. Comment:	5 4 3 2 1 N/A
Was approachable and willing to provide me with necessary support and guidance. Comment:	5 4 3 2 1 N/A
Appeared committed to the RIT co-op program. Comment:	5 4 3 2 1 N/A
Provided me with a satisfactory work experience. Comment:	5 4 3 2 1 N/A
Did your employer review RIT's Employer Evaluation form with you?	Yes
My overall evaluation of the co-op experience:	5=Strongly Agree, 3=Neutral, 1=Strongly Disagree
The work assignment met my expectations. Comment: My projects progressed in the second half of the co-op. I never got bored with what I was doing.	5 4 3 2 1 N/A
The work assignment enhanced my overall education at RIT. Comment:	5 4 3 2 1 N/A
The work assignment helped me further develop my leadership skills. Comment:	5 4 3 2 1 N/A

I would recommend this employer to other co-op students. Comment:	5 4 3 2 1 N/A
Overall, I'm very satisfied with this work experience. Comment: I do feel like this is a job I would consider after graduation.	5 4 3 2 1 N/A

Additional Info

Describe the major achievements of your work assignment. Developing a system of Job Aids (preventative maintenance procedures) for the functional area. Coordinating and executing tactical projects for the area. Performing time studies and updating speed models for several area machines. Completing small assignments as others in the department requested them.
How well did your academic preparation assist you in this assignment? Please note specifically any courses and other RIT experiences that were particularly helpful. Comment on subject areas that should be added or strengthened in the curriculum. My first co-op experience taught me how to approach machine technichians, and form relationships with them to make the job easier and more effective. Microsoft Excel is the program most often used by the IE's. I would have liked to have further education into the abilities of this program. Also, "macros" are something I have been introduced to on this job.
What contemporary issues were you exposed to in this work assignment? Comment on technical, organizational, and societal issues. Experienced a round of layoffs that affected the employees emotionally, and conflicted with their work progress. The department went through a few reorganizational processes due to people taking leave, being layed off, or moving to a different department.
Describe the positive experiences - both personal and professional - related to this work assignment. There are a lot of recent college grads, and people under the age of 30 working on this site. I was able to relate to them, and feel like I was part of the company instead of just the college intern kid. Also, I have the opportunity to go back and work there in the future.
Describe the negative experiences - both personal and professional - related to this work assignment. Far from home and family (this could be a positive too!).
How has this work experience influenced your thinking and plans concerning your RIT coursework and/or career plans? While most enjoyed working at [Company], a few recent college grads expressed that they wanted to "move on" from [Company]. I am not sure if they felt unchallenged, or were just ready for a new job experience after spending a few years with [Company]. This made me consider what an MBA might do for me by increasing the types of jobs a could fill, or increasing the different styles of projects I could take on at one company to prevent a long term job with one company from getting repetitive, or uninteresting after several years.

Figure 1.
Student Co-op Report (V3.0.). Student, employer and job information fields not shown.

Rate the student on his/her performance in the following areas. 5 = Excellent, 1 = Poor.

Quality of Work: Accuracy, thoroughness, timeliness Comment: Completes all work on time with attention to detail and high degree of quality	5 4 3 2 1 N/A
Quantity of work: Volume, pace, and effort Comment: completes all tasks on schedule	5 4 3 2 1 N/A
Ability to learn: Grasps and retains new skills and concepts Comment: Fast learner who is willing to take on responsibility.	5 4 3 2 1 N/A
Initiative: Originates ideas and seeks new responsibilities, proactively seeks assistance Comment: allways ready to take on new projects with great enthusiasm	5 4 3 2 1 N/A
Ability to apply knowledge of mathematics, science, and engineering Comment: as necessary	5 4 3 2 1 N/A
Ability to design and conduct experiments, as well as to analyze and interpret data Comment: works well with engineering and vendors responsible for new design.	5 4 3 2 1 N/A
Ability to design a system, component, or process to meet desired needs Comment: Completed and provided input for continuous improvement projects	5 4 3 2 1 N/A
Ability to function on multi-disciplinary teams Comment: accomplishes all tasks on time as required.	5 4 3 2 1 N/A
Ability to identify, formulate, and solve engineering problems Comment: Great self starter who needs very little guidance	5 4 3 2 1 N/A
Understanding of professional and ethical responsibility Comment: Fully competent	5 4 3 2 1 N/A
Ability to communicate effectively, written and oral Comment: communicates well with all	5 4 3 2 1 N/A
Ability to use the techniques, skills, and modern engineering tools necessary for engineering practice Comment: Applies skills as required with projects that are assigned.	5 4 3 2 1 N/A
Quality of technical preparation Comment: performs quality work	5 4 3 2 1 N/A
Respect for diversity and a knowledge of contemporary professional, societal, and global issues Comment: Fully competent	5 4 3 2 1 N/A
The ability to understand own strengths and weaknesses and receive feedback Comment: always willing to receive feedback	5 4 3 2 1 N/A
Ability or potential to lead others and/or projects, set and achieve goals, create change and inspire confidence Comment: Has continuously demonstrated leadership skills on projects that have been assigned to her.	5 4 3 2 1 N/A
Overall Performance Comment: Outstanding	5 4 3 2 1 N/A
If available and appropriate, would you offer a regular employment position to this student upon graduation?	5 4 3 2 1 N/A

Additional Comments

Strengths Has a great attitude, knowledge and willingness to learn, she will be a great asset to any organization.
Recommended areas for improvement
Feel free to enter any additional comments you would like to add about this student's performance in the box below.

Figure 2.
Employer Co-op Report (V3.0). Student, employer and job information fields not shown.

The on-line system also provides a report function that allows a search and filter of co-op data by: student name, degree program, quarter co-op was completed, student year level, co-op number (e.g., 1st, 2nd, 3rd co-op experience), gender, and company name. Individual student and employer reports can be viewed as well as summary statistics for each question (i.e., mean, median, score ranges) that are produced for given student groups for a selected time period (Figures 3 and 4). All of the data can also be saved to an Excel file for additional customized analysis. This feature allows each program to develop additional unique evaluation metrics.

The on-line co-op assessment tools and data collection capabilities have created a rich database of information and an enhanced ability to assess and evaluate student performance, based on inputs from both students and employers. In addition, it allows for the development of student feedback profiles and employer co-op profiles. For example, the student feedback profiles are provided to companies that are interested in knowing how their co-op program is perceived by RIT students. These student profiles allow companies to benchmark and improve their competitiveness.

The employer co-op profiles, which indicate the companies that participate in the co-op program, can illustrate company distributions by industry sector, geographic location, size as well as other relevant variables of interest. This allows us to determine how our students are meeting the needs of various market segments and market our program to prospective students and employers. This employer profile has also been used to guide event planning and to support student and alumni outreach. Our executive and honors engineering programs at RIT include domestic and international trips to visit engineering companies within various regions of the world (e.g., Seattle, Silicon Valley, Germany). This employer profile has allowed us to identify potential locations to visit as well as student and company contact information.

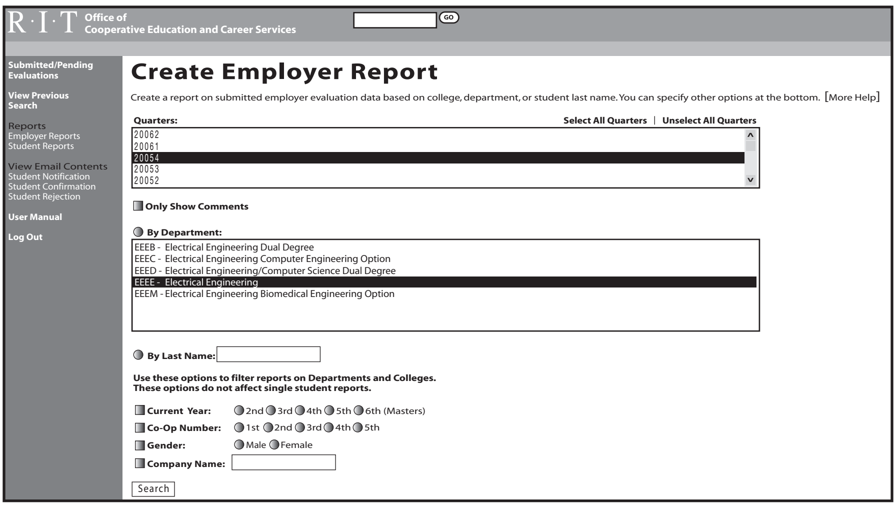

Figure 3.
Report function that provides search and filter capabilities by various factors.

Student Reports

View Email Contents
Student Notification
Student Confirmation
Student Rejection

User Manual

Log Out

Report on departments: EEEE Quarters: 20054

Question	Mean	Median	Lowest	Highest	Count
Quality	4.27	4.0	3.0	5.0	73
Quantity	4.29	4.0	2.0	5.0	73
Learn	4.42	5.0	3.0	5.0	73
Initiative	4.08	4.0	2.0	5.0	73
Knowledge	4.2	4.0	2.0	5.0	73
Experiment	4.14	4.0	2.0	5.0	73
SystemDesign	4.08	4.0	3.0	5.0	73
TeamWork	4.38	5.0	3.0	5.0	73
ProblemSolving	4.09	4.0	3.0	5.0	73
Ethic	4.23	4.0	2.0	5.0	73
Communication	4.14	4.0	2.0	5.0	73
Applying Skills	4.25	4.0	3.0	5.0	73
Preparation	4.12	4.0	2.0	5.0	73
Respect	4.31	4.0	3.0	5.0	73
Self-Awareness	4.01	4.0	2.0	5.0	73
Leadership	3.94	4.0	2.0	5.0	73
Overall	4.27	4.0	3.0	5.0	73

Question	Mean	Median	Lowest	Highest	Count
Hours	40.59	40.0	35.0	70.0	73

Hours - Hours per week:
Quality - Quality of Work: Accuracy, thoroughness, timeliness
Quantity - Quantity of work: Volume, pace, and effort
Learn - Ability to learn: Grasps and retains new skills and concepts
Initiative - Initiative: Originates ideas and seeks new responsibilities, proactively seeks assistance
Knowledge - Ability to apply knowledge of mathematics, science, and engineering
Experiment - Ability to design and conduct experiments, as well as to analyze and interpret data
SystemDesign - Ability to design a system, component or process to meet desired needs
TeamWork - Ability to function on multi-disciplinary teams
ProblemSolving - Ability to identify, formulate, and solve engineering problems
Ethic - Understanding of professional and ethical responsibility
Communication - Ability to communicate effectively, written and oral
ApplyingSkills - Ability to use the techniques, skills, and modern engineering tools necessary for engineering practice
Preparation - Quality of technical preparation
Respect - Respect for diversity and a knowledge of contemporary professional, societal, and global issues
Self-Awareness - The ability to understand own strengths and weaknesses and receive feedback
Leadership - Ability or potential to lead others and/or projects, set and achieve goals, create change and inspire confidence
Overall - Overall Performance

Figure 4.
Employer report filtered by program (EEEE = Electrical Engineering) and time period (20054=Summer, 2006).

Assessment and Evaluation of Data

Employer Data

Once all of the data has been saved in an Excel file, metrics can be developed, analyzed, and evaluated based upon program specific outcomes. For example, a defined metric goal that is used to evaluate student performance in the Industrial Engineering program is: At least 90% of students will be rated three (3) or greater by their employer for each competency identified in the assessment instrument. This type of evaluation provides affirmation, in this case by the employer, that students possess the knowledge and skills required by the ABET-EAC criteria (Figure 5). Based on these data, and our defined performance metric, students are demonstrating all the knowledge and skills required by the ABET criteria. Above and beyond the ABET requirements, however, we are interested in understanding, among other things, if students are receptive to employer feedback. According to our assessment of the evaluations, only 80% of the students are adequately receptive to feedback (Figure 5 – Self-Awareness/Feedback). Action could consequently be focused on counseling students in better understanding their own strengths and weaknesses and becoming more receptive to constructive feedback.

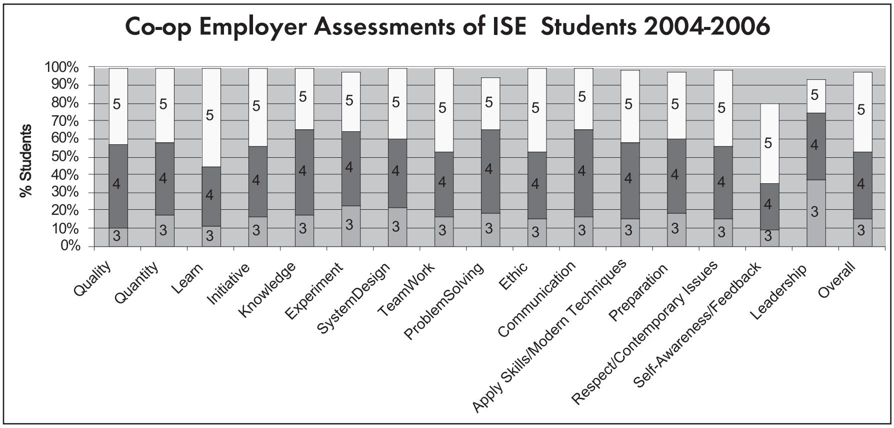

Figure 5.
Co-op employer ratings of Industrial Engineering students (graduation year 2004-2006) at RIT. Percentage of students receiving a rating of 3, 4, or 5.

Through OCECS, students are required to attend orientation sessions before they can co-op. These sessions provide training in many topics, including: cover letter and resume writing, interviewing, job search skills, business etiquette, and defining career goals. In coordination with the OCECS, we could ensure that there is a topic related to the importance of, and receptiveness to, feedback. More work can also be done with individual students who do not receive adequate ratings (i.e., ratings of 1 or 2), to better understand if there are individual issues that need to be addressed.

In Mechanical Engineering, it was observed that the "Leadership" question response set on the employer reports had a higher incidence of "NA" responses (50%) than any other category. Trailing close behind in "NA" responses was the "Self-Awareness/Feedback" category (47%). All other categories had a "NA" response of less than 20%. This implies that a significant number of students are not getting an opportunity to learn and demonstrate leadership skills on their co-op. Further, when reading the comments under employer assessment of areas that students should improve, a very common topic was in the area of planning, time management, task tracking, and project management. Mechanical Engineering has recently implemented the Design Project Management class, and also has integrated a required project management experience into the required capstone program for dual degree students (e.g., BS/Master of Engineering). However, it is still too early to determine the impact of these changes.

The data related to the "Self-Awareness/Feedback" category is consistent with the pattern demonstrated by Industrial Engineering students. The OCECS co-op orientation sessions described above are conducted at the college level across all departments, so the proposed changes related to counseling students in better understanding their own strengths and weaknesses and becoming more receptive to feedback, are also supported by this data set from Mechanical Engineering students.

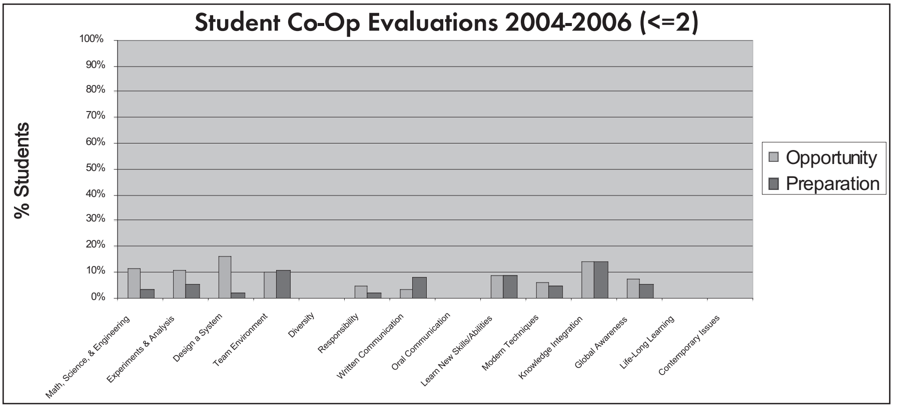

Figure 6.
Student co-op evaluations of Industrial Engineering students (graduation year 2004-2006) at RIT. Percentage of students rating their opportunity and preparation across various categories at 1 or 2.

Student Data

Similar analyses can also be done with the student data, to ensure that students feel they are prepared with respect to required knowledge and skills they must apply during a co-op experience. The student work report allows examination of the student's perception of the opportunity to perform as well as their preparation across various competencies (Figure 6). In comparing preparation and opportunity, it is important to look for gaps. Ideally, opportunity should meet or exceed preparation. A defined metric goal that could be used to evaluate student performance based on employer co-op reports could be: At least 90% of students rate both their preparation level and opportunity to perform at three (3) or greater; or inversely, no more than 10% of the students rate preparation and opportunity less than three (i.e., 1 or 2).

If we consistently find that the opportunity is low in various categories, we can work in conjunction with OCECS to develop stronger opportunities in these areas. For example, in Figure 6, over 15% of the students indicated the opportunity for systems design to be low. While this is somewhat expected in initial work experiences, we could work to find additional co-op experiences that allow more systems design opportunities, particularly for upper-level students during their final co-op experiences. These more significant experiences would most likely also require students to integrate knowledge from several different subject areas, and more than 10% of Industrial Engineering students indicate the opportunity to integrate knowledge to be low (Figure 6).

Individual student evaluations of both employers and work assignments provide extremely useful data in managing and enhancing employer co-op partnerships. Of particular interest is the student's overall assessment of the quality of the co-op work assignment. This perspective in not only helpful in assessing the experience for individual students but also in identifying potential trends and patterns among employers that merit intervention by appropriate university officials. A pattern of low quality ratings by students for a certain employer, for example, will automatically prompt OCECS staff outreach to that employer. The goal will be to share and discuss, in general and confidential terms, opportunities to potentially revise work assignments and/or change work settings to enhance the experience for both the participating student and employer. Employers have consistently demonstrated an openness to this type of dialog and a willingness to consider and act upon jointly determined recommendations.

Actions Based on Assessment and Evaluation

Over the last 6 years, the assessment and evaluation of the co-op data, in conjunction with other data (e.g., course materials evaluations, student exit surveys), has allowed for positive curricular change. For example, Industrial Engineering students had been indicating in their student co-op reports that skills in AutoCad and Access were important to obtain co-op assignments and perform successfully. In 2001, AutoCad and Access skill development was integrated into the first year curriculum. Following this curricular change, it can be seen that student ratings of their co-op experiences in their ability to use modern engineering tools, indicate that third year students felt better prepared than fifth year students (Figures 7 and 8 – Use modern tools). In addition, the opportunity increased, which is most likely due to their increased skill set in this area. This trend may be attributed to these curricular changes that occurred before third year students began their cooperative education. The positive effect of this change is also confirmed by student focus group data and co-op data that have been collected since that time (Figures 5 and 6 – Apply Skills/Modern Techniques).

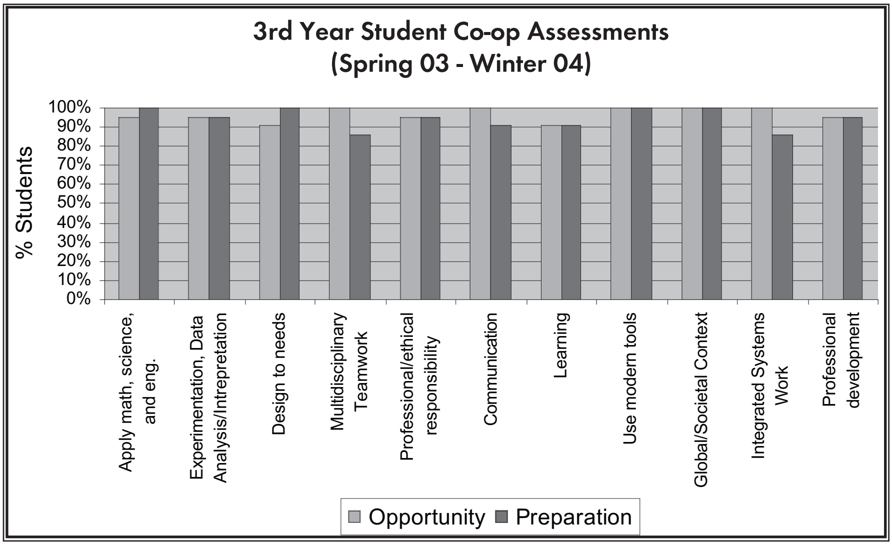

Figure 7.
Percentage of 3rd year Industrial Engineering students rating their opportunity and preparation on co-op a three or greater from Spring 2003 through Winter 2004.

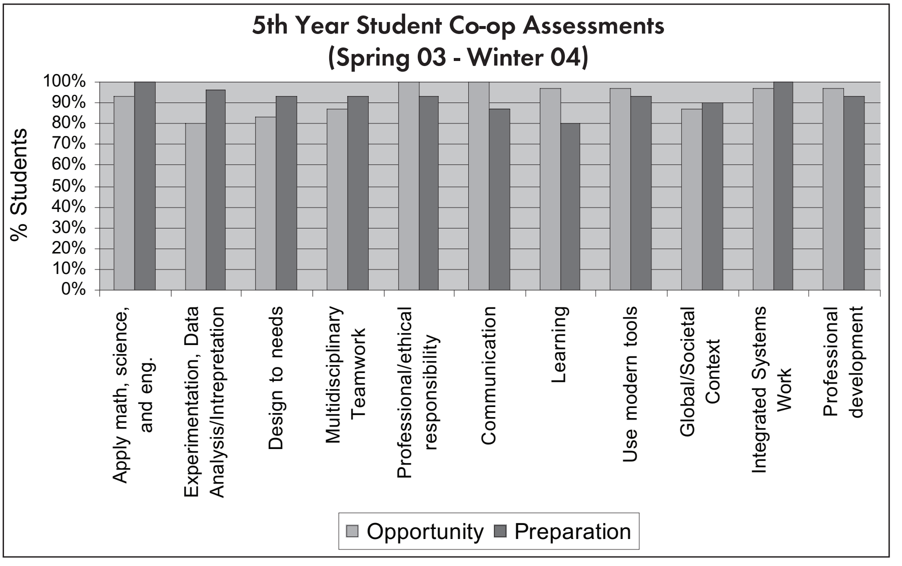

Figure 8.
Percentage of 5th year Industrial Engineering students rating their opportunity and preparation on co-op a three or greater levels from Spring 2003 through Winter 2004.

There has also been a concerted effort to incorporate the handling of contemporary issues into more coursework. Several required Industrial Engineering courses now provide students with some knowledge of "contemporary issues". There has also been a significant increase in the number of professional electives that are now offered. Several of the new offerings are related to contemporary topics currently facing engineers (e.g., lean manufacturing, manufacturing systems, project management, sustainable design). This effort is confirmed by student co-op data that indicates that 3rd year students felt better prepared in their understanding of solutions in a global and societal context than 5th year students (Figures 7 and 8 – Global/Societal Context). This effort has also been confirmed by recent student co-op data (Figures 5 and 6 –Respect/Contemporary Issues and Global Awareness). Student exit survey and student focus group data also confirms that students are supportive of new required courses and elective course offerings.

Conclusions

Cooperative education is considered to be a key component of the educational experience of engineering students in the Kate Gleason College at RIT. As such, in order to capture the entire educational experience of students, it is imperative that we evaluate the cooperative education experience. The structured data collection, assessment and evaluation of these experiences can also provide many benefits including evidence that students are meeting program outcomes; assistance in identifying individual students that may lack the appropriate skills; and feedback on the work experience to the employing organization in order to ensure quality work assignments.

The assessment and evaluation of cooperative education experiences is necessary but not sufficient in the development of a comprehensive assessment plan. In the Kate Gleason College at RIT, employer and student co-op data, in addition to course material evaluations, student exit surveys, alumni surveys, and employer and student focus groups are used to assess student learning. The employer and student co-op data provide critical evidence of student performance, particularly with respect to broader objectives and outcomes such as ethics, life-long learning, communication and contemporary issues (Bartkus, 2001; Parsons, et. al., 2005).

This information can also be valuable in corroborating other data, and in better identifying and understanding emerging trends. Data triangulation, in which multiple methods are used for assessment, will increase the likelihood that the assessment is valid and reliable (Felder and Brent, 2003). The development of a structured evaluation system allows for powerful data mining capabilities and enhanced support of various programs that rely on student and alumni outreach. These data, as part of an overall continuous improvement process, enable each engineering program to assess and evaluate strengths and weaknesses in order to drive change and enhance quality.

Acknowledgments

The authors wish to acknowledge the faculty and staff who have helped lead and develop the KGCOE co-op assessment process since its inception. This group includes Vincent Amuso, Louise Carrese, Roy Czernikowski, Edward Hensel, Mike Jackson, Steve LaLonde, Harvey Palmer, Fernando Naveda, and Brian Thorn. In addition, we would like to thank the software engineering students who have developed the implementation software: Shaun Collins, Patrick Flanagan, Neil Girard, Jennifer Kotski, Philip Makara, Tom Small, Michael Sorenson, Peter Vertenten, Dan Volpe, Chris Woodbury. Finally, we would also like to thank the corporations which have provided valuable feedback for the growth and improvement of the assessment process at RIT.

References

ABET (2007). History - Outcomes assessment and continuous improvement, Retrieved August 8, 2007, from http://www.abet.org/history.shtml .

Bartkus, K.R. (2001). "Skills and cooperative education: A conceptual framework," *Journal of Cooperative Education*, 36, 17-24.

Blair, B.F., Millea, M., and Hammer, J. (2004). "The impact of cooperative education on academic performance and compensation of engineering majors," *Journal of Engineering Education*, 93, 333-338.

Felder, R.M. and Brent, R. (2003). "Designing and teaching courses to satisfy the ABET criteria," *Journal of Engineering Education*, 92, 7-18.

Gardner, P.D., Nixon, D., and Motschenbacker, G. (1992). "Starting salary outcomes of cooperative education graduates," *Journal of Cooperative Education*, 27, 16-26.

Parsons, C.K., Caylor, E. and Simmons, H. (2005). Cooperative education work assignments: The role of organizational and individual factors in enhancing ABET competencies and co-op workplace well-being, *Journal of Engineering Education*, 94, 309-318.

CHAPTER 6

FACILITATING LIVE ASSESSMENT OF OUTCOMES IN ENGINEERING EDUCATION

Charles J. Malmborg
Rensselaer Polytechnic Institute

Introduction

This chapter describes how the effectiveness of an educational assessment system is enhanced through the creation of interactive, computer based assessment tools when they are introduced directly into the faculty committee process. As described in this chapter, such tools can facilitate calibration of target performance levels, assessment of individual outcomes, cross cutting assessment, and they can be applied to drill down in the assessment database to address specific problem areas.

The opportunity for the development of these tools arose from the development of educational objectives and outcomes differentiating the unique strengths and characteristics of an undergraduate engineering program at Rensselaer. As programs across the country have climbed the assessment learning curve, they have addressed a number of new challenges. Among these challenges is overcoming the extended time frame associated with the academic year. That is, many useful assessment measures associated with student cohorts are observed on an annual basis. The relatively long lag between the implementation of these measures and the collection of a critical mass of data creates difficulties in calibration of assessment processes and target performance levels, slows identification of trends, and generally slows impact analyses of curricular and pedagogical initiatives. Depending on how assessment results are introduced to the committee process, the corresponding delays in feedback loops can make it difficult to use assessment results to support meaningful discussion and group decision making.

This chapter describes how these problems have been addressed by the faculty committee responsible for the undergraduate program in Industrial and Management Engineering (IME) at Rensselaer Polytechnic Institute. An interactive tool has been developed to eliminate feedback delays, facilitate the committee process of group decision making, and support in-depth analysis of assessment results. This software tool is designed specifically for the assessment system associated with the IME program. Although this system reflects the IME program's unique history, character and strengths, the design principles underlying the assessment software tool are readily adaptable to other engineering programs. The "live" assessment process created through its application has energized and accelerated the assessment process by enabling identification of the easiest opportunities for curricular reform and effective analysis of the more difficult ones.

The next section provides background information of the IME program and the development of its assessment system in response to ABET Engineering Criteria 2000. The third section describes the decision support tool developed specifically to implement this assessment system and facilitate the group process. The fourth section describes the impact of the tool on curricular development for the IME program. The final section offers a summary and conclusions.

Background Information

Rensselaer Polytechnic Institute has offered an undergraduate program in Industrial Engineering (IE) since 1933. Initially, Rensselaer's IE program was created by replacing accounting and social sciences courses with courses in mechanical and electrical engineering within a Business Administration program. In 1938, the first two years of the IE program were replaced by common core requirements for all engineering disciplines making the program a true engineering program as opposed to a modified Management program. Following the trend toward greater specialization seen in most engineering disciplines during the late 1930's, fundamentals of production systems, organizational administration, and plant layout courses were introduced in the IE program by 1940 to replace courses in history and economics. In 1946, the program was renamed to Management Engineering and changed once again with the addition of courses in manufacturing processes including metal casting, forming and welding. By 1960, the IE program at Rensselaer had returned to the School of Management. Subsequently, courses in personnel management, industrial relations, and work methods were added to the curriculum to replace mechanical engineering courses, and an accounting course was re-introduced to the curriculum. The unique character of Rensselaer's IE program was further differentiated in the late 1950's when it became among the first of its type to require a computer programming course. More management oriented technical electives in data analysis and data processing were added to the curriculum during the sixties along with the courses in operations research and statistics that were being introduced in most other IE programs in the United States. During the 1970's, the program further increased its emphasis on data analysis applied to management decision making and, by 1982, the program was renamed as Industrial and Management Engineering, and formally relocated back to the School of Engineering where it had resided from the late 1930's through the late 1950's.

In 1987, the administration at Rensselaer identified the IME program as the focal point for the creation of a totally new inter-school department focused on decision sciences. Faculty members from the School of Management with interests in information systems and statistics were brought together with faculty members from the Schools of Engineering and Science with interests in systems engineering and optimization to form the new department of Decision Sciences and Engineering Systems (DSES). The collective interest areas of the multi-disciplinary faculty of the new DSES department spanned computation in the analysis and synthesis of data to generate information for decision-making models and decision model building from a total integrated systems perspective. The new department sought to assert these themes in both its research thrusts and academic programs. It also sought to emphasize its relevance by focusing its methodological research and academic programs on solving problems in global services and global manufacturing. Both the curriculum and courses in the IME undergraduate program steadily evolved during this period to reflect faculty strengths in these methodological and application domain areas and the changing learning styles of IME students.

The development of DSES Master's curricula in manufacturing and services initially proceeded on parallel tracks. The department launched a successful Master's program in Manufacturing Systems Engineering in 1992 and a new concentration in Service Systems within the IME Master's program in 1994. The same differentiation was initially reflected in new research centers created by the DSES faculty. In 1990, the Center for Services Research and Education (CSRE) was created by faculty members in DSES to coordinate academic programs from several disciplines across campus and provide a focal point for multidisciplinary research focused specifically on service sector issues. In 1994, the Electronics Agile Manufacturing Research Institute (EAMRI) was created by DSES faculty members to study problems in globally distributed manufacturing supply chains. During the mid 1990's, evolving trends in DSES faculty research and course development emphasizing a balance between service and manufacturing applications resulted in the eventual convergence of these domain application interests. The department became a focal point for the emergence of a new multidiscipline designated as "service systems engineering" (ABET, 2005), postulating the parallels and contrasts in decision making systems for service and manufacturing systems. In 2004, DSES revised its Master's degree in Manufacturing Systems Engineering to become the first program to offer a Master's degree in Service and Manufacturing Systems Engineering. Just as IME undergraduate courses were being developed and revised to emphasize parallels between manufacturing and service systems, they were also being modified to emphasize the use of computation in data analysis and optimization to reflect the methodological roots of the DSES faculty. New computationally focused courses were added to the IME curriculum in such areas as information systems and discrete event simulation. New computational components were being added to courses in operations research, statistics, and quality control. Some of the unique applications in the use of computation in IME courses are described in Bullington et al, 2000; Heragu et al, 2001; Heragu et al, 2003.

When the ABET 2000 initiative was launched in the mid 1990's, ABET accredited engineering programs were charged with identifying their unique educational objectives and outcomes (ABET, 2005). Like most other engineering undergraduate programs, a version of ABET specified outcomes 3a-3k were adopted for the IME program. However, the consensus of the program constituents was that the educational objectives of the IME program must effectively facilitate learning and assessment (Deek et al, 1999; Gronlund, 1999) but also reflect the program's unique history and strengths in the systems approach in model building and problem solving, the application of computing in decision making, in-depth knowledge of manufacturing/service systems, and the ability to apply principles of management. Subsequently, the following program objectives were approved in 2006 after a long consensus-building process:

- *Objective 1* - Graduates of the IME program will have a solid foundation in all of the fundamental areas of industrial and management engineering emphasizing a total integrated systems perspective and reflecting the unique strengths of Rensselaer's program including in-depth knowledge of manufacturing and service systems, effectiveness in the management of people and systems, and the creative application of computing and other technologies.
- *Objective 2* - Graduates of the IME program will be creative and innovative designers of systems, processes, facilities, services, products and equipment with strong analytical skills and a sufficient practical understanding of real world problems to be skillful at identifying, modeling, analyzing and solving challenging problems.
- *Objective 3* - Graduates of the IME program will be effective oral communicators, good technical writers and have a solid foundation for using communication media of all types to facilitate their strengths as contributors and leaders of diverse teams.
- *Objective 4* - Graduates of the IME program will be broadly educated in the humanities, social sciences and engineering professionalism which will inform their socially responsible and ethical practice of industrial and management engineering. They will understand the importance of life long learning and be capable and motivated to pursue continued growth and learning throughout their careers.
- *Objective 5* - Graduates of the IME program will have a solid foundation in math and science which they can effectively apply in the practice of industrial and management engineering.

As described in Felder and Brent (2003), ABET defines program outcomes based on educational objectives to represent, "more specific statements of program graduates' knowledge, skills, and attitudes that serve as evidence of achievement of the program's educational objectives." A total of 12 program outcomes were identified by the DSES faculty for the IME program objectives including:

- *Outcome 1* - IME students will demonstrate the ability to apply a total integrated systems perspective to the practice of industrial and management engineering.
- *Outcome 2* - IME students will demonstrate the ability to apply knowledge of manufacturing and service systems to the practice of industrial and management engineering.
- *Outcome 3* – IME students will demonstrate the ability to apply in-depth knowledge of computing to the practice of industrial and management engineering.
- *Outcome 4* - IME students will demonstrate the ability to manage people and systems.
- *Outcome 5* - IME students will demonstrate the ability to design innovative products, services, facilities, equipment, processes and systems.
- *Outcome 6* - IME students will demonstrate the ability to identify, model, analyze, and solve challenging real-life problems.
- *Outcome 7* - IME students will demonstrate a solid foundation in math and science.
- *Outcome 8* - IME students will demonstrate strong communication skills with emphasis on technical writing and interpersonal communication.

- *Outcome 9* - IME students will demonstrate the ability perform effectively on diverse teams, both as leader and contributor.
- *Outcome 10* - IME students will be informed citizens broadly educated in the humanities and social sciences.
- *Outcome 11* - IME students will demonstrate that they are prepared to practice engineering in a socially responsible and ethical manner.
- *Outcome 12* - IME students will learn in a stimulating environment that prepares and motivates continued growth and learning.

To assess these twelve outcomes, the DSES faculty developed an array of measurement instruments. Based on the many advantages of portfolios for use in assessment(Christy and Lima, 1998; Panitz, 1996), a requirement for IME undergraduates to create a career portfolio was introduced as part of the capstone design requirements. This included a formal statement of individual career goals, a resumé and transcript along with a summary of unique accomplishments chronologically summarizing awards, participation in undergraduate research programs, special training, descriptions and documentation of extracurricular activities, co-op and internships. The portfolio requirement also includes a documented history of published articles and completed project assignments in course work, internships and co-op assignments documenting the student's skills and capabilities in engineering.

Additional measurement instruments developed by the DSES faculty included a survey of co-op employers of IME students, surveys of alumni and graduating seniors, an evaluation form for industry sponsors of capstone design team projects, a team evaluation form focused on specific skills for use by course instructors, and an internship/co-op assessment survey to be completed by IME students. The DSES faculty also utilized the extra question option on the Individual Development and Educational Assessment (IDEA) course evaluation form to add outcomes-based survey questions to routine instructor evaluations in IME courses. Individual student performance as measured by grades was also used in a supplementary assessment role.

Following sound principles of assessment (Aldridge and Benefield, 1998; Panitz 1996), the instruments supporting the assessment of the twelve outcomes were designed to achieve a balance of individual student and student group measures, (e.g., transcripts and portfolios versus team evaluation reports and IME industry sponsor surveys), direct and indirect measures, (e.g., portfolio evaluations and co-op employer evaluations versus senior and alumni surveys), and reporting from a diversity of program constituents including the faculty, students, employers and alumni. The assessment system was also designed to assure that each outcome was evaluated, at least in part, from specific curriculum requirements applicable to all IME students as well as making constructive use of elective requirements, e.g., student evaluations and grades from technical elective courses, etc.

A Decision Support System for Live Assessment

To exploit the array of assessment instruments developed for the IME undergraduate program, the undergraduate committee in DSES realized that an innovative, data-driven decision support system was needed to provide program-level assessment. This led to the development of a customized decision support tool in the form of an interactive computer program providing flexible access to a series of assessment databases created from the data obtained from the assessment instruments. This tool is described in this section. It is used directly during undergraduate committee meetings to report overall assessment results as well as to analyze results parsed by individual educational outcomes. As an interactive system, this tool facilitated the necessary what-if analyses for the initial calibration of target performance levels while allowing users to drill down into lower level assessment data for investigation of specific performance issues. For example, performance indicators proposed for Outcome 5 for the IME program include:

- 80% of graduating seniors achieve a 2.5 or higher GPA in related courses
- 90% of senior portfolios are rated as adequate or higher in design
- 80% of alumni respond with 3 or higher on the survey question on design skills
- 80% of seniors respond with 4 or higher on the senior survey design question
- 80% of industry sponsors rate design skills on capstone projects at 3 or higher

The decision support tool enables the undergraduate committee to validate these measures by changing the percentages and re-evaluating the corresponding results. In this way, it is possible to establish a meaningful baseline of performance for the program with respect to the assessment data available. It also enables the committee to identify trends in performance over time. It would be difficult to fully utilize the available assessment data without this interactive capability which facilitates "live" assessment during the committee discussion and decision making processes. Thus, the system supports the undergraduate committee in the continuous improvement of the IME curriculum based on appropriate synthesis and reporting of the data collected through surveys, industry sponsor and employer evaluations, transcripts and other assessment instruments comprising the underlying databases. In 2007, the system reported seven years of results from assessment data. This enabled the identification of trends in performance that were associated with various changes introduced to the IME in different years. The use of the decision support system has contributed to significant changes in the IME program over the past four years including a major re-structuring of the curriculum as described in the next section as well as numerous revisions in course content and pedagogy. The features of this system and an overview of its impact on the IME program are described in this section.

The entry screen of the DSES ABET educational assessment system is summarized in Figure 1.

The system enables the user to analyze assessment results by individual outcome or examine a series of over-arching measures in addition to reviewing program objectives and examine sample sizes associated with the assessment instruments used by the system. For example, the IME transcript analyzer provides direct (anonymous) access to the transcripts of program graduates by class year. This feature can be used to evaluate how IME graduates perform in any combination of program courses designated by the user. For example, Figure 2 illustrates how the user can assess the performance of IME undergraduates in selected math and science courses. Users can also apply the system to review how industry sponsors of capstone design projects rate the performance

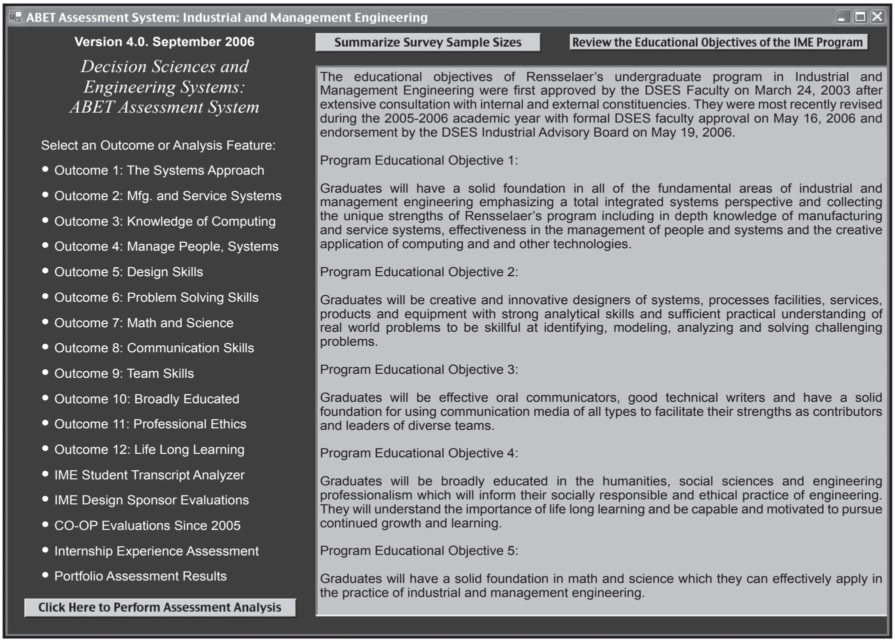

Figure 1.
DSES ABET Software System Entry Screen

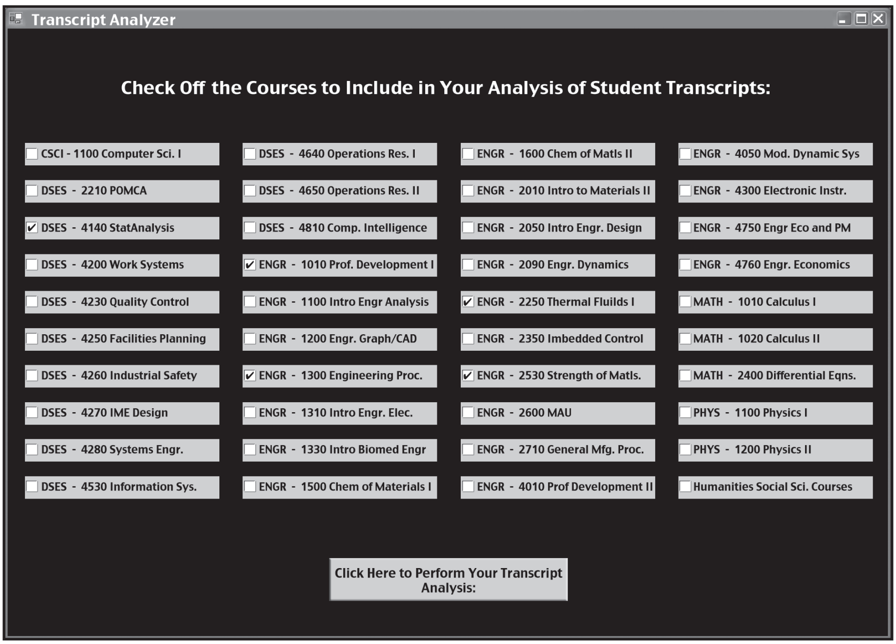

Figure 2.
Transcript Analyzer Feature

The Following Courses Were Selected for Inclusion in the Transcript Analysis:

DSES - 4140 Stat Analysis
ENGR - 1010 Prof. Development I
ENGR - 1300 Engineering Proc.
ENGR - 2250 Thermal Fluids I
ENGR - 2530 Strength of Matls.

The GPA Values for the Selected Combination of Courses are Summarized Below:

Percentages include grades of those class members actually taking the course(s). Therefore, they may not always sum to 100% especially with elective course(s).

Year	Avg. GPA	Median GPA	Below 2.0	2.0-2.49	2.5-2.99	3.0-3.49	Above 3.5
2001	3.47	3.48	0.00%	0.00%	13.51%	35.14%	51.35%
2002	3.51	3.60	0.00%	0.00%	14.63%	21.95%	63.41%
2003	3.45	3.36	0.00%	0.00%	10.20%	44.90%	44.90%
2004	3.49	3.45	0.00%	2.94%	8.82%	38.24%	50.00%
2005	3.51	3.61	0.00%	2.86%	8.57%	28.57%	60.00%
2006	3.43	3.49	0.00%	0.00%	25.00%	25.00%	50.00%

Figure 2.
Transcript Analyzer Feature (contd.).

Rensselaer's Co-Up student performance instrument was revised in 2005. The revised form provides student ratings on ten performance dimensions.

Average IME student ratings are summarized below for each performance dimension where the following four-point rating scale is used

	Excellent	Above Avg.	Average	Below Avg.
TEAMWORK: Works well with others, values diversity, strives toward team goals.				
2005	80%	2016	0%	0%
2006	29%	57%	14%	0%
ATTITUDE: Enthusiastic, positiveand displays interest in work.				
2005	60%	20%	20%	096
2006	43%	14%	43%	0%
INITIATIVE: Self-motivated, diligent, seeks additional work when necessary.				
2005	00%	20%	0%	0%
2006	14%	29%	43%	14%
DECISION-MAKING: Evaluates options, displays maturity, uses good judgement.				
2005	80%	20%	0%	096
2006	14%	43%	29%	14%
TECHNICAL SKILLS: Proficient, adept in field of assigned responsibilities.				
2005	80%	20%	0%	0%
2006	28%	28%	43%	0%
LEADERSHIP QUALITIES: Posesses potential to lead and direct others.				
2005	0%	80%	20%	0%
2006	14%	57%	14%	14%
COMMUNICATION: Can express self well to superiors and employees across all levels.				
2005	40%	60%	0%	0%
2006	14%	29%	29%	29%
ACCOMPLISHMENT OF OBJECTIVES: Task oriented, persistent, works to task completion.				
2005	60%	40%	0%	0%
2006	43%	29%	29%	006
QUALITY OF WORK: Takes pride in work, displays neatness, thoroughness and accuracy.				
2005	80%	20%	0%	0%
2006	43%	43%	14%	0%
OVERALL PERFORMANCE:				
2005	80%	20%	0%	0%
2006	14%	29%	57%	0%

Figure 3.
Co-op and Capstone Project Evaluation Screens.

of IME undergraduate teams or how employers of IME undergraduates on co-op assignments are evaluated by their employers (see Figure 3).

To illustrate how the system is used to assess individual outcomes, consider the assessment matrix for outcome 6 that is presented in Table 1.

Table 1
Assessment Matrix for Outcome Six – Design Skills. IME students will demonstrate the ability to identify, model, analyze, and solve challenging real-life problems.

Program Outcomes	Implementation Strategy	Performance Criteria	Assessment Methods	Timeline	Feedback
What are the desired program outcomes? What should students know and be able to do?	What program activities (curricular and co-curricular) help produce the desired outcome?	What level of performance is acceptable?	What assessment methods will be used to collect data? How will data be evaluated and interpreted?	When will outcomes be measured?	Who needs to know the results? How can the program be improved?
Students should develop skills to: (1) recognize what constitutes meaningful engineering problems, (2) convert a physical problem to its mathe-matical equivalent where appropriate, (3) solve this mathe-matical problem, and (4) relate the result to the physical problem at hand	Core courses and discipline specific courses DSES-2210, DSES-4140, DSES-4200, DSES-4230, DSES-4250, DSES-4260, DSES-4270, DSES-4280, DSES-4640, DSES-4650, DSES-4810	80% of students should average a grade of 2.5 or better in this group of core and upper-level courses which are intended to contribute to the development of problem-solving ability.	Course grades.	Annually	Grade profiles and interpretations, course by course; will be reported to the department chair and Undergraduate Advisory Committee.
	Senior Portfolios	At least 95% of individual senior portfolios exhibit adequate problem solving and at least 60% receive good or better ratings.	Individual student portfolios	Annually	
	CO-OP survey on students' overall quality of work	90% of industry supervisors rate the students' quality of work "above average" or higher	CO-OP Evaluation Report	Annually	
	Special IDEA form question: "Is this course useful in developing problem solving ability?"	80% of students respond 4 or higher that these courses have been useful in developing problem solving ability.	IDEA form	Annually	
	Survey of Alumni 3-6 years after graduation.	80% of alumni surveys respond 4 or higher.	Alumni survey.		

This outcome states that, “IME students will demonstrate the ability to identify, model, analyze and solve challenging real life problems”. The user can enter any target performance levels for analysis. Based on the specific performance levels entered by the user, the system reports the corresponding assessment results for that outcome. For example, the target performance parameters applicable to outcome six during 2006 required that: 80% of graduating seniors achieve a combined GPA of 2.5 or greater on the associated courses itemized in the assessment matrix, the UAC subcommittee on portfolio evaluation agrees that 100% of individual student portfolios exhibit adequate levels of problem solving skills, 90% of industry co-op employers rate IME students’ quality of work as “above average” or higher, 80% of student respondents in the IDEA course evaluation survey respond with a rating of 4 or higher that these courses have been effective in enhancing problem solving ability, and 80% of alumni survey respondents report a 4 or higher that their education adequately prepared them in the area of design skills.

These targets are entered into the assessment software as shown in Figure 4 and the system generates the corresponding results illustrated in Figure 5. In this case, the results suggest that the IME undergraduate program was generally meeting target performance levels with the exception of student grades in associated courses which are very close to the target in every year except 2001. These results motivated the UGAC to introduce additional assessment measures including modification of the Team Evaluation Report, the IME Design sponsor survey and student portfolios to increase the diagnostic value of assessment for this outcome.

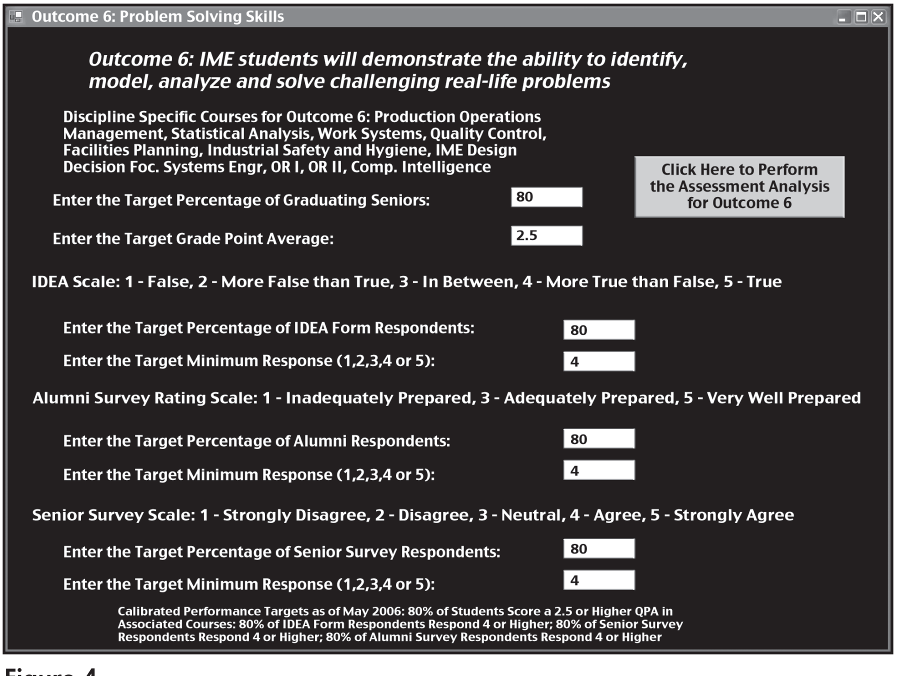

Figure 4.
Outcome Six Input Screen

Assessment Results for Transcript Data:

Year	Target GPA Value	Target Percent	Actual. Percent	Average GPA	Median GPA
2001	2.50	80.00%	67.57%	2.91	2.75
2002	2.50	80.00%	70.05%	3.19	3.32
2003	2.50	80.00%	75.51%	2.98	3.07
2004	2.50	80.00%	73.53%	3.04	3.21
2005	2.50	80.00%	82.86%	3.13	3.21
2006	2.50	80.00%	85.71%	3.12	3.21

Assessment Results for Alumni Survey Data:

	Target	Target	Actual	Rating/GPA	Respondent Rating Distribution:				
Year	Rating	Percent	Percent	Ratio	1	2	3	4	5
2001	4	80.00%	80%	1.51	0.0%	0.0%	20.0%	20.0%	60.0%
2002	4	80.00%	84%	1.39	0.0%	0.0%	16.0%	24.0%	60.0%
2003	4	80.00%	100%	1.57	0.0%	0.0%	0.0%	31.6%	68.4%
2004	4	80.00%	97%	1.53	1.8%	1.8%	0.0%	23.7%	72.8%
2005	4	80.00%	94%	1.43	2.0%	0.0%	3.9%	35.3%	58.8%
2006	4	80.00%	95%	1.45	2.5%	0.0%	2.5%	32.5%	62.5%

Assessatent Results for Senior Survey Data:

	Target	Target	Actual	Rating/GPA	Respondent Rating Distribution:				
Year	Rating	Percent	Percent	Ratio	1	2	3	4	5
2001	4	80.00%	100%	1.49	0.0%	0.0%	0.0%	66.7%	33.3%
2002	4	80.00%	100%	1.28	0.0%	0.0%	0.0%	90.0%	10.0%
2003	4	80.0096	93%	1.39	0.0%	3.6%	3.6%	67.9%	25.0%
2004	4	00.00%	94%	1.42	0.0%	5.5%	0.0%	50.0%	44.4%
2005	4	80.00%	91%	1.39	0.0%	0.0%	8.8%	47.1%	44.1%
2006	4	80.00%	100%	1.43	0.0%	0.0%	0.0%	52.6%	47.4%

Assesssent Results from the 'pre-2005' CO-OP Evaluation Reports:

Quality of Work:

Year	Outstanding	Very Good	Average	Below Avg.	Very Poor
2001	57.1%	35.7%	7.1%	0.0%	0.0%
2002	11.1%	77.8%	11.1%	0.0%	0.0%
2003	50.0%	50.0%	0.0%	0.0%	0.0%
2004	20.0%	80.0%	0.0%	0.0%	0.0%

Assessment Results for IDEA Form Data:

IDEA Form Statement: This course is useful in developing problem solving ability.

IDEA Form Rating Scale:

1. False, 2. More False than True, 3. In Between, 4. More True than False, 5. True

	Target	Target	Actual	Rating/GPA	Respondent Rating Distribution:				
Year	Rating	Percent	Percent	Ratio	1	2	3	4	5
2001	4	80.00%	0%	0.00	0.0%	0.0%	0.0%	0.0%	0.0%
2002	4	80.00%	0%	0.00	0.0%	0.0%	0.0%	0.0%	0.0%
2003	4	80.00%	096	0.00	0.0%	0.0%	0.0%	0.0%	0.0%
2004	4	80.80%	81%	1.36	1.3%	0.0%	18.2%	45.0%	34.8%
2005	4	80.00%	78%	1.29	3.5%	5.9%	12.6%	37.5%	40.4%
2006	4	80.00%	77%	1.29	4.4%	5.9%	13.2%	35.3%	41.2%

Figure 5.
Outcome Six Output Screen.

Today, the DSES undergraduate committee exploits this software tool for real time dynamic assessment in its group decision making process. This strategy introduces assessment results into the UAC discussion process using projection onto an overhead screen where each committee member can suggest alternative cuts on the assessment database to perform "what if" analyses. The first version of this decision support system was introduced in September 2003 with subsequent upgrades of the software introduced in 2004, 2005 and 2006. It has been continuously adapted to ongoing revisions in program objectives and outcomes over this period. The software system has been extremely effective in facilitating "live assessment". During regular undergraduate committee meetings, the software is run on a laptop and projected onto the conference room screen enabling committee members to simultaneously view the results during the committee discussion. Committee members can spontaneously suggest different ways to examine the data. For example, a committee member might observe underperformance on mathematics courses relative to target levels. The system could then use the system's transcript analyzer feature to drill down and examine the performance of IME students in individual mathematics courses or customized combinations of courses by individual graduating class. These types of features have not only been essential to effective calibration of target performance levels but, more importantly, they have been extremely useful for gaining the insights needed for finding curricular solutions in several problem areas.

The Impact of the Decision Support Tool

When it was first introduced, the software was used extensively for calibration of target performance levels based on the assessment results observed between 2001 and 2003. It was ultimately able to support an extensive IME curriculum re-development effort by providing the first truly effective tool for data driven curriculum reform. It facilitated the DSES undergraduate committee's capability to exploit the "low hanging fruit" associated with the curriculum reform process by pointing out obvious areas of need. To illustrate the dramatic impact that these tools have had on the curriculum development process during the first year of use, consider the following chronological summary of the educational assessment activities excerpted from the minutes of six undergraduate committee meetings from September through December 2003:

September 17, 2003 - The new assessment software developed during summer 2003 was presented to the DSES UAC. It was explained how this software, developed with inputs from the faculty, students, alumni and Industrial Advisory Board is specifically designed to maximize the utility of assessment results in curriculum development and evaluation.

October 1, 2003 - The UAC reviewed outcome 1 using the software system. The conclusion was that the curriculum was meeting the systems perspective objective. The UAC also reviewed assessment results relative to outcome 2, where it was determined that students needed greater understanding of the parallels between manufacturing and services. The possibility of an advanced supply chain course was subsequently re-visited.

Greater infusion of service operations examples in IME courses was also suggested. The UGAC also reviewed outcome 3, where assessment results suggested that more emphasis on practical computing was needed. The 50/50 response reported on this outcome from the alumni survey was considered a red flag. Subsequent suggestions for infusion of computing including the use of VBA Excel in OR II and the use of SAS in the Statistical Analysis course.

October 15, 2003 - Further discussion of outcome 3 assessment results led to a proposal to swap the C Programming for Engineers course for the current Computer Science I course and the addition of a stand alone simulation course.

October 29, 2003 - The UAC used the assessment software to examine outcomes 4, 5 and 6 leading to the following re-calibrations in target performance levels:

- Outcome 4 – 80% was changed to 90% on senior survey results.
- Outcome 6 – Faculty grading appears significantly harsher than students' self assessment. No subsequent action plan was suggested.

November 12, 2003 - The UAC applied the assessment software to examine program performance with respect to educational outcomes 7 and 8. It was agreed to change the outcome 7 grade parameters from 80/2.5 to 90/2.0 as a recalibration step. The assessment results from outcome 7 also suggested that mathematics and science is a potential problem area for IME's that would be discussed at a future meeting. Assessment results for outcome 8 suggested that IME's are not achieving program objectives in the area of team skills, but more specific assessment data would be more diagnostic. Subsequently, a revised team evaluation form was designed for the assessment system that would replace the most recent, September 2003 version.

December 3, 2003 - Initial assessment of educational outcomes was completed through re-calibration of outcomes 9-12 using the assessment software. The results are summarized below:

- Outcome 9 - We appear to be doing all right – no changes recommended.
- Outcome 10 - Change alumni question to 50% in the assessment software and change the senior survey question parameter to 80% responding 4 or higher.
- Outcome 11 - Change the course mix to exclude H&SS courses and add DSES.4260, change the alumni survey parameter from 70% to 80%, change senior survey parameter from 80% to 90%
- Outcome 12 - Change the percent of alumni participating in continuing education to 60% from 50%.

The assessment software has continued to be used to guide the UAC's curriculum development work in subsequent academic years for "fine tuning", trend analysis, and identifying increasingly useful measurement opportunities. It has supported many ad hoc assessment initiatives relating to courses using its transcript analyzer and identifying specific problem areas leading to the direct application of measurement instruments outside of the software tool. Based on use of the software tool over several years, the

DSES undergraduate committee established very specific performance targets for the IME undergraduate program and has undertaken many specific measures in response to the results as summarized below:

Outcome 1 – Systems Thinking:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) 80% of student respondents in the IDEA course evaluation survey respond with a 4 or higher that these courses have been effective in enhancing the ability to apply the systems approach.

C.) 80% of alumni survey respondents respond with a 4 or higher that they are able to apply the system approach in problem solving.

D.) 80% of graduating senior survey respondents report with a 4 or higher that they understand and can apply the systems approach.

The system results for Outcome 1 have generally supported the conclusion that the IME undergraduate program has been consistently meeting or exceeding target performance levels over the past six years.

Outcome 2 – Manufacturing/Services:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) 80% of student respondents in the IDEA course evaluation survey respond with a 4 or higher that these courses have been effective in enhancing the manufacturing and service systems.

C.) 80% of alumni survey respondents respond with a 4 or higher that they are able to apply knowledge of manufacturing and service systems in problem solving.

D.) 80% of graduating senior survey respondents report with a 4 or higher that they have knowledge of both manufacturing and service systems.

During the past four years, system results for this outcome have motivated the development of a new course sequence in production operations and value chain management.

Outcome 3 – Computing Skills:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) 80% of student respondents in the IDEA course evaluation survey respond with a rating of 4 or higher that these courses have been effective in enhancing the ability to apply knowledge of computing.

C.) 80% of alumni survey respondents respond with a 4 or higher that they are able to apply knowledge of computing.

D.) 80% of graduating senior survey respondents report with a 4 or higher that they have the ability to apply knowledge of computing.

During the past four years, system results for Outcome 3 have motivated the changes in the computer programming requirements for the IME program, the development of a new course in discrete event simulation and the addition of computing modules to a number of existing IME courses.

Outcome 4 – Management Skills:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) 60% of alumni survey respondents respond with a 4 or higher that their education prepared them to manage others.

C.) 90% of graduating senior survey respondents report with a 4 or higher that they have been prepared to become effective managers.

D.) 90% of industry supervisors rate IME students "above average" or higher in judgment, enthusiasm and overall performance

During the past four years, system results for this outcome have motivated the addition of a new management elective course requirement to the IME program.

Outcome 5 – Design Skills:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) UGAC subcommittee on portfolio evaluation agrees that 100% of individual student portfolios exhibit adequate levels of design skills.

C.) 80% of alumni survey respondents report with a 3 or higher that their education adequately prepared them in the area of design skills.

D.) 80% of graduating senior survey respondents report with a 4 or higher that they have developed the ability to design products, services facilities, equipment processes and systems.

E.) 80% of student teams are rated at 3 or higher on design skills in the capstone design course.

During the past four years, system results reported for Outcome 5 have motivated the development of additional diagnostic measures to assess IME student design skills including the IME capstone sponsor evaluation and the student portfolio evaluations measure.

Outcome 6 – Problem Solving:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) UGAC subcommittee on portfolio evaluation agrees that 100% of individual student portfolios exhibit adequate levels of problem solving skills.

C.) 90% of industry co-op employers rate IME students' quality of work as "above average" or higher.

D.) 80% of student respondents in the IDEA course evaluation survey respond with a rating of 4 or higher that these courses have been effective in enhancing problem solving ability.

E.) 80% of alumni survey respondents report a 4 or higher that their education adequately prepared them in the area of design skills.

During the past four years, system results have reinforced that the IME program is generally meeting target levels with the exception of student grades in associated courses which are very close to the target in most years. These results motivated the introduction of additional assessment measures including modifications to the Team Evaluation Report, the IME Design sponsor survey and student portfolios to increase their diagnostic value.

Outcome 7 – Math/Science:

A.) 90% of graduating seniors achieve a combined GPA of 2.0 or greater in associated courses in mathematics and science.

B.) 80% of alumni survey respondents report a 4 or higher that their education adequately prepared them in the area of math and science.

C.) 80% of graduating senior survey respondents report a 4 or higher that they have been adequately prepared in math and science.

During the past four years, system results have indicated slight underperformance relative to targets, especially in more recent years. Subsequent committee discussions of problems for IME students relative to mathematics and science preparation have resulted in several hypotheses including over reliance on software versus analysis in calculus courses. A number of proposals have been developed based on in-depth analysis of transcript data and survey results including increased emphasis on first principles such as expectation in statistics courses, retuning to a two-course sequence in statistics, and imposing more rigorous use of probability (and mathematics generally) in IME courses. To date, no curricular actions have been taken, but it is an ongoing area of discussion within the UGAC.

Outcome 8 – Communication Skills:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) UGAC subcommittee on portfolio evaluation agrees that 100% of individual student portfolios exhibit adequate levels of communication skills.

C.) 90% of industry co-op employers rate IME students' ability to relate with others as "works well" or better.

D.) 90% of alumni survey respondents report a 4 or higher that their education adequately prepared them in the area of communication skills.

E.) 80% of graduating senior survey respondents report a 4 or higher that they have developed strong communication skills.

F.) 80% of student teams are rated at 4 or higher on writing and presentation skills in the capstone design course.

During the past four years, system results for communication skills suggest strong performance in every dimension except the alumni survey which, in this case, may be the most valid measure. In reviewing this result, the undergraduate committee elected not to take immediate action pending implementation of a new Institute level communication requirement introduced in 2006. The committee will carefully monitor IME alumni survey results as well as communication skills exhibited in IME student portfolios in future years to assess whether this initiative is having the desired impact.

Outcome 9 – Team Skills:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) 90% of industry co-op employers rate IME students' dependability as "above average" or higher.

C.) 60% of student respondents in the IDEA course evaluation survey report a rating of 3 or higher that these courses have been effective in enhancing team skills.

D.) 80% of student teams are rated at 3 or higher on team skills in the capstone design course.

E.) 90% of alumni survey respondents report a 4 or higher that their education adequately prepared them in the area of team skills.

During the past four years, system results have supported the conclusion that the IME program is meeting target performance levels in this area.

Outcome 10 – Educational Breadth:

A.) 80% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) 80% of alumni survey respondents report a 3 or higher that they had adequate or better exposure to humanities and social sciences.

C.) 50% of alumni survey respondents report that one or more of their humanities and/or social sciences courses significantly influenced their perspective or world view.

D.) 90% of graduating senior survey respondents report a 4 or higher that they are informed citizens broadly educated in the humanities and social sciences.

During the past four years, system results reveal performance close to target levels although some deterioration in senior survey responses is evident in recent years. Strong performance on transcripts suggest that recent IME students do well academically in humanities and social science courses but may not be deriving much benefit from the experience. This has motivated a discussion within the undergraduate committee where several hypotheses have emerged including excess emphasis on minors and dual majors which severely restrict electives. Other possible explanations includepoor course selections (e.g., electing "easy" course options) due to inadequate advising or an overemphasis on "skills" courses (e.g., writing, foreign languages) relative to more broadening course selections. Discussions of possible implications for student advising are continuing.

Outcome 11 – Professional Ethics:

A.) 90% of graduating seniors achieve a combined GPA of 2.5 or greater in associated courses.

B.) 80% of alumni survey respondents report a 3 or higher that treatment of ethical issues in their undergraduate education was adequate or better.

C.) 90% of graduating senior survey respondents report a 4 or higher that they are prepared to practice engineering in a socially responsible and ethical manner.

Assessment results suggest performance is meeting target levels with the exception of the deterioration in senior survey results during the last two years possible due to increasing awareness of ethical issues among IME students and a corresponding failure within the program to deliver adequate content on the issue. Subsequently, the committee member serving as the representative to the School of Engineering Curriculum Committee (SoECC) referred the issue to that group for possible school-level action.

Outcome 12 – Lifelong Learning:

A.) 90% of alumni survey respondents report a 4 or higher that their undergraduate education stimulated them to grow and learn.

B.) 90% of industry co-op employers rate IME students' learn "readily" or "quickly".

C.) 60% of recent alumni should report that they have participated in graduate courses or other professional development activities.

D.) 80% of graduating senior survey respondents report a 4 or higher that they learned in a stimulating environment that prepared them for continued growth and learning.

E.) 50% of graduating senior survey respondents report that they joined a relevant professional society prior to graduation.

The system results for this outcome have generally supported the conclusion that the IME undergraduate program has been consistently meeting or exceeding target performance levels over the past six years.

Summary and Conclusions

This chapter has described how the program level objectives and outcomes of the Rensselaer's IME undergraduate program reflect the unique history and character of the program. The subsequent formulation of these educational outcomes motivated the development of a specialized collection of assessment instruments designed to achieve a balance of individual student and student group measures, direct and indirect measures, and to report from a diversity of program constituents including the faculty, students, employers and alumni. The assessment system was also designed to assure that each outcome was evaluated, at least in part, from specific curriculum requirements applicable to all IME students as well as making constructive use of elective requirements, e.g., student evaluations and grades from technical elective courses, etc. To make maximum use of this collective knowledge base, the undergraduate committee responsible for the IME undergraduate program developed a software system to facilitate live, dynamic assessment of educational outcomes. The system directly facilitated the committee process for calibration of target performance levels and flexible analyses of assessment results to guide the curriculum development process. Following its introduction, this system had an immediate and dramatic impact on the IME undergraduate curriculum. In subsequent years, it has played an important role in the fine tuning and continuous improvement of the program. It has enabled the DSES undergraduate committee to successfully solve the easiest 80% of curriculum development challenges and enabled it to focus the majority of effort in recent years to deal with more difficult issues.

References

Accreditation Board for Engineering and Technology. (2005). Criteria for Accrediting Engineering Programs, Effective for Evaluations During the 2006-2007 Accreditation Cycle.

Aldridge, M.D. & Benefield, L.D. (1998). "A Model Assessment Plan," *ASEE Prism*, 22-28.

Bullington, S., Heragu, S.S. & Malmborg, C.J. (2000). „Preparing students for careers in material handling," *Journal of Engineering Education*, 89 (10), 439-441.

Christy, A.D. & Lima, M.B. (1998). "The use of student portfolios in engineering instruction," *Journal of Engineering Education*, 87 (2), 143-148.

Deek, F.P., Hiltz, S.R., Kimmel, H. & Rotter N. (1999). “Cognitive assessment of students’ problem solving and program development skills,” *Journal of Engineering Education*, 88 (3), 317-326.

Felder, R.M. & Brent, R. (2003). “Designing and teaching courses to satisfy ABET engineering criteria,” *Journal of Engineering Education*, 92 (1), 7-25.

Gronlund, N.E. (1999). *How to Write and Use Instructional Objectives*, 6th edition, Prentice Hall.

Heragu, S.S., Graves, R.J., Malmborg, C.J. & Jennings, S. (2001). „Teaching material handling via multimedia tools,“ Proceedings of the International Conference on Engineering Education, Oslo, Norway.

Heragu, S.S., Graves, R.J., Malmborg, C.J., Jennings, S. & Newman, D. (2003). „Multi-media tools for materials handling classes,“ *European Journal on Engineering Education*. 28 (3), 375-393.

McGourty, J. Sebastian, C. & Swart, W. (1998). “Developing a comprehensive assessment program for engineering education,” *Journal of Engineering Education*, 87 (4), 355-361.

Panitz B. (1996). “The student portfolio: A powerful assessment tool,” *ASEE Pris*m, 24-29, March-April.

Tien, J.M. & Berg, D. (2003). “A case for service systems engineering,” *International Journal of Systems Engineering*, 12, (1), 13-39.

CHAPTER 7

STRUCTURED AND SYSTEMATIC ASSESSMENT: A SUCCESSFUL AND SUSTAINABLE CIVIL ENGINEERING EXAMPLE

Allen C. Estes
California State Polytechnic University
Ronald W. Welch
The University of Texas at Tyler
Stephen J. Ressler
The United States Military Academy

Introduction

Assessment and continuous improvement have become a major part of the accreditation process for engineering programs in the United States. ABET Inc., the accrediting agency for engineering programs, has established some guidelines and requirements for this assessment. The truly successful programs have embraced the philosophy that improvement only comes from introspection, examination, consultation with constituents, and then reasoned action based on the analysis and synthesis of the data received. Each engineering program is unique and must therefore create an assessment program that incorporates the needs of its university and captures the strengths and areas for improvement of the curriculum, facilities, faculty, resources, and students that comprise the program. A program that cobbles together an assessment program solely for accreditation purposes will gain little from it and will experience significant frustration. Good assessment takes time and effort which are both precious resources, but to be sustainable over time, the process also has to be efficient. Although someone needs to lead the effort, the work can be effectively shared among various faculty members, which also leads to greater buy-in from the faculty at large.

This paper shares the systematic assessment program used by the civil engineering program at the United States Military Academy over the last eight years. The program is characterized by a process that builds on the university assessment system, a flexible slow loop assessment cycle, a highly structured fast loop cycle, an advisory board that has evolved over time, a course assessment process that collects data which in turn rolls up into an annual program assessment, and almost a decade of documented results. For each outcome and objective, there are a series of performance measures and desired standards that are based on student performance, survey results, and instructor ratings. Based on the results, a rating is assigned each year and recommendations are made for the future. The follow-up reporting on those recommendations closes the feedback loop and starts a new cycle of assessment.

The Assessment Model

The United States Military Academy (USMA) has described its assessment process for the curriculum and instruction in its widely circulated publication, "Educating Army Leaders for the 21st Century" (USMA, 2000). The academic program goals are developed from the needs of the Army. Those goals are attained through articulating a learning model that includes the structure, process, and content of the desired learning experience; designing an appropriate curriculum; designing the individual courses that comprise the curriculum; and implementing the instruction. The university assessment system consists of four phases which are linked to the curriculum and instruction steps as shown in Fig. 1. The USMA Civil Engineering (USMA CE) program has adopted the university model using the same four phases of assessing the learning model, program design, implementation, and outcomes for its program assessment. Because there is a requirement to develop program outcomes and objectives that meet the needs of program constituents, a phase 0 element was added to accommodate this.

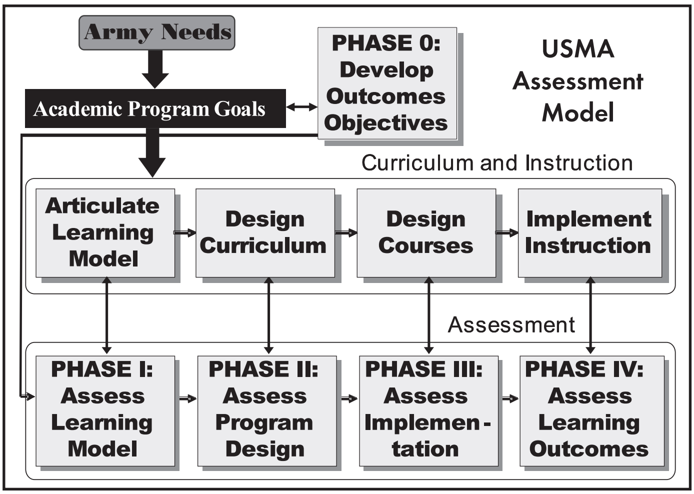

Figure 1
The USMA assessment process (USMA, 2000) on which the USMA CE program assessment process is based

The USMA CE program has adopted a two-loop cycle suggested by the ABET literature as shown in Fig. 2. While ABET no longer uses the two loop cycle, it remains a valuable means to separate the major program changes from the routine changes made on a year-to-year basis. The slow loop is completed every three years – immediately after an accreditation visit and at the mid-point between visits. This allows any major changes to be implemented and assessed prior to an accreditation visit. The slow loop encompasses phases 0, I, and part of II of the assessment model where changes to the objectives, outcomes, and learning model are made. Major revisions of the curriculum occur in the slow loop. The conduct of the slow loop assessment is totally flexible and the format is based on the issues that arise over a three year period.

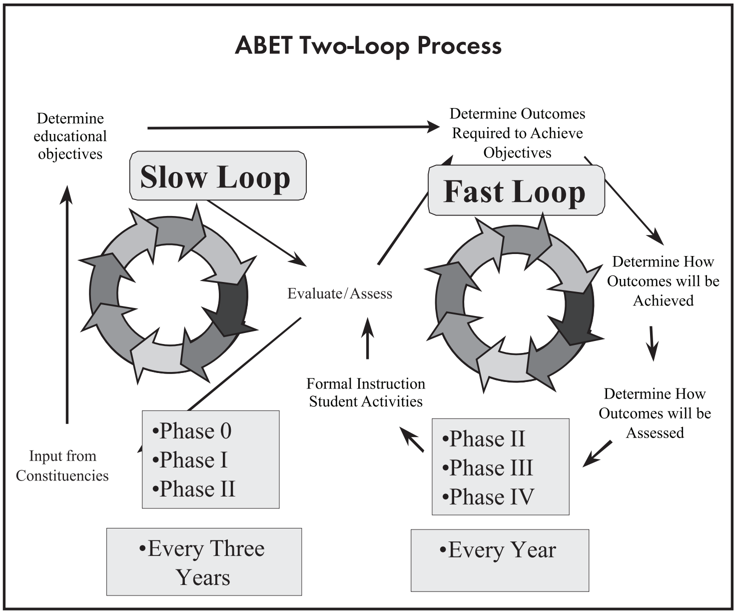

Figure 2
The USMA CE program assessment process adopts the two loop process that appeared in early ABET literature

The 2003 slow loop assessment (Estes and Welch, 2004) involved a zero-based, bottom up look at the CE curriculum caused by some changes in the institutional priorities. The process involved six teams, working independently and returning to a larger group, to iteratively devise a new curriculum. The result was the development of seven new courses in the Civil and Mechanical Engineering programs and the largest curriculum change in two decades. Because this new curriculum was still being assessed as the new courses were developed and experiencing initial offerings, the 2006 slow loop assessment involved no major curriculum changes. The relevant issues were collected and addressed using GroupSystems (2006) software to assemble input in an efficient manner. The software enables a group dialogue on the computer where all parties are communicating simultaneously and the conversation is recorded in writing. Decisions and program changes were made in subsequent faculty meetings and provided to the advisory board for input. The program outcomes and objectives were revised. Three new program outcomes were added in response to the American Society of Civil Engineers Policy 465 which has established a body of knowledge (ASCE, 2004a) for civil engineers and makes the master's degree or equivalent a requirement for professional licensure.

In contrast, the fast-loop assessment is conducted annually and follows a rigid, systematic format. The process shown in Fig. 3 is sustainable because the data are collected in the same manner every year and minor changes are made based on the

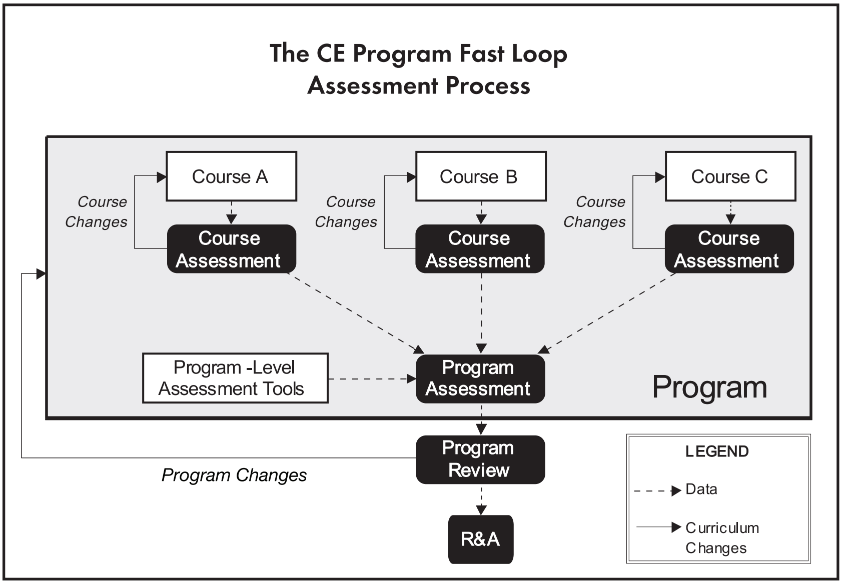

Figure 3
The USMA CE fast loop process is a systematic annual assessment that is highly structured and relies heavily on input from program constituents and from the assessment of individual courses in the curriculum

input from constituents. Annual input is collected at the program level in the form of student surveys, graduate surveys, Fundamentals of Engineering exam results, and advisory board minutes. The CE program conducts annual course assessments for every course in the curriculum. Because there is a consistent systematic format for the course assessment, data regarding student performance, course objectives, and instructor ratings can be collected and assembled. The fast loop assessment culminates in a formal program assessment briefing to the department head. The briefing covers minor changes to the curriculum resulting from the latest round of course assessments (Phase II from Fig. 1); implementation in terms of faculty performance, student performance and resources (Phase III); and the assessment of program objectives and outcomes (Phase IV). The most substantial portion of the briefing is the implementation. Student performance assessment includes the capstone project, independent study projects and competitions, summer intern experience, student chapter activities and student awards. The faculty are assessed based on qualifications, teaching ratings, professional society participation, service activities, scholarship, and support to the Army over the previous year. Student and faculty statistics are analyzed in terms of enrollments, diversity, and quality. Resources are assessed based on facilities, budget, laboratories, computers, support staff, and external support. The briefing begins with the recommendations made at the previous program assessment along with a status report on their implementation. The briefing ends with the new recommendations based on the annual assessment.

Progress on these new recommendations will become part of the next program assessment, which is how the feedback loop is closed. The program briefing is the record of the annual assessment and is the first document in the annual assessment notebook that contains all of the raw data used in the assessment. A summarized version of the results is provided to the Dean in the annual Review and Analysis briefing.

Program Constituencies and Their Input

The program constituents are the customers, the clients – those whom the program is designed to serve. The USMA CE program has identified its constituents as the Army, the Corps of Engineers, the current faculty, the students, the graduates/alumni, and the civil engineering profession. The USMA CE program is one of the few programs that lists a specific branch of the Army (i.e., Corps of Engineers) as a constituent. The constituencies were involved in the development of program outcomes and objectives and continue to provide survey and advisory board input as to whether they need to be revised. Fig. 4 shows these constituencies and the formal input that they provide. The CE program is able to take advantage of many institutional level surveys to collect data.

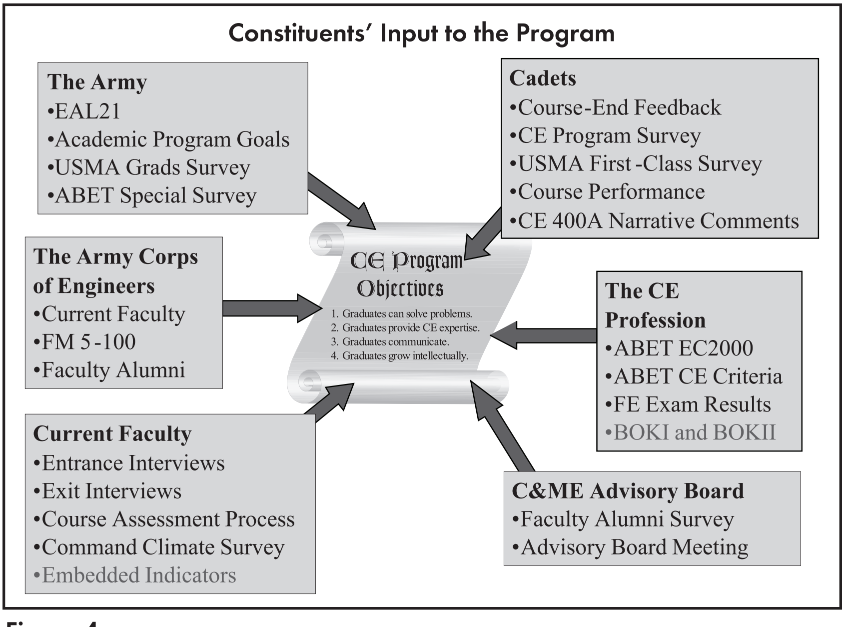

Figure 4
The USMA CE program has identified it constituencies and solicits input from each using a variety of tools

Because the Army is the industry into which each of the USMA graduates will enter upon graduation, the institution puts extensive thought into the needs of the Army. The Army needs leaders of character who possess ethics, leadership and team skills, versatility, communication skills, and dedication to lifelong learning and who understand technology, information systems, history, people and organizations and cultures (USMA, 2000). The USMA academic program goals are directly based on these needs. Because the institution is so focused on this area, there is very little the CE program needs to do to discover the needs of the Army. USMA seeks input from Army leaders on the quality of its graduates through surveys sent to graduates and commanders of graduates for year groups several years after graduation. A special tri-annual institutional survey is sent to graduates directly in support of accreditation preparation. The programs provide input on what questions to ask. Data on graduates from the Civil Engineering program can be separated from the graduates at large. The data are the most useful tools available for assessing program objectives.

The program is more directly engaged in determining the needs of the Corps of Engineers, the branch of the Army that most graduates will choose. The doctrinal field manual FM 5-100 Engineer Combat Operations (U.S. Army, 1993) is a major source of what graduates are expected to do, with particular emphasis on sustainment engineering. Because most faculty members are also Corps of Engineer officers, they provide feedback on behalf of both the faculty and the Corps of Engineers. Most faculty return to the field Army after teaching, and their survey input is collected.

The current faculty provide input through a variety of means to include entrance surveys, exit interviews, an institution-wide command climate survey and input at various faculty meetings. Faculty members prepare the course assessments through which so much of the program data are collected. Student feedback is obtained in every course through web-based end of course surveys that evaluate the effectiveness of the course and their individual instructor. Some questions are common throughout the institution, which allows a comparison of performance across departments. The questions asked at the CE program level are directly correlated to a model that defines excellent teaching (Estes et.al,. 2006; Estes et.al,. 2005a; Estes and Ressler, 2003). These data can be compared across courses and over time to assess the quality of teaching in the CE program.

Students also complete exit surveys at both the program and institutional levels at the time of graduation. In addition, the students address the appropriateness of the program outcomes and objectives in a journal entry and survey administered in CE400, the Civil Engineering professional practice course.

The needs of the civil engineering profession are obtained through accreditation criteria. While the EAC provides the general criteria (ABET, 2006), the American Society of Civil Engineers (ASCE) writes the program criteria (ABET, 2006; ASCE, 2004b) that are specific to civil engineering programs. Most recently, additional input has been provided through the body of knowledge (BOK I and BOK II) efforts supported by ASCE committees (ASCE, 2004a). Active faculty participation on professional society educational and technical committees provides input as well.

Finally, input is received through annual meetings of the CE advisory board which is comprised of members of various program constituencies. The advisory board has evolved significantly over the past six years. The initial advisory board consisted of department alumni who returned to West Point for a designated weekend, received an overview of the program and completed a survey form. The next iteration was a board of designated individuals who represented specific constituencies (faculty, students, and outside members who represented alumni, the Army, other institutions and the profession). The CE program director chaired the one day meeting, asked the board for input on specific issues, and recorded the comments in formalized minutes. Today the board consists of 12 very prominent outside representatives from industry, academia, and the Army. The board leader is a one of these members. After some preparatory work, the board meets annually for a day at West Point. They receive update briefings from the CE program director, interview students, interview faculty, meet in executive session and present their thoughts to the CE program leadership. A written report follows, and the CE program director responds with written feedback to the report. As the board evolves, the quality of the input and the influence the board has with the rest of the institution has grown as well.

Program Objectives

As defined by the accreditation criteria (ABET, 2006), program objectives are defined as "broad statements that describe the career and professional accomplishments that the program is preparing graduates to achieve." With considerable input from program constituents, the USMA CE program developed the following objectives:

1. As Army leaders, graduates solve complex, multi-disciplinary problems effectively, to include:
    - recognizing and fully defining the physical, technological, social, political, and economic aspects of a complex problem;
    - using a methodical process to solve the problem;
    - demonstrating creativity in the formulation of alternative solutions;
    - using appropriate techniques and tools to enhance the problem-solving process;
    - working effectively on teams; and
    - developing high-quality solutions that consider the technological, social, political, economic, and ethical dimensions of the problem.
2. Graduates provide appropriate civil engineering expertise to the Army, when called upon to do so.
3. Graduates communicate effectively.
4. Graduates continue to grow intellectually and professionally—as Army officers and as engineers.

The slow loop assessments are used to update and revise these objectives. Constituent surveys and advisory board meetings provide much of the input data. Changes to objectives must be made slowly, as there is significant lag time between implementation and ability to assess the effect. The objectives were not changed in the 2003 slow loop and were modified slightly in 2006.

The process of ongoing evaluation of the extent to which these program objectives are being attained is accomplished through survey data assembled in the fast loop process. Direct measures of performance are much more difficult to obtain than for outcomes because the attainment occurs several years after graduation. Institutional surveys have been the best tool and have provided some excellent data on professional society participation, professional licensure and attainment of masters degrees on the parts of graduates. Since the institutional survey polls commanders as well as graduates, data are attained on graduate performance from their current employers. Based on the survey results, the program director provides an annual rating for each program objective on an annual basis with the results shown in Fig. 5. A rating of 4 denotes successful accomplishment of the objective.

Evaluation of Objectives: Summary

#	Civil Engineering Program Objective	02	03	05	06
1	**Solve complex multi -disciplinary problems effectively**	4	4	4	4
2	**Provide appropriate civil engineering expertise to the Army**	3+/ 4-	3+/ 4-	3+ /4-	3+ /4-
3	**Communicate effectively**	4	4	4	4
4	**Continue to grow intellectually and professionally –as Army Officers and engineers**	5	4+	5	5

Figure 5
Based on the survey results from graduates and their employers, an annual rating from 1 to 5 is given which evaluates the degree to which the CE program objectives are being attained

Program Outcomes

Program outcomes are defined as "statements that describe what students are expected to know and be able to do by the time of graduation." (ABET 2006) Each program is different and is expected to create outcomes that meet the needs of that specific program and enable the graduates to ultimately attain the program objectives. To ensure that certain standards are met within any program, the accreditation criteria 3(a-k) specify some minimum attainments that must be included within every program's outcomes. The current USMA CE program outcomes are:

1. Graduates design civil engineering components and systems.
2. Graduates demonstrate creativity, in the context of engineering problem-solving.
3. Graduates solve problems in the structural, construction management, hydraulic, and geotechnical discipline areas of civil engineering.
4. Graduates solve problems in mathematics through differential equations, calculus-based physics, and general chemistry.
5. Graduates design and conduct experiments, and analyze and interpret data.
6. Graduates function effectively on multidisciplinary teams.
7. Graduates describe the roles and responsibilities of civil engineers and analyze the issues they face in professional practice.

8. Graduates use modern engineering tools to solve problems.
9. Graduates write effectively.
10. Graduates speak effectively.
11. Graduates incorporate knowledge of contemporary issues into the solution of engineering problems.
12. Graduates draw upon a broad education necessary to anticipate the impact of engineering solutions in a global and societal context.
13. Graduates are prepared and motivated to pursue continued intellectual and professional growth—as Army officers and engineers.
14. Graduates explain the basic concepts of management.
15. Graduates explain the basic concepts of business and public policy.
16. Graduates are leaders of character.

The outcomes assessment process consisted of developing program outcomes, documenting input from constituencies, identifying where in the curriculum each outcome was addressed, creating performance measures and desired standards for each outcome, evaluating the student performance against these measures on an annual basis, and then making program decisions/changes based on these results. Because the program has control of its students through graduation, it is much easier to obtain good data on which to assess student attainment of outcomes than it is for objectives.

As with objectives, the slow loop process is used for revision of outcomes based on input from constituents. No changes to outcomes were made in 2003, but the 2006 slow loop resulted in some substantial changes. The ASCE Policy 465 and the creation of a body of knowledge for civil engineers created new requirements for the USMA CE program. These requirements currently appear in new civil engineering program criteria that have been approved by the Engineering Accreditation Commission and should take effect for programs evaluated in 2008 (ASCE, 2005a). The addition of outcomes 14, 15, and 16 are directly attributable to this policy. Outcomes 14 and 15 are taken directly from the new program criteria and outcome 16 was modified to reflect the unique emphasis on leadership at the Military Academy. The latest supplement to the Body of Knowledge (ASCE, 2005b) used the cognitive levels associated with Bloom's taxonomy (Bloom, 1956) to classify the desired attainment level in various outcomes. The USMA CE program outcomes were reworded to choose action verbs that more clearly define the cognitive level being sought. Several USMA CE faculty members are serving on ASCE Committee on Academic Prerequisites for Professional Practice sub-committees that are implementing this policy. The Curriculum sub-committee has formally assessed the current USMA CE curriculum with regard to compliance with the new body of knowledge (Welch and Estes, 2006a; Estes et.al., 2005b).

The course assessment process helps identify the contributions of various courses to the overall program outcomes. Fig. 6 shows the results where course directors have submitted a rating of 1 (no contribution) to 5 (very large contribution) for each outcome. Those courses that provide a rating of 4 or 5 for a particular outcome become a good source for a direct measure of student performance. Similarly, the course notebooks in which samples of student work are assembled should include examples that support the outcome. If no course attains a rating of at least 3 for a particular outcome, a curriculum change to include the outcome may be in order. The outcomes shown in Fig. 6 reflect those in effect prior to the recent slow loop assessment.

	Course											
Program Outcome	**CE300**	**CE364**	**CE371**	**CE380**	**CE400**	**CE404**	**CE460**	**CE471**	**CE483**	**CE491**	**CE492**	**Max Value**
Apply the engineering thought process to design CE components and systems	5	5	5	3	3	5	3	5	4	4	5	5
Creativity	3	2	4	2	4	4	5	4	4	2	5	5
Proficiency in structural engineering	4	4	3	1	2	5	3	4.5	5	5	5	5
Proficiency in environmental engineering	1	1	2	3	1	1	3	1	1	1	3	3
Proficiency in hydrology & hydraulic engineering	1	1	2	5	2	1	1	1	1	1	4	5
Proficiency in geotechnical engineering.	1	3	4	1	1	1	2	1	2	2	4	4
Proficiency in mathematics	4	5	3	4	1	3	3	3	4	4	4	5
Proficiency in calculus-based physics	4	3	2	2	1	2	1	3	3	3	1	4
Proficiency in general chemistry	1	1	2	1	1	1	1	1	2	1	1	2
Design and conduct experiments, analyze and interpret data in civil engineering discipline areas	4	4	4	1	1	3	2	2	5	1	2	5
Function on multidisciplinary teams	1	1	3	4	4	2	4	3.5	1	1	5	5
Roles and responsibilities of civil engineers and the issues / professional practice	2	2	4	3	5	4	5	3	3	2	5	5
Use the modern engineering tools necessary for engineering practice	3	4	4	5	1	5	3	3	4	5	5	5
Write effectively	3	5	3	3	4	4	3	3	5	3	5	5
Speak effectively	3	2	2	2	2	2	4	2	3	2	5	5
Knowledge of contemporary issues	2	2	3	4	3	3	3	3	3	2	3	4
Broad education to understand the impact of engineering solutions in a global/societal context	2	3	4	3	3	3	3	2	3	2	3	4
The preparation for and willingness to pursue continued intellectual and professional growth	3	3	3	3	5	2	4	3	3	3	4	5

Assessed by Course Director: *1=No Contribution 2=Small Contribution 3=Average Contribution 4=Large Contribution 5=Very Large Contribution*

Figure 6
Ratings are provided through the course assessment process that rate the contributions of each course toward attainment of the CE program outcomes

The USMA CE program has developed performance measures for each outcome using the data that are collected on an annual basis in the fast loop process. The sources of data in order of priority from most to least credible are F.E. exam scores, outside agency evaluations, student performance on course requirements, survey data, instructor course assessment ratings and satisfactory course completion. The best data are results from the Fundamentals of Engineering examination. The exam is administered nationally, is unbiased by faculty members, and almost every USMA CE major takes the exam. Performance data are provided in each of the subject areas covered on the exam and can be used to assess attainment of some program outcomes.

Outside visitors provide credible data because they do not hold the same bias as faculty. Experts from industry and professional practice are typically invited to the Academy on Projects Day to evaluate student capstone and independent study projects. The evaluators complete grade sheets that are tailored to correlate to program outcomes. The degree to which students communicate orally and in writing is certainly evaluated. It is an opportunity to attain direct measurement data on some of the more difficult-to-quantify outcomes such as creativity and understanding contemporary issues.

Many programs grapple with how to attain direct measures of student performance. Much of the accreditation literature (ABET, 2005; Rogers, 2003; Rogers, 2004) have indicated that survey data and course grades are useful but not sufficient to demonstrate attainment of outcomes. The next section of this paper specifically addresses how the USMA CE program uses embedded indicators to provide direct measures of performance. Surveys still provide useful data that can contribute to the overall assessment of a program outcome. USMA surveys all of the graduating seniors and the CE program conducts a more targeted survey. The questions typically require a Likert scale response that provides a numerical score that can be compared against other questions and to the same question from previous year groups. Often questions and their responses can be directly applicable to a specific outcome.

In the course assessment process, the faculty member is making judgments about the degree to which students attained the course objectives. When these course objectives can be tied directly to a program outcome, the instructor rating becomes another data point to consider in making an overall assessment. This is particularly useful in laboratory courses where some course objectives relate directly to outcome 5: design and conduct experiments.

While course grades are considered a weak indicator, the data can be useful. Course grades as an assessment tool is much enhanced by the institutional policy prescribing the use of criterion-referenced grading. With norm-referenced grading (i.e., grading on a curve where a student's grade is assigned relative to the performance of other students), grades are a poor assessment measure, because there is no clear connection between the level of performance and the grade. With criterion-referenced grading, the connection is explicit because each student is assessed relative to the learning outcomes.

When a particular required course such as international relations, economic policy, or physics clearly correlates to a particular outcome, successful completion of that course is a data point that should at least be considered. Desirable standards are created for these data points and performance measures are created for each program outcome. Fig. 7 shows an example of the performance measures for program outcome 1: graduates design civil engineering components and systems. There are five questions on the Civil Engineering First Class Survey (CE1CS) and the USMA First Class Survey (1CS) that relate to this outcome. The students rated their abilities on a Likert scale of 1 (strongly disagree) to 5 (strongly agree). The desired average response is between 4 (agree) and 5 (strongly a gree). Fig. 7 shows that the standard was met on four out of the five questions for the most recent survey. The benchmark is the average response to the same question from Civil Engineering students over the past five years. On the USMA survey, the responses of the civil engineers can be compared to those of the rest of the student body. The embedded indicators are direct measures of student work. The performance on the capstone project as determined by the embedded indicator was slightly lower than the standard, but the reports from the judges who helped grade the capstone and independent study presentations were highly favorable.

Assessment of Outcomes:

- **OUTCOME 1: Graduates design civil engineering components and systems**

Survey Data:

Tool	#	Item	Std.	Bench mark	Avg. Resp.	Remarks
CE1CS	15	I can formulate a sound methodology for solving an engineering problem.	4/5	4.36	4.32	0 of 38 cadets disagreed.
CE1CS	16	I can design a CE component or system to meet a need	4/5	4.43	4.37	1 of 38 cadets disagreed.
1CS	83	The ability to think of creative solutions to real world problems	4/5	4.29	4.00	Lower than the average response for all cadets (4.20).
1CS	63	I can solve basic real-world technical problems.	4/5	4.66	4.26	Higher than average response of all cadets (4.18)
1CS	80	I can confront ambiguous situations.	4/5	4.30	3.83	Slightly lower than average response of all cadets (4.20). Lower than the benchmark

CE492 Embedded Indicators:

Correlation Rating	Performance Score	Performance Standard	Comments
11.8	3.4	3.5	Minimally satisfactory performance, Strong correlation

Correlation between CE492 tasks and program objective is sufficient if correlation rating is > 1.0
Performance scores range from 1 to 5 where 1-2 is unsatisfactory, 2-2.5 is minimally unsatisfactory, 2.5-3.5 minimally satisfactory, 3.5-4.5 good, 4.5-5.0 outstanding

Other Course Embedded Indicators

Course	Embedded Indicator	Performance Score	Performance Standard	Comments
CE300	PS #9 (Beam Design)	94.7%	80%	Strong performance
CE364	EDP	N/A	80%	Will be collected during 2006 -2007 cycle

Figure 7
Each CE program outcome has a series of performance measures based on surveys, direct performance measures, course assessments, and instructor ratings that are used to assess the degree to which an outcome is being attained

Based on this performance by the CE majors in the class of 2006, the CE program director assessed the attainment of this outcome as 4 on a rating scale of 1 to 5 (Fig. 8). The same process was conducted for the all of the 13 outcomes in existence at the time and the results are shown in Fig. 8, along with the ratings over the past five years. It might appear that performance has declined slightly, but in reality, the system has evolved and the standards for attaining a rating of 5 have increased.

Assessment of Outcomes: Summary

#	Civil Engineering Program Outcome	01	02	03	05	06
1	Apply engineering thought process	5	4	4	4	4
2	Demonstrate creativity	5	4	4	4	4
3	Proficiency in structural, const., hydro, and geotech disciplines	4	4-	4-	4-	4-
4	Proficiency in mathematics, calculus-based physics, and chemistry	4	4	4	4-	4
5	Design & conduct experiments, analyze and interpret data	5	4	4	4	4
6	Function on multi-disciplinary teams	5	4	4+	4+	4+
7	Appreciate roles of civil engineers and issues in professional practice	5	5	5	4+	4+
8	Use modern engineering tools to solve problems	4	5	5	4+	4
9	Write effectively	5	4	4	4	4
10	Speak effectively	5	4	4	4+	4+
11	Demonstrate knowledge of contemporary issues	4	4-	4-	4	4
12	Understand engineering solutions in a global and societal context	4	4	4-	4	4
13	Motivation to pursue continued intellectual and professional growth	4	4	4-	4+	4+

Figure 8
The CE program director assigns a rating to each program outcome on an annual basis that evaluates the degree to which each outcome is being attained

Embedded Indicators

Programs are encouraged to develop assessment systems that are sustainable over time and avoid creating data collection systems solely for the purpose of accreditation. Embedded indicators are direct measures of student performance based on assignments already in the curriculum. They can be questions from an exam, a specific essay, a design problem, a group project, or even an entire final exam. The objective is to identify areas that are already being assessed that correlate directly to a specific program outcome. The score on that assignment becomes a direct measure data point for assessment. The embedded indicator should not be taken as proof that an outcome is being attained. There may be many other opportunities in the curriculum to attain the outcome. Rather it is a snapshot in time and is useful only as a single indicator.

The most recent addition to the USMA CE program has been the identification of embedded indicators for every outcome. For outcome 3 that requires graduates to solve problems in the geotechnical sub discipline of civil engineering, the final examination score in CE371, Soil Mechanics and Foundation Design is a relevant embedded indicator. In CE400, Professional Practice of Civil Engineering, the students are required to write ten journal entries. One of the journal entry topics specifically addresses the roles and responsibilities of the civil engineer professional. The score on that essay becomes a direct measure for attainment of outcome 7. Students are required to use a variety of software packages. The AutoCAD problem set in CE390, the Site

Civil course, is an embedded indicator for outcome 8, using modern engineering tools to solve engineering problems. Fig. 7 shows that the two embedded indicators for outcome 1 were problem set #9 from the CE300 (Engineering Mechanics and Design) which involved the design of a simple beam and the engineering design problem from CE364 (Mechanics of Materials) which required the design of a series of roof t-beams that require a load analysis and design based on shear, moment and deflection. Data for these indicators will be collected through course assessment process.

The capstone course CE492, Design of Structural Systems, is unique because it is a culminating design experience and incorporates many of the CE program outcomes. A special tool was designed to capture the student performance and how it relates to the various program outcomes. As described in Meyer et.al., 2005, Meyer et.al., 2006, and Welch and Estes, 2006b, the capstone design is graded where a fixed number of points are allocated to over 50 different areas that include site plan, assumptions, load calculations, social implications, floor plans, architectural layout, cost estimates, construction schedule, quality of presentation, etc. A correlation matrix is created that quantifies the relative contribution of each graded part to the program outcomes. After the tool is developed, the instructor simply enters the scores on each item for each design group and the results are shown in Fig. 9.

CE492 Embedded Indicators: 2006

Measure of Correlation

3	Strong Correlation
2	Good Correlation
1	Acceptable Correlation
<1	Poor Correlation

Measure of Assessment

90%	Excellent
80%	Good
70%	Marginally Satisfactory
<70%	Unsatisfactory

Assessment of Standards

90%	to	100%	Excellent (5)
80%	to	90%	Good (4)
70%	to	80%	Marginally Satisfactory (3)
65%	to	70%	Marginally Unsatisfactory (2)
0%	to	65%	Unsatisfactory (1)

CE492 Assessment of the CE Division Program Outcomes

Team #	EDP Grade	Assessment: 5 (High) - 1 (Low)																	Group Average
1	91.3%	4	4	4	4	4	4	4	3	5	5	5	5	4	4	4			4.3
2	95.8%	4	4	4	4	4	4	4	4	4	5	5	4	4	4	5	4	5	4.2
3	94.7%	4	4	4	4	3	4	4	4	4	5	4	4	4	4	4	4	4	4.0
4	74.9%	3	2	3	3	3	3	3	2	5	5	4	4	4	3	4	4	4	3.5
5	73.7%	2	2	2	1	3	2	3	1	4	4	4	4	3	1	3	2	1	2.5
6	88.3%	4	4	3	3	3	4	4	3	5	5	5	5	4	3	4	4	4	3.9
7	89.4%	4	4	4	4	4	4	4	4	5	4	5	4	4	4	4	5	5	4.2
8	79.2%	3	4	3	2	3	4	4	2	4	4	4	4	4	3	5	5	5	3.7
9	91.1%	4	5	4	4	5	5	4	4	5	5	5	5	5	4	5	5	5	4.6
10	71.1%	2	3	1	2	2	2	3	1	4	4	4	4	3	1	3	1	1	2.4
11	89.1%	4	5	4	4	4	4	4	4	5	5	5	5	5	4	4	3	3	4.2
12	88.9%	3	2	2	1	2	3	3	3	4	4	1	3	3	1	4	4	4	2.8

Criteria Average	3.4	3.6	3.2	3.0	3.3	3.6	3.7	2.9	4.5	4.6	4.3	4.3	3.9	3.0	4.1	3.8	3.8
Measure of Correlation (>1.0 is Acceptable)	11.8	4.0	9.1	2.1	0.6	6.7	3.4	3.0	5.2	2.4	1.6	1.9	5.2	10.2	6.5	3.2	2.3

Assessment Average	3.7
Correlation Average	4.7
EDP Average Grade	85.6%

O5: Design and conduct experiments
O8: Can use modern engineering tools to solve problems
O9: Can write effectively

Figure 9
The CE program has developed an embedded indicator tool that provides a direct measure of student performance with respect to each CE program outcome on the annual capstone design project

Based on the average of each group's performance, two scores emerge for each program outcome. The first score (criteria average) reflects student performance on those tasks in the capstone design that relate to a specific outcome and the second score (measure of correlation) records the extent to which the outcome is covered in the culminating design experience. For those outcomes where the correlation score is low, the outcome should be attained through other courses in the curriculum. Figs. 7 and 9 show that for outcome 1 (design civil engineering components and systems), the correlation between the capstone design and this outcome is the highest (11.8) of any program outcome but the student performance (3.4) is satisfactory but slightly below the desired standard of 3.5. These embedded indicator tools for individual courses and the capstone design provide relevant direct measure data points for outcome assessment.

Course Assessments

As shown in Fig. 3, a major component of the fast loop assessment is the course assessment process (ABET, 2006; Estes, 2004) where every course in the CE program is examined once a year. The formal assessment takes place in a one hour meeting attended by the CE program director, group directors, course directors and interested faculty members. Prior to the meeting, the course director prepares an assessment report in a prescribed format that is distributed to the attendees in advance of the meeting. The assessment consists of three parts. The first is the course description which consists of the verbatim course description from the university course catalogue, current and projected enrollment numbers, course objectives, current textbook, outline of course lessons with their respective contribution to course objectives, outline of laboratory experiences, summary of graded requirements, grading policy, and a reporting of group work, computer usage, active learning, curriculum integration, facilities assessment, and embedded indicators for that course. The second part is course assessment. Both the course director and the students (using the end of course web-based survey) rate the degree to which the course objectives were achieved on a 1 (unsatisfactory) to 5 (excellent) scale. The student ratings on the quality of instruction are included in graphical form, along with a summary of the narrative comments from students. The data are compared to previous years for the same course. Student performance is recorded and compared with the past in terms of incoming grade point average of students, grades in the course, and results on the final examination. Results of time surveys which record the amount of time students are spending on the course are included. Finally, the course director rates the contribution of the course to the program outcomes. Fig. 6 was based on a roll-up of this assessment from the individual courses.

The final part of the course assessment is recommended changes. The course assessment process is the official venue for making changes to courses. This allows faculty members who teach prerequisite and follow-on courses to provide input into course changes that might affect them. The course director addresses whether the previous year's changes were effective and then makes recommendations for new changes based on the findings in part 2. Decisions are made at the course assessment meeting and are recorded on a memo cover sheet. The entire course assessment document is then placed in the course notebook.

The course assessment process takes considerable time and effort, especially the first time through. Because the reporting is done in the same systematic way each year, it becomes much less onerous to update a previous report. The standardized format makes it easy to consolidate data for the fast loop program assessment. When several courses are interrelated, their respective course assessment meetings are often conducted together to facilitate cross talk and coordination.

Ongoing Efforts

While the assessment system within the CE program has evolved and improved, there are a number of courses taught outside the department that contribute to program outcomes, but are not being effectively examined. Because all accredited engineering programs at the Academy are in the same situation, the USMA level ABET committee is addressing this challenge. The USMA curriculum is broad based where every student, regardless of major, is required to take calculus, physics, chemistry, psychology, English, foreign language, history, international relations, geography, information technology, philosophy, law, leadership, and economics. The ABET committee is attempting to meet with outside departments to specifically quantify the contributions that these courses make toward attainment of the ABET 3(a-k) outcomes. Every accredited engineering program can then use those data in its own program assessment.

There is an additional ongoing effort to create rubrics wherever possible. A rubric helps provide a word picture description to the Likert values used on surveys to help gain greater consistency and understanding from various raters. Rubrics are currently being developed for the course-outcome matrix, team projects, the overall assessment of outcomes and outside evaluator sheets.

Staying Current and Communicating

As EC2000, the outcomes-based continuous improvement approach to accreditation continues to evolve, the standards for assessment keep getting higher and the requirements change somewhat from year to year. Every program needs to stay involved in order to keep up with the latest developments. The USMA CE program faculty attempt to stay current by attending and making presentations at such forums as the ABET national meeting (Welch and Estes, 2005a), the ASEE Annual Convention (Estes, 2003; Estes et.al., 2003; Welch and Estes, 2003; Estes and Klosky, 2002), the ASCE National Conference (Estes, 2004a; Welch and Estes, 2005b; Estes, 2004b) and the Rose Hulman Best Practices Workshop (Estes and Welch, 2004; Welch and Estes, 2006b). Currently, three out of the twenty faculty members in the USMA CE program are ABET evaluators, and one member conducts ABET evaluator training for other institutions. There is no better way to stay current and see what other programs are doing than to serve as an evaluator of other programs. While one or two leaders will orchestrate and lead the assessment effort within the CE program, every attempt is made to involve all faculty members in the process. This divides the work load, educates the faculty members on the process, and helps facilitate support and buy-in from the entire faculty.

Conclusions

The United States Military Academy CE program has developed and successfully sustained a program assessment process that effectively encompasses slow loop and fast loop evaluations that are each designed to perform very different functions. The system is based on the university assessment system and has eight years of documented data and analysis. The annual assessment culminates in a formal briefing that addresses previous recommendations, reports on the results of data collected over the year, and recommends changes to the program based on that analysis. The foundation of the process is annual course assessments in every course that allow data to be consolidated in a standardized format. Outcomes and objectives are assessed based on a comparison of student performance to prescribed performance measures, and an annual rating is given for each individual outcome and objective. Additional input is provided by an external advisory board whose composition and role have evolved over the past six years. The CE program has remained current on changes in assessment requirements through participation on professional society committees, presenting papers at workshops and conferences, and serving as accreditation evaluators for other programs. There are still improvement and adjustments to be made. The 2006 CE program assessment listed 25 specific recommended changes that range from major to extremely minor. The USMA CE assessment program will continue to evolve as a process of continuous improvement is sustained.

References

ABET. (2005). *Guidelines to institutions, team chairs and program evaluators on interpreting and meeting the standards set forth in criterion 3 of the Engineering Accreditation Criteria*. ABET Criterion 3 White Paper, Retrieved May 2004 from http://www.abet.org/pev.shtml accessed 19 September 2005.

ABET. (2006). *Criteria for accrediting engineering programs effective for evaluations during the 2006-2007 accreditation cycle*. Baltimore, MD: Engineering Accreditation Commission, Accreditation Board for Engineering and Technology.

ASCE. (2004a, October) *Civil engineering body of knowledge for the 21st Century*. Reston, VA: Committee on Academic Prerequisites for Professional Practice, American Society of Civil Engineers.

ASCE. (2004b, October). *Interpretation of the ABET engineering criteria for civil and similarly named programs*. Commentary, Version 1.1. Reston, VA: Committee on Curricula and Accreditation, American Society of Civil Engineers.

ASCE. (2005a, September). *Levels of achievement applicable to the body of knowledge required for entry into the practice of civil engineering at the professional level*. Reston, VA: Draft Report of the Levels of Achievement Subcommittee to the ASCE Committee on Academic Prerequisites for Professional Practice, American Society of Civil Engineers.

ASCE. (2005b, October). *Program criteria for civil and similarly named engineering programs–ASCE's proposed changes to the criteria for accrediting engineering programs, effective for evaluations during the 2008-2009 accreditation cycle*. Reston, VA: Work Product of the Accreditation Committee of the Committee on Academic Prerequisites for Professional Practice, American Society of Civil Engineers.

ASCE. (2006, November). *Development of civil engineering curricula supporting the body of knowledge for professional practice*. Reston, VA: Draft Report of Curriculum Committee of the Committee on Academic Prerequisites for Professional Practice, American Society of Civil Engineers.

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives*. New York: Longman.

Estes, A.C. (2003, June). *Trials of a beginning ABET evaluator: An unauthorized tell-all biography*. Proceedings of the American Association of Engineering Education (ASEE) Annual Conference, Nashville, TN.

Estes, A.C. (2004a, October). *Course portfolios: A tool for continuity and assessment*. Paper presented at the 2004 ASCE Civil Engineering Conference and Exposition, Baltimore, MD.

Estes, A.C. (2004b, October). *The role of student chapters in improving civil engineering programs*. Paper presented at the ASCE Civil Engineering Conference and Exposition, Baltimore, MD.

Estes, A.C. & Klosky, J.L. (2002, June). *Designing experiments in a civil engineering curriculum*. Proceedings of the ASEE Annual Conference, Montreal, Quebec.

Estes, A.C., LaChance, E. & Evans, M.D. (2003, June). *The role of student chapters in improving CE programs*. Proceedings of the ASEE Annual Conference, Nashville, TN.

Estes, A.C. & Ressler, S.J. (2003, June). *Teaching assessment: How do you do it?* Proceedings of the ASEE Annual Conference, Nashville, TN.

Estes, A.C. & Welch, R.W. (2004). *Program assessment: Conquering the slow loop*. Baltimore, MD: Best Assessment Processes Proceedings CD (2004), ABET, Inc.

Estes, A.C., Welch, R.W. & Ressler, S.J. (2005a). The ExCEEd teaching model. *Journal of Professional Issues in Engineering Education and Practice*, 131: 1-5.

Estes, A.C , Welch, R.W. & Meyer, K.F. (2005b, June). *Will ten pounds fit into a five pound bag?* Proceedings of the ASEE Annual Conference, Portland, OR.

Estes, A.C., Welch, R.W. & Ressler, S.J. (2006). The assessment of teaching. *Journal of Professional Issues in Engineering Education and Practice*, 132: 218-222.

GroupSystems (2006). GroupSystems II software, Information retrieved November 14, 2006 http://www.groupsystems.com/ .

Meyer, K.F., Estes, A.C., Welch, R.W. & Winget, D. (2006, June). *Program assessment the easy way: Using embedded indicators to assess program outcomes*. Proceedings from the ASEE Annual Conference, Chicago, IL.

Meyer, K.F., Morris, M., Estes, A.C. & Ressler, S.J. (2005, June). *How to kill two birds with one stone – Assigning grades and assessing program goals at the same time*. Proceedings from the ASEE Annual Conference, Portland, OR.

Rogers, G. (2003). *Do grades make the grade for program assessment? Assessment Tips with Gloria Rogers*. Retrieved October 20, 2007 from http://www.abet.org/Linked%20Documents-UPDATE/Assessment/Assessment%20Tips4.pdf

Rogers, G. (2004). *How are we doing? Assessment Tips with Gloria Rogers*. Retrieved October 20, 2007 from http://www.abet.org/Linked%20Documents-UPDATE/Assessment/Assessment%20Tips4.pdf

Rogers, G. (2006). *Direct and indirect assessments: What are they good for?" Community Matters, Assessment 101, Assessment Tips with Gloria Rogers*. Retrieved October 20, 2007 from http://www.abet.org/Linked%20Documents-UPDATE/Newsletters/06-08-CM.pdf

Welch, R.W. & Estes, A.C. (2006a, June). *Assessing current programs against the new BOK*. Proceedings of the ASEE Annual Conference, Chicago, IL.

Welch, R.W. & Estes, A.C. (2006b). *Systematic program assessment: Using embedded indicators and closing the feedback loop*. Baltimore, MD: Best Assessment Processes Proceedings CD (2006), ABET Inc.

Welch, R.W. & Estes, A.C. (2005a, October). *Assessment of squishier outcomes: Open ended problem solving through client-based projects*. Paper presented at the ABET Annual Meeting, San Diego, CA.

Welch, R.W. & Estes, A.C. (2005b, April). *Project based independent study capstone course*. Proceedings of the 2005 Structures Congress and Exposition, ASCE, New York, NY.

Welch, R.W. & Estes, A.C. (2003, June). *Client-based project for any senior – A mark of excellence for any program*. Proceedings of the ASEE Annual Conference, Nashville, TN.

U.S. Army (1993). *FM 5-100, engineer combat operations field manual*. Washington D.C.: autbor.

U.S. Military Academy (USMA). (2000). *Educating army leaders for the 21st century*. West Point, NY: author.

CHAPTER 8

EMBEDDED ASSESSMENT: EASING THE FACULTY WORKLOAD

Daina Briedis
Michigan State University

Introduction

The purpose of this paper is to illustrate an approach to program outcomes assessment that follows principles of best practices yet minimizes faculty workload. The paper is directed to faculty in engineering, computing, technology, and applied sciences programs, and the message is three-fold and simple:

1. Assessment, whether it is done to satisfy ABET criteria for accreditation or as a normal part of education, is becoming pervasive and demonstrably valuable for student learning and program improvement.
2. Demonstration of effective assessment processes is necessary for accreditation, and the standard for compliance with the new criteria has been raised subsequent to first- and second-round accreditation visits. More specifically, it is anticipated that *direct assessment* of outcomes will become an expectation in assessment processes.
3. To minimize workload, faculty can use many of the routine modes of evaluation of student performance as opportunities for assessment with the addition of a few simple tools and by developing strategies for efficient assessment cycles. Some up-front work is a necessity.

This message is not new, but to date many faculty are reluctant to accept the inevitable - assessment should and will eventually become as normal a part of teaching responsibilities as classroom pedagogy and grading student work. The purpose of this paper is to illustrate that this can be done relatively easily and efficiently with a positive return in improvement of student learning for minimal (but concerted) investment.

There is no single approach to assessment that is specified by ABET. Programs are encouraged to develop processes that are most effective within their unique constraints and academic environments. This paper offers suggestions for a streamlined design of a sustainable assessment process, discusses the importance of *direct* assessment of student work, the need for serious preparation, and proposes tools that may be used to accomplish these tasks. Several examples are provided as to how this can be successfully implemented in small to large engineering, technology, computing, or applied sciences programs.

ABET and Outcomes Assessment

The approaches described in this paper deal with assessment of program *outcomes*—the knowledge, skills and attributes that students should demonstrate by the time of graduation. Therefore, the focus is campus- or curriculum-based assessment. This paper does not address *objectives*, which describe career and professional accomplishments of program alumni. Evaluation of achievement of objectives usually requires different tools than those used for program outcomes, although some overlap does exist.

The 2008-09 criteria of all four ABET commissions specify from nine (computing programs) to eleven (applied sciences, engineering, and engineering technology) program outcomes and require that programs demonstrate that their graduates possess the attributes included in these outcomes (ABET, 2007a; ABET, 2007b; ABET, 2007c; ABET, 2007d). Programs are free to add to these outcomes and to tailor them to their particular program characteristics and curricular thrusts. Any outcomes that are added must be assessed as well. ABET criteria for all commissions (2008-09 accreditation cycle) further require that the student achievement in these outcomes be demonstrated and measured; the various commission criteria use language such as requiring "determination of the extent to which program outcomes are being met" and "the degree to which program outcomes are attained" (ABET, 2007a; ABET, 2007b; ABET, 2007c; ABET, 2007d). Within these phrases is the crux for needing direct assessment.

Direct and Indirect Assessment

Many programs are struggling to find effective measures to bring their outcomes assessment processes into full compliance with the requirements of the "outcomes and assessment" and "program improvement" criteria. This is substantiated by the relatively large number of citations for shortcomings relative to some aspect of these criteria (Briedis, 2005). One cause of this is that, while faculty members are actually very familiar with direct evaluation of student work (we do this each time we give an exam, grade a project, or evaluate a presentation), faculty will understandably default to the least demanding assessment method when it comes to assessing ABET outcomes. In most cases, the choice has been surveys and similar indirect measures. Surveys are easy to use, some are available commercially, and analysis of results can be quick, painless, and executed by user-friendly software. However, indirect measures generally do not provide as much useful information as direct measures do for evaluating achievement in student learning (Rogers, 2006c).

Direct assessment (performance appraisals, examinations, portfolios, simulations) is based on direct examination or observation of student performance. Indirect assessment (surveys, interviews, focus groups) provides opinions or perceptions about student knowledge or skills and typically relies on self-reporting (Rogers, 2006c). While neither assessment method is perfect and both have limitations, direct assessment provides a stronger measure of student learning than indirect assessment; indirect assessment is self-reported, is usually not rated on a calibrated scale, and therefore carries more bias than direct assessment. Most importantly, a vital assessment program should use multiple assessments methods–both direct and indirect. Using multiple methods allows for triangulation and confirmation of evidence of student learning, which is the bottom line not only for ABET but also for faculty truly seeking to improve their programs.

While it is true that, in the past, programs may have passed through accreditation visits with only surveys and faculty opinions as evidence of student achievement in outcomes, indications are that this will no longer be the case. It is safe to say that, in future accreditation visits, programs relying *only* on indirect measures for outcomes assessment will likely be cited as having shortcomings.

Grading versus assessment

As will be mentioned again later in this paper, assessment is not grading, although, conversely, grading can sometimes be used for assessment. Grading of projects, homework, quizzes, and exams typically involves norm-based scoring of a student's overall performance on any one assignment (Nichols, 1995). Projects, homework, quizzes, and exams represent an evaluation of integration of student knowledge and skills usually on a variety of questions or problems. When a student is assigned a grade, the grade represents a norm-based evaluation of this integrated performance, and student performance on a specific outcome (assessment) cannot usually be disaggregated from the grade. For example, at a minimum, student learning in heat transfer requires an integration of problem-solving skills, applied mathematics, computational tools, and design skills. If a student receives an average grade ("C" or 2.5) in a heat transfer course, the assigned grade provides very little specific information about the student performance in any of the outcomes area.

Alternately, grading may be used as assessment if the object being graded is specifically designed to allow students to demonstrate an outcome. This will be discussed and illustrated later in the paper.

Formative and summative assessment

Most faculty are familiar with formative and summative assessment. Formative assessment provides the instructor with ongoing feedback that monitors student progress through the learning process (National Research Council, 2000). Summative assessment measures what students have learned at the end of this process. In this context, assessment of ABET outcomes is summative. This does not rule out formative assessment as a mechanism for outcomes assessment, but good design of the assessment tool and data overload may become issues in using the more frequent and less formal types of formative assessment methods.

New program evaluator training

ABET has received feedback concerning inconsistencies in program evaluations, particularly in regard to evaluation of objectives and assessment of outcomes. ABET's accreditation processes include a thorough network of report comparison, editing, and consistency checks, but ABET will readily acknowledge that a measure of variability still exists among the reasons for which shortcomings are cited. The variability may be attributable to multiple sources, one of which is the on-site judgment of the program evaluator (PEV). The first line of program evaluation takes place through the eyes of the PEV—the program evaluator is the face of ABET on campus.

Program evaluators are volunteers, and each PEV carries a slightly different perspective into a visit. Therefore, ABET has begun its improvement efforts for program evaluation by focusing on program evaluator training. As a good steward of its own improvement processes, ABET has embarked on a new program to promote excellence in its volunteers (PAVE=Partnership to Advance Volunteer Excellence (ABET, Inc.). Key elements envisioned for this program relative to PEVs and accreditation visits are a) the selection and training of PEVs based on required competencies and b) new training that is experiential and consistent across all commissions. Most relevant to this paper is the fact that program evaluators are now being trained to look for direct assessment as part of their evaluation of program compliance with the intent of ABET criteria.

Faculty who have attended assessment workshops or training programs are also becoming savvy in understanding that indirect measures alone are insufficient for effective outcomes assessment; those programs interested in best practices will be intrinsically motivated to use direct methods as a key element of their assessment processes. If programs still plan on seeking ABET accreditation in the future, assessment will be required, and if it must be done, why not gather valuable information? The remainder of this paper addresses ideas on how to run an effective, "minimally intrusive" assessment process without simply going through the motions.

Assessment Program Design

Pet-Armacost and Armacost (2005) list several traits of a good assessment plan. These include faculty ownership of the process, the use of multiple measures and sources consistent with program constraints in time and resources, realistic goals and timetables, and the necessity for review and possible re-design of the assessment program itself.

Faculty are at the heart of several of these traits and are key in sustaining an assessment program. While it is clear that there must be an individual or committee with the responsibility of overseeing assessment efforts within a program, broad faculty participation and understanding of the processes are vital. No one faculty member should be singly responsible for assessment of an outcome. Program outcomes require an integration of learning experiences; therefore, it is the responsibility of the combined program faculty to assess the student learning across the curriculum.

First, faculty are the most qualified to directly assess student performance as they work most closely with students in the classrooms and laboratories. Secondly, these same faculty are the most logical channels for effecting any curricular or pedagogical changes that are deemed necessary for program improvement. Therefore, for an assessment process to even begin to be palatable and sustainable, faculty assessment workload should be minimized. This is as important for faculty who have significant research responsibilities as it is for faculty who carry heavy teaching loads. The design of an efficient assessment process must acknowledge faculty time as a valuable commodity and a significant constraint.

Designing an Assessment Process to Avoid Faculty Overload

Minimizing faculty effort can be accomplished by using the "familiar" and keeping the workload manageable. Strategic decisions must be made by program faculty when determining when and how to assess outcomes. These decisions will depend on resources: the number of faculty in the program, the number of students in the program, the availability of assessment support (learning centers, institutional research programs, etc.), and other factors particular to a program and institution. Local and external experts can contribute to assessment process design.

All faculty members should have a good understanding of which courses prepare students in which outcomes whether the faculty member teaches the course or not. Many programs represent this distribution in an outcomes map, in which the expected contribution to student learning in an outcome area is prioritized for each course in the curriculum (e.g., Figure 1). Other programs have pre-requisite "trees" that clearly display continuity and interdependence among courses in a curriculum and allow tracking of learning in the outcome areas (Froyd, Layne, Yurttas, & Ford, 2006). Of course, not all courses will address all outcomes; however, when integrated across an entire curriculum, all outcomes must be covered–preferably in multiple courses. This distribution offers variety not only for student learning opportunities towards the outcomes, but also offers options for allocating the assessment load among courses and faculty.

Options in assessment cycles

Not all outcomes need to be assessed each time a course is taught. Some programs iterate regularly through their list of outcomes and assess a given number each year; for example, the shaded cells in Table 1 show one possible cycle of outcomes assessment.

Other programs have prioritized their outcomes and assess the higher priority outcomes more frequently. If a program is striving to improve student performance in a particular outcome, the frequency of assessment may be increased in order to check on the progress of improvement efforts. This is a case in which formative assessment

Table 1
Six-Year Outcomes Assessment Cycle

Year/Outcome	1	2	3	4	5	6
A						
B						
C						
D						
E						
F						
G						
H						
I						
J						
K						

	Program Courses													
Program Outcome	**100**	**150**	**200**	**250**	**300**	**325**	**375**	**401**	**410**	**425**	**450**	**475**	**485**	**490**
Appl. of math, science, eng knowledge														
Experiments and data														
Design														
Teams														
Problem-solving														
Prof. and ethic														
Communication														
Global/societal context														
Lifelong learning														
Contemporary issues														
Engineering tools														

may be appropriate not only as a means of evaluating student progress, but also of evaluating the effectiveness of instruction.

For programs with a small number of faculty members, the assessment may also be rotated from course to course so as to distribute the load evenly among the faculty, even providing faculty an occasional "assessment-free" semester. Regardless of the model, assessment should be done in the framework of the academic calendar and within the ABET accreditation cycle. It is mandatory, at a minimum, to have assessed all ABET- and program-designated outcomes over the six years between accreditation visits.

Sampling

Even despite the diminished burden that a rotating assessment cycle provides, faculty of large programs may still be initially overwhelmed by the thought of directly assessing the work of all of their classroom students. Not only is it efficient to rotate through a cycle of outcomes as described above, but it is also acceptable to assess the work of a sample of students (Briedis, 2002; Nichols, 1995; Rogers, 2006a). Some programs use cohorts or samples of a cohort to conduct longitudinal and/or comparative studies. In other programs (Briedis, 2002), a different sample of students is taken for each different assessment conducted in a course. As long as a class is large enough so that a sample will be statistically meaningful across important student traits (GPA, gender, ethnicity, co-op and non co-op, part-time or full-time, other parameters), assessing a smaller number of students can significantly diminish faculty workload. Programs with small numbers of students may need to assess the entire population in order to obtain meaningful results.[1]

These approaches are justified by remembering that assessment for ABET purposes is holistic—to conduct program, not individual student assessment. Classroom assessment can be integrated to provide program assessment. Therefore, the proposed models are appropriate for achieving the goal of improving the quality of programs based on reliable data. The "minimalist" strategies encourage faculty involvement, but also avoid data overload for the assessment coordinator who must collect and analyze the data, and present results to the faculty, administrators, and advisory boards. The following pages will further address the collection of meaningful data and suggest tools by which this can be realized in familiar territory.

---

[1] The concept of assessing the work of a sample of students is not to be confused with the requirement of all ABET commissions that samples of student work be part of the display materials provided during the campus visit. The ABET Accreditation Policy and Procedure Manual (ABET, 2007e) requires that representative samples of student work be available during the on-site accreditation visit. The purpose of these materials, which include course outlines, syllabi, textbooks, and **graded** student work, is to provide the team with information needed to make a qualitative evaluation of the program. In addition, the guidelines for the self-study questionnaires for each commission request samples of student work demonstrating performance in outcome areas. How the requirement for display of **graded** student work is to be reconciled with the request for demonstration of student outcomes performance is best left to consultation with the team chair and program evaluator prior to the institutional visit.

Tools and Opportunities—Preparing for Embedded Assessment

Many occasions exist in the classroom for the direct assessment of program outcomes in student work. This so-called "embedded" assessment allows a faculty member to double up on a student evaluation that is already in use in a classroom or elsewhere in the curriculum and exploit it as an assessment opportunity.

When considering assessment of coursework, it is important to understand that sment is NOT equivalent to grading. Student work must be viewed in a different light to enable its use for outcomes assessment. As described by Nichols (1995) and Rogers (2006a), program assessment is focused on providing demonstration of student knowledge or skill directly linked to a specific program outcome. Grades depend on a faculty member's expectations for a course, course content, and the particular grading policy. As mentioned earlier, grades are norm-based (Nichols, 1995) and reflect a student's standing relative to others within a class or on a test. Assessment is based on a student's absolute performance on a specific measurable aspect of an outcome.

Performance metrics

Outcomes must be measurable. This is required by ABET criteria and is a means by which a program can benchmark improvement in student learning. Constructing a set of descriptors, or performance criteria, to make outcomes *measurable* is one of the most important steps in implementing and sustaining an effective assessment process (Rogers, 2006b) and can also be one of the more difficult and time-consuming tasks related to assessment. However, when designed well, a good set of measures provides for consistency and reliability of data throughout the outcomes assessment process.

The general statements of outcomes provided in the ABET criteria are not descriptive enough for productive assessment. Program faculty should understand for themselves, in the context of their own programs, what each outcome will "look like" when students are demonstrating its traits (performance criteria). Furthermore, rubrics, or a measurement scale, should be developed to describe several levels of achievement for each performance criterion. Diagrams of this hierarchy for two outcomes, their performance criteria, and rubrics are shown in Figures 2 and 3.

Each outcome should be parsed into several performance criteria that describe observable features of an outcome in familiar terminology (Rogers, 2006b; Dodge, 2001). A single outcome may have numerous performance criteria, but it is recommended (Rogers, 2006b) that the focus be narrowed to three to five performance criteria that are relevant to instruction and student learning in the particular program. Four performance criteria will be used here.

At the next hierarchal level, rubrics provide the rating scale that describes levels of possible student performance for each performance criterion. A numerical score is assigned to each level, e.g., 4 = exemplary; 3=meets expectations; 2= below expectations; 1=unacceptable; or 5=Superior proficiency; 4=Good proficiency; 3=Proficient; 2=Progressing to proficiency; 1=Not Proficient. Other rating scales may be used, and the user should decide how many levels of achievement are needed. With rubrics, the faculty member can rate student performance in each of the four performance criteria, with a level (1 through 4 or 1 through 5) assigned for each one.

Figures 2 and 3 show a possible breakdown for Engineering Criteria Outcomes (a) and (i). These figures are intended only to demonstrate the hierarchy of outcome to performance criteria to rubrics. Depending on how a program faculty interprets these outcomes, the statements of performance criteria and rubrics may be completely different than those given in the examples. Similar hierarchies should be developed for each ABET- and program-designated outcome. Ideally, all faculty members and instructors should be involved in this development exercise, and should be in general agreement on the content, as the performance criteria and rubrics should serve across the curriculum. Additional helpful information on developing performance criteria and rubrics is available elsewhere (Dodge, 2001).

Threshold levels for acceptable student performance also need to be set. For the example given in Figure 2 for life-long learning, a faculty may agree that a level of 2.5 or above for at least 75% of the students on each performance criterion could be considered acceptable. The thresholds for all outcomes need not be the same. Outcomes given high priority in a program may have more rigorous standards than lower priority outcomes. It is to be expected that some iteration on thresholds will be required as faculty become familiar with performance criteria, rubrics, and their correlation to student performance.

Embedding and In Situ Assessment

Once careful preparation is completed, the time-saving part of embedded assessment begins. Faculty members may use the performance criteria and rubrics in existing classroom practices to directly assess student performance. Since the instructor is grading the assignment anyway, a clear advantage of this method is that the extra effort required by instructors is incremental to the effort required to grade the assignments.

A faculty member can simply use the documentation for the relevant outcome, its associated performance criteria, and rubrics to assess any sample of student work that demonstrates that outcome. The statements for performance criteria and their rubrics may even be incorporated into a grading sheet to allow these specific traits to be evaluated during the normal grading process. However, as discussed earlier, it is very important to understand that grading and assessment are not equivalent. The rubric scores may be part of the overall grade, but they should be extracted separately from the grade to provide a measurement of the student performance in the outcome under consideration. The following example may be useful to illustrate this point.

For example, let us suppose that a mid-term exam has been administered in which the second problem has been designed to test the outcome on the ability to apply mathematics in problem-solving. In this particular case, the problem requires

- development of a differential equation to describe a physical system
- the linearization of the non-linear differential equation, and
- transformation of the differential equation into the Laplace domain.

Either during or after the grading of the entire exam, the faculty member selects a representative sample of students; for small classes, it may be necessary to assess the work of all of the students. The pertinent exam problem is assessed using the levels described by the rubrics (Figure 2), and the individual student scores for each relevant performance criterion are recorded. Alternatively, the faculty member may have already included the performance criteria on a grading sheet. In this case, the scores on the specific performance criteria are simply extracted from the grading sheet and recorded separately. Student names are not associated with assessment results. A similar process

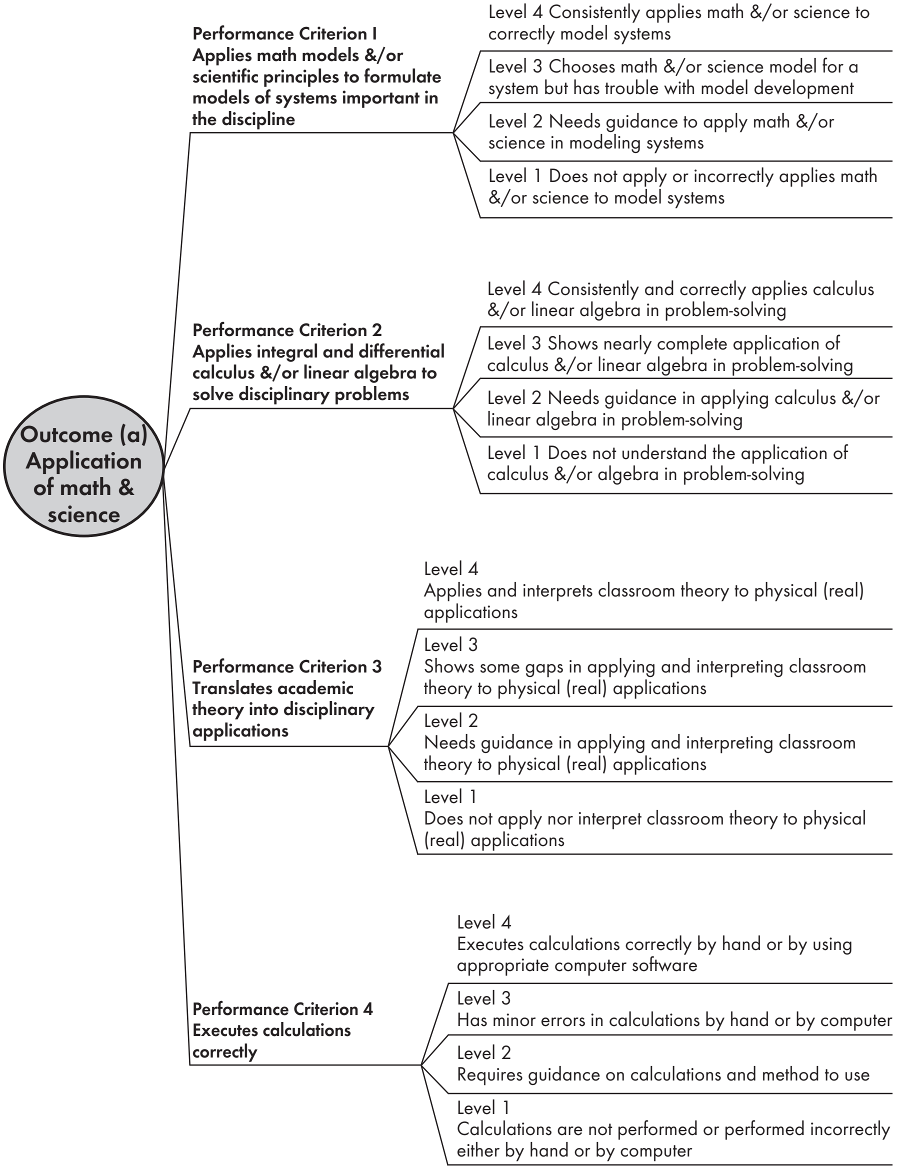

Figure 2
Example of an outcome (a), its performance criteria, and rubrics

is conducted for selected student work by all faculty members assigned according to the program's assessment cycle so that outcomes results are obtained across the curriculum.

Good practice suggests that assessment be repeated for different types of student work during the course of a semester. This avoids the "snapshot" phenomenon in which

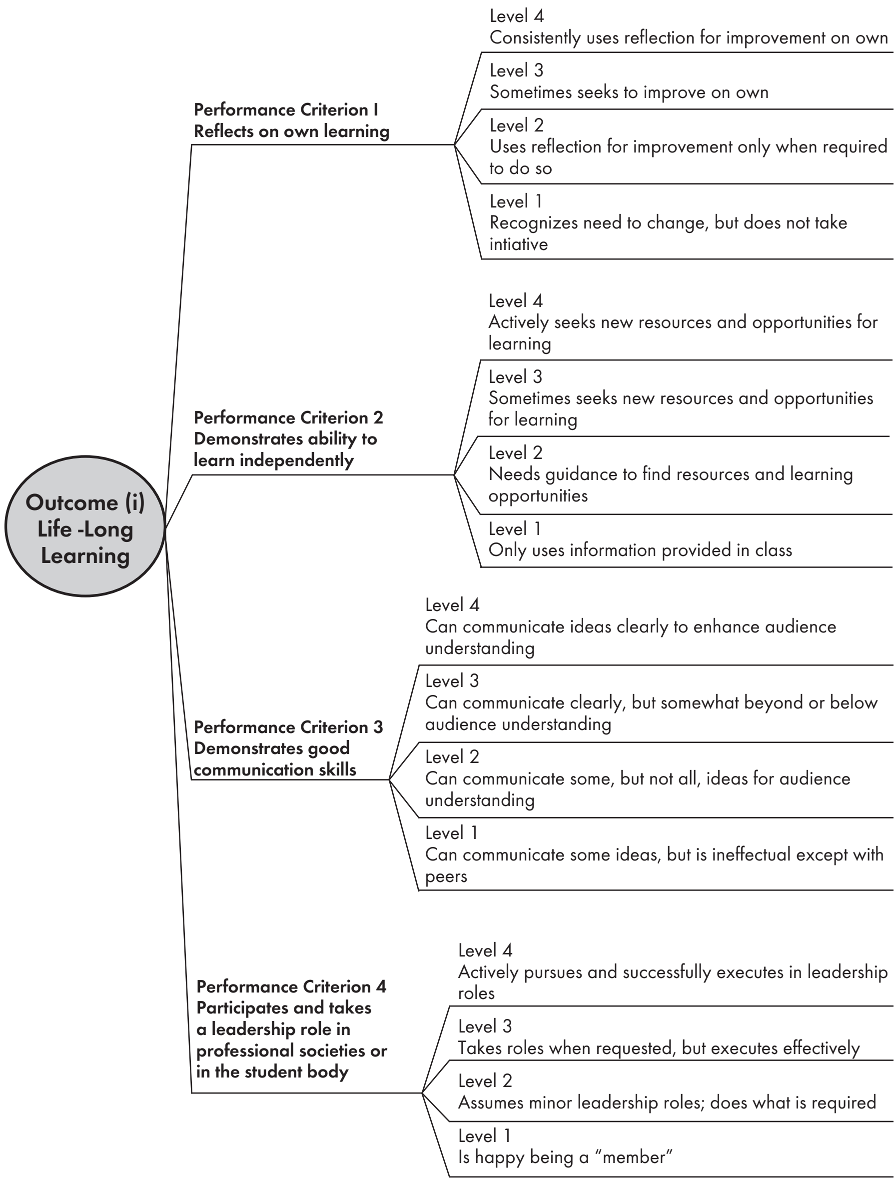

Figure 3
Example of an outcome (i), its performance criteria, and rubrics

a single assessment may not be an accurate indicator of student skills. This approach allows not only for multiple direct assessments of an outcome, but may also double as formative course assessment (Angelo & Cross, 1993) for improvements during the course offering.

The assessment coordinator, or other responsible staff or faculty member, collects the assessment data for all scheduled outcomes and collates the results most preferably by performance criteria. The performance criteria data may be averaged or viewed as a function of course, student class, or any other parameter that is important to student learning and its improvement. It is best not to average the data across all performance criteria of an outcome, since specific information on the root cause of any weaknesses may be lost. For any quality assessment process, these direct assessment results should be analyzed together with other forms of assessment that can validate the need for change.

Opportunities for direct embedded assessment

The example above cited the use of a mid-term examination for assessment of mathematics skills. This approach may be applied to targeted assignments, homework, examinations, projects, and presentations throughout the curriculum that may serve as key determinants of student achievement in specific program outcomes.

Capstone experiences are ideal opportunities for multiple assessments since, by definition, students are integrating previous learning to solve usually design-oriented problems. Furthermore, since these experiences are frequently done in teams and usually involve written reports and oral presentations, the professional outcomes (teamwork, communication, global/societal context, life-long learning, contemporary issues) can also be assessed. Laboratory experiences may also fall into this category, especially if the laboratory class requires significant application of previous learning.

Seminar courses may provide additional unique opportunities to assess the professional skills, which are not always explicitly covered in traditional technical courses. The author's program includes a junior seminar course in which many of the professional skills are considered. One assignment is an essay on an ethics case study. This allows opportunity to assess not only understanding of professional and ethical responsibility, but also serves as a formative assessment of student abilities in writing. Similarly, other assignments are designed to allow demonstration of student knowledge and understanding of contemporary issues and the societal impact of the discipline.

Capstone examinations are required by some engineering and technology programs. These include exams such as the Fundamentals of Engineering (FE) exam, the Fundamentals of Surveying exam, or program-generated comprehensive examinations. While the outcomes of the FE exam currently only provide information on broad topic areas and do not provide information down to the performance criteria level, the topic-specific information provided by most state boards can provide some data on achievement of subject mastery. These data may serve as supporting evidence for direct assessment on campus. At some institutions, the FE exam is not required; students self-select, and the sample becomes biased. Differences of opinion abound as to whether the FE or FS exam provides sufficient information for assessment of program outcomes (Watson, 1998; Nirmalakhandan, Daniel, & White, 2004).

Although rare, some programs write comprehensive examinations that are graduation requirements for all students (Maffia, 2004). This provides a unique opportunity for the exam writers to focus problems, or parts of problems, on particular outcomes. While not all outcomes can be assessed with this method, it is convenient for evaluating integrated student learning.

Another excellent assessment opportunity is afforded by student portfolios. Portfolios are a collection of student work in a course or through a curriculum that demonstrates the skills students have learned. Students use portfolios to document their best work, their longitudinal progress, and for reflection on their learning. By choosing a sample of students from which to collect portfolios, achievement of outcomes can be demonstrated and assessed. Faculty may even require portfolio assignments that address specific outcomes. However, this may be a more difficult endeavor for faculty and students than other means of embedded assessment. Portfolio assessment can go beyond the normal grading routine because portfolios are usually assessed when complete and often by those not involved in the course instruction (RosE Portfolio System). This requires a time commitment that most faculty members are not willing or able to make.

Using technology

As the Internet and various personal electronic devices continue to become a greater part of university instruction, the convenience that they provide can also be valuable in streamlining assessment processes. Several programs have developed web-based assessment forms which allow electronic entry of assessment data and its subsequent storage in a database (Briedis, 2005). Web-based systems will also send automated e-mail reminders to faculty responsible for assessment during any given semester. Application of such technology adds convenience and lightens the workload on those responsible for assessment analysis and implementation of change.

Does Assessment Work?

Given the front-end work that is necessary for effective embedded assessment, it is rational for faculty to question whether it is all worth it. Beyond satisfying ABET requirements, does assessment really help the programs? The answer to this question depends a great deal on how open faculty members are to change. Entrenchment will not breed improvement. However, assessment success stories have been told in programs where faculty members are willing to accept assessment results and undertake improvement in new directions.

In the author's program, a weakness was identified in student performance in the application of mathematics to engineering problem solving. The weakness appeared to worsen as students progressed through the curriculum. These results were validated by student self-assessment and by cooperative education employer feedback that indicated that student employees were lacking in certain computing skills. Focus groups conducted with groups of students indicated that the source of this weakness might be two-fold:

1. because of the large gap between the offering of our own applied mathematics courses and the basic math courses taken by first- and second-year students, and
2. by the lack of incorporation of some of the entry-level computing skills in our own disciplinary courses (specifically MATLAB).

The first problem was solved simply by moving an applied mathematics course to an earlier point in our program. The attempted solution to the second problem has developed into a fruitful collaboration between the computing sciences program and several engineering programs. A strong link has been established between the introductory course and its importance in upper level disciplinary courses, not only for students, but for faculty as well. Many faculty members have increased implementation of MATLAB in their courses due to a refresher MATLAB short course that was offered by the computing sciences faculty and subsequent technical support. The collaboration has also resulted in the inclusion of more disciplinary examples and specifically requested course content in the introductory computing course. While the study of the outcomes of this improvement process is still underway, preliminary assessment results of student performance in and attitude towards the application of mathematics in chemical engineering has improved significantly (Briedis, Miller, Ofoli, Sticklen, & Urban-Lurain, 2006; Briedis, Miller, Ofoli, Sticklen, Walton, & Urban-Lurain, 2007).

Conclusion

For the faculty of programs desiring continued ABET accreditation, assessment will become of way of life. It is likely that the standards for compliance with outcomes and assessment criteria will become more rigorous in the near future. In this context, both program evaluators and knowledgeable faculty understand that best assessment practices require direct assessment of student work. With concerted investment in thoughtful development of assessment cycles and performance metrics, the "burden" of assessment can be minimized for program faculty so that assessment processes may be fruitful and valuable in enhancing student learning.

References

ABET. (2007a). *Criteria for accrediting applied sciences programs*. Retrieved August 27, 2007 from www.abet.org.

ABET. (2007b). *Criteria for accrediting computing systems programs*. Retrieved August 27, 2007 from www.abet.org.

ABET. (2007c). *Criteria for accrediting engineering programs*. Retrieved August 27, 2007 from www.abet.org.

ABET. (2007d). *Criteria for accrediting engineering technology programs*. Retrieved August 27, 2007 from www.abet.org.

ABET. (2007e). *Accreditation policy and procedure manual*. Retrieved August 27, 2007 from ww.abet.org.

Angelo, T.A. & Cross, K.P. (1993). *Classroom assessment techniques*. San Francisco: Jossey-Bass Publishers.

Briedis D. (2002). Developing effective assessment of student professional outcomes. *International Journal of Engineering Education*, *18*(2), 208-216.

Briedis, D. (2005, June). *The raised bar: Second round accreditation experiences*. Proceedings of the ASEE Annual Conference, Portland, OR.

Briedis, D., D. Miller, R. Ofoli, J., Sticklen,J., & Urban-Lurain, M. (2006, June). *Course assessment for evolutionary curricular reform: Infusing MATLAB vertically through the curriculum*. Proceedings of the ASEE Annual Conference, Chicago, IL.

Briedis, D., Miller, D., Ofoli, R., Sticklen, J., Walton, S.P., & Urban-Lurain, M. (2007, June). *Course assessment for curricular reform II*. Proceedings of the ASEE Annual Conference, Honolulu, HI.

Dodge, B. (2001). *Creating a rubric*. Retrieved September 8, 2007 from http://edweb.sdsu.edu/webquest/rubrics/rubrics.html

Engineering education: Assessment methodologies and curricula innovation. Retrieved September 8, 2007 from http://www.engrng.pitt.edu/~ec2000/ec2000_attributes.html.

Froyd, J., Layne, J. Yurttas, L. & Ford, D. (2006, June). *Designing a process for department curricular reform*. Proceedings of the ASEE Annual Conference, Chicago, IL.

Maffia, G. J. (2004, October). *Comprehensive testing and outcomes assessment in the chemical engineering program*. Proceedings of the 34th ASEE/IEEE Frontiers in Education Conference, Savannah, GA.

National Research Council, (2000). *How people learn*. Washington, D.C.: National Academy Press.

Nichols, James O. (1995). *The departmental guide and record book for student outcomes assessment and institutional effectiveness*. New York: Agathon Press.

Nirmalakhandan, N., Daniel, D., & White, K. (2004). Use of subject-specific FE exam results in outcomes assessment. *Journal of Engineering Education*, *93* (1), *73-77*.

PAVE: Partnership to advance volunteer excellence. Retrieved September 8, 2007 from http://www.abet.org/pave.shtml

Pet-Armacost, J. & Armacost, R.L. (2005, October). *System design for program assessment to support ABET accreditation requirements*. Proceedings of the ABET 2005 Annual Meeting, San Diego, CA.

Rogers, G. (2006a). *Using course or test grades for program assessment*. Retrieved March 19, 2007 from http://www.abet.org/assessment.shtml.

Rogers, G. (2006b). *Rubrics: what are they good for anyway. part III*. Retrieved March 19, 2007 from http://www.abet.org/assessment.shtml.

Rogers, G. (2006c). *Direct and indirect assessments; what are they good for?* Retrieved March 19, 2007 from http://www.abet.org/assessment.shtml.

RosE Portfolio System (REPS) Zone. Retrieved March 20, 2007 from http://www.rose-hulman.edu/REPS/.

Rubrics for web lessons. Retrieved March 5, 2007 from http://edweb.sdsu.edu/webquest/rubrics/weblessons.htm

Watson, J. L. (1998). An analysis of the value of the FE examination for the assessment of student learning in engineering and engineering science topics. *Journal of Engineering Education*, *87* (3), 305-311.

CHAPTER 9

TWENTY YEARS OF ASSESSMENT: A RETROSPECTIVE CASE STUDY

Barbara M. Olds
Colorado School of Mines

Introduction

At the Colorado School of Mines (CSM), outcomes assessment has been a part of life for twenty years, first as the result of a state mandate and then as a response to ABET's EC2000 and regional accreditation requirements. This paper traces the four phases in CSM's assessment journey, from the original conception to the present day, focusing in each section first on the national (largely ABET) accreditation and assessment policy level and then on how policy played out locally at CSM. Lessons learned are discussed for each phase as well.

Phase I: The Beginning

The concept of outcomes assessment for higher education was definitely in the air in the mid to late 1980's as states, regional accreditation groups and professions began to stress the importance of measuring the value added of a college education. For example, the first National Assessment Conference of the American Association for Higher Education took place in 1985 and the periodical *Assessment Update* began publishing in 1989. However, institutions of higher education in states for which assessment had been mandated found themselves confused about their marching orders and uncertain about where to start. The questions faced in South Carolina and articulated by Johnson (1992) were typical:

> What were the practical implications of the assessment mandates, in terms of new activities and costs? What models of assessment programs were available? How were assessment programs to be planned and designed? ...What kinds of results could be expected and how interpreted? And most important to the entire endeavor, how could assessment results best be utilized for educational program improvement?" (p. 11)

Accrediting entities such as the Accreditation Board for Engineering and Technology (A.B.E.T., now ABET, Inc.) were also concerned about the approach to accreditation that had developed over the previous decades as traced thoroughly by Prados, Peterson, and Lattuca (2005) in a recent article in the *Journal of Engineering Education*. According to Prados et al, "by the late 1980s, engineering employers and visionary educational leaders were recognizing that the effective preparation of engineers for twenty-first century practice demanded fundamental changes in the dominant engineering science paradigm (p. 167)." The emphasis on engineering science, which had developed after World War II, produced graduates with strong technical skills, but without, in many cases, the "professional skills" many industry leaders believed necessary for success in the future. A number of studies and reports produced around this time (McMasters, 2004; David et al, 1987; National Academy Press, 1985) emphasized the need for such skills as "communication and persuasion, an ability to lead and work effectively as part of a team, an understanding of the non-technical forces that profoundly affect engineering decisions, and a commitment to lifelong learning" (Prados, Peterson & Lattuca, p.168).

However, according to Prados et al, "Despite its best intentions, the pre-1990 ABET could well be characterized as a protector of the status quo (p. 168)," maintaining adversarial relationships with engineering schools, relying on measurable outcomes such as seat time, developing increasingly prescriptive criteria, and in general developing a reputation for "bean counting" at the expense of innovative programs. Fortunately, ABET leadership recognized the validity of these criticisms and formed an Accreditation Process Review Committee (APRC) to advise it on ways to build more flexibility into the accreditation process while still maintaining the strong emphasis on educational quality that ABET had insisted on over the years. Thus the road that was to culminate in EC 2000 was under construction.

The Colorado School of Mines, like many other institutions of higher education, was also thinking about assessment in the late 1980s. At that time, Colorado, like many other states, became interested in higher education accountability and assessment and passed legislation (HB1187) requiring the Colorado Commission on Higher Education (CCHE) to develop an accountability policy and report annually on its implementation (Olds & Pavelich, 1996). In addition, the legislation required that institutions of higher learning be held accountable for improvements in student knowledge between entrance and graduation; that these improvements be publicly announced and available; that institutions express clearly to students their expectations of student performance; and that these improvements be achieved through effective use of time, effort and money. The state required each institution to report assessment of general education, discipline-specific education, retention and completion, alumni/student satisfaction, after-graduation performance, minority student statistics, and costs. According to the timeline established by CCHE, each institution was required to submit its institutional goals and objectives for approval in 1988 and then submit an assessment plan after the goals were approved. In 1989 the first assessment reports were submitted. As a "stick" to encourage compliance, the legislation stipulated that CCHE could retain two percent of an institution's appropriation if it found the assessment report "unsatisfactory."

Unlike several other states in which assessment was mandated at approximately the same time, Colorado allowed each institution to develop an individual assessment plan appropriate for its size, student body, mission and goals. After considerable input from alumni, recruiters, faculty, and students, CSM chose initially to develop a portfolio assessment program. The program was based on maintaining comprehensive longitudinal records for a sample of CSM students, for whom parental permission had been obtained. In brief, each year a random sample of incoming students was selected for whom CSM developed portfolios. For those students the School collected and reported such typical quantitative data as SAT and ACT scores and GPAs; in addition, the portfolios included samples of classroom work from a variety of courses as well as surveys and other feedback on the students' satisfaction with the institution. Each spring the portfolios were evaluated by an ad hoc faculty Assessment Committee whose summary provided the heart of an annual report to the campus and CCHE.

The Assessment Committee, comprised of members from disciplines across campus, met regularly during the academic year to discuss assessment issues and then for two days after the end of the school year to evaluate first and second year students' portfolios. The committee's evaluations and recommendations, along with those of departmental assessment committees who evaluated students' work in the major, formed the basis for the annual assessment report to the CSM community and to CCHE.

Based on CSM's institutional mission and goals, as defined in our Profile of the CSM Graduate (Colorado School of Mines, 2006), the committee decided to assess the following areas: technical knowledge and ability; communication skills, critical thinking and intellectual development; ability to self-educate; familiarity with the humanities and social sciences; and leadership and teamwork. Working with the core departments, the Assessment Committee identified specific materials to collect for the portfolio. The material was collected directly from faculty teaching the courses; once students (or their parents) had agreed to participate in the program, they had virtually no involvement in it.

A critical task was establishing the guidelines for evaluating the collected student work. This occupied a great deal of the Assessment Committee's time and talent in the first several years of the program. The Committee basically developed rudimentary rubrics (although they were not called that at the time) that turned out to be straightforward, analytical, and gave reproducible results.

In the decade or more in which CSM collected and evaluated student portfolios, they led to several sustained and critical changes at the institution. For example, the faculty evaluating first and second year writing noticed spottiness in the quality of student work and inconsistency in the grading standards employed by faculty. This was attributed to problems with the writing program, which had lost its leadership. The School committed to hiring a new Writing Program Director and other communication experts who redesigned and oversaw our writing-across-the-curriculum efforts, leading to solid improvement in this area (see the Campus Writing Program's website at http://www.mines.edu/Academic/lais/wc/mission2.htm).

In using the portfolio approach for a decade, CSM concluded that there are some clear advantages to portfolios, as articulated by Prus and Johnson (1994). These include that portfolios can be used to assess learning longitudinally; that multiple outcomes can be measured using a single portfolio; that portfolios are a more "authentic" form of assessment than many others; and that the process of reviewing portfolios can be a positive faculty development activity. However, there are also disadvantages, the most pertinent perhaps being cost in terms of both time and effort.

Lessons Learned from Phase I

The first, state-mandated, phase of assessment activity at CSM was generally positive in that it put the institution ahead of the EC2000 curve; it developed an "assessment mindset" at the School; and it put the mechanisms in place for responding to accreditation changes by identifying a coordinating group, the Assessment Committee. On the other hand, it became clear over time that the portfolio program was very resource intensive and would eventually collapse under its own weight; that the Assessment Committee wasn't "official" and therefore had no real power; and that CSM, like most other institutions, was generally naïve about both assessment and portfolios (and legalities—the rudimentary consent form would probably not pass an IRB today).

Phase II – The Development and Early Implementation of ABET 2000

ABET continued to examine its accreditation practices during the early to mid-1990s. Following a series of NSF-funded stakeholder workshops in 1994 and a synthesis workshop in 1995, ABET published the *Vision for Change,* which called for "fundamental changes in accreditation criteria, evaluation procedures, and selection and training of those who carry out the accreditation process" (quoted in Prados, Peterson & Lattuca, p. 169). Also in 1995 ABET's Board of Directors approved the publication of Engineering Criteria 2000 (EC2000) for public comment. EC2000 represented a fundamental shift in accreditation protocol away from "bean counting" and towards continuous improvement.

Pilot evaluations were conducted at five volunteer institutions in 1996 and 1997 using the new criteria. These visits helped "to refine the self-study questionnaires, assess the level of implementation of the continuous improvement processes, and train future team chairs and program evaluators" (Prados, Peterson & Lattuca, p. 169). After the pilot visits, institutions scheduled for an accreditation visit over the next three years could choose whether they wanted their ABET visit under the old or the new criteria. After that point, all evaluations were conducted using EC2000. ABET faced a number of challenges during the transition period including the need to retrain over 1000 volunteers, the difficulty of obtaining faculty buy-in to the process, and the need to train faculty to undertake their new roles as assessment "experts."

During the early 1990s, CSM continued to refine its portfolio assessment plan, and several members of the Assessment Committee became involved in engineering assessment activities at the national level. For example, Barbara Olds and Ron Miller, members of CSM's Assessment Committee, worked with Gloria Rogers of Rose-Hulman Institute of Technology (now at ABET, Inc.) to design and implement the Best Assessment Processes Symposium series, which began in 1998 and continues to the present (see http://dev.rose-hulman.edu/assessment2007/). Over a decade of BAP Symposia, hundreds of engineering, technology, and computer science faculty have had the opportunity to learn about current assessment processes from a variety of experts and exchange ideas with colleagues struggling with the same assessment questions.

Assessment was also an increasingly popular topic at meetings such as the annual American Society for Engineering Education Conference and the Frontiers in Education meeting. Olds served as guest editor for a special edition of the *Journal of Engineering Education* (April 1998), which focused on assessment issues. She and Miller (1998) contributed an article to the special issue titled, "An Assessment Matrix for Evaluating Engineering Programs." The matrix article was based on the assessment plan that the CSM Assessment Committee had adopted in preparation for a 2000 ABET visit under the new EC2000 criteria, which CSM had chosen to use. The matrix is replicated in Table 1. Olds and Miller argued that the program assessment matrix "provides faculty members (especially ones with little assessment experience) with a structure for developing their assessment plan using a series of questions" (p. 173) related to program goals, objectives, performance criteria, implementation strategy, assessment methods, timeline, and feedback. The CSM Assessment Committee held a series of workshops for CSM faculty and department heads in preparation for the 2000 ABET visit using the matrix as the organizing principle, and all programs agreed to use the matrix in developing their assessment plans and reporting results in their ABET self-studies.

Table 1
Program Assessment and Evaluation Matrix

Educational Objectives: What are the overall goals of the program? How do they complement institutional and accreditation expectations?

Student Outcomes	Performance Criteria	Implementation Strategy	Evaluation Methods	Logistics	Feedback
Using measurable language, what should your students know and be able to do?	How will you know the outcomes have been achieved? What level of student performance will indicate students possess the outcome?	How will the outcomes be met? What program activities (curricular and co-curricular) help your students meet each outcome?	What assessment methods will you use to collect data? How will you interpret and evaluate the data?	When will you measure? How often? Who will collect and interpret the data and report the results?	Who needs to know the results? How can you convince them the outcomes were met? How can you improve your program and your assessment process?

Lessons Learned from Phase II

ABET made an accreditation visit to the CSM campus in the fall of 2000 and, by all accounts, the visit went very well. All programs received full six-year accreditation and, more importantly, all programs appeared to buy into the concept of assessment and accreditation. This phase was the "honeymoon" of assessment activities at CSM to date for several reasons:

- A well tested assessment system was in place on the campus and had been for over a decade. A sizable number of faculty members were becoming comfortable doing assessment and were using assessment results to make positive changes in their programs.
- The use of the assessment matrix allowed all programs to follow a similar process, but to vary their approach within that process to meet individual program needs.
- The Assessment Committee had been in existence for over a decade, and several members of the committee were developing national and international reputations for their work in the area.
- Faculty workshops well ahead of the ABET visit allowed for much broader dissemination of assessment activities across the campus.
- There was strong administrative support for assessment efforts and the word was clearly received by all programs that this was an activity in which each program was expected to participate fully.

Phase III – The Honeymoon is Over

ABET continued to evaluate the impact of the changes brought about by the implementation of EC2000. In 2002 ABET reconvened stakeholder workshops to discuss the progress of its reform efforts. In the workshops, the following concerns, among others, were noted (Prados, Peterson & Lattuca, p.170):

- The apparent focus on quantity of data rather than assessment of quality.
- Sustainability questions
- Sustained commitment by ABET, institutional, and professional society leadership
- Inconsistent performance by evaluators

These concerns were also voiced by participants in the annual Best Assessment Processes symposia and in many conversations among engineering educators. In some cases, programs admitted, among friends, to using a "smoke and mirrors" approach to assessment, hoping that producing large volumes of material would make up for lack of depth in the evaluation of the material. In other programs, faculty burnout became a big issue. Faculty in some cases devised very elaborate assessment plans that were in danger of collapsing under their own weight at any time, especially if the "inventor" moved on to other endeavors. Sustainability questions also arose over financing assessment and over how to compensate faculty who spent large amounts of time on assessment activities. There were also questions about the long-term commitment of ABET to stay the course on assessment. More than a few cynics argued that "if we just wait this out, it will disappear in the way that all reform movements do." Finally, there were concerns over the inconsistent quality of the evaluators who represented ABET in accreditation visits. Since ABET now had to retrain its large evaluator force, there is no question that this was a key, and challenging, piece of the transition to EC2000.

However, there were also positive findings from the stakeholder workshops, including:

- "growing acceptance of the value of the systematic engagement of external constituencies in improving program quality;
- growing awareness of the value of outcomes-based assessment for improving program quality;
- increased faculty attention to student learning in improving program quality; and
- growing involvement of industry at the program level." (Prados, Peterson & Lattuca, p. 170)

In an effort to better understand the impact of EC2000, in 2002 ABET commissioned a study by the Center for the Study of Higher Education at Pennsylvania State University. The study (Lattuca, Terenzini & Volkwein, 2006) was designed to answer two questions (p.1):

- What impact, if any, has EC2000 had on student learning outcomes in ABET-accredited programs?
- What impact, if any, has EC2000 had on organizational and educational policies and practices that may have led to improved student learning outcomes?

The resulting 426-page report, *Engineering Change: A Study of the Impact of EC2000*, presents and analyses the results of extensive surveys of engineering graduates, faculty members, program chairs, and employers in seven engineering fields, both in the aggregate and broken out by discipline.

The report concludes that "today, according to the accumulated evidence in *Engineering Change*, engineering education in the United States has changed substantially since the mid-1990s. Engineering programs and faculty members have modified their curricula, teaching methods, professional development practices, program assessment and decision making, and, to some extent, their hiring, promotion, and tenure criteria. The findings from the EC2000 study also strongly suggest that improvements in student learning have resulted from these changes in engineering programs" (p. 123). The study also concludes that there are multiple dimensions to a college experience and that, while classroom work may be the most influential, outside-of-class experiences, such as internships and study abroad, can also have a profound effect on learning.

Following the successful ramp-up to its 2000 ABET visit, the CSM campus community, like many others who have experienced the intensive pressure of preparing for an accreditation visit, fell into an assessment "slump." During a period of several years, the Assessment Committee met only infrequently and programs, while they continued to gather assessment information, did so largely in a desultory manner and made little use of the information gleaned. However, with the specter of another ABET visit in 2006 and rumors that ABET would be "grading" harder on the second go-round under EC2000, activity was rekindled across campus. As part of the preparation for the 2006 ABET visit, the Assessment Committee was re-energized under new leadership. Each

ABET-accredited program met with the Committee to outline its assessment activities including program objectives and outcomes, assessment methods, materials collected, evaluation process, and plans for closing the feedback loop. These sessions allowed the Committee to update its understanding of the state of assessment on the campus and to make recommendations to the programs on ways to improve their assessment processes. Ultimately, a second successful ABET accreditation visit was held in the fall of 2006.

Lessons Learned from Phase III

After the "slump" following the 2000 ABET visit, the CSM assessment effort gradually reemerged. Lessons learned from this third phase included the recognition that assessment enthusiasm must be tempered with a realistic appraisal of how to balance assessment with all of the other activities to which faculty members are committed. Thus, there was a need at CSM to re-evaluate the amount of effort being put into assessment activities in order to avoid campus-wide burnout. As a result, programs reflected on their processes and adapted a variety of approaches that led to more efficient, but also potentially more useful, assessment processes. For example, some programs decided that it was redundant to collect information on every student every semester for some outcomes. Instead, they moved to collecting material once a year or even once every two or three years with no loss of momentum.

At the institutional level, the decision was made to participate in the National Survey of Student Engagement (NSSE) every three years rather than annually. The NSSE "annually obtains information from random samples of first-year and senior students about the nature of their undergraduate experience" (NSSE, 2006). Participation allows an institution to compare its students' experiences with those at peer institutions as well as with the overall NSSE sample. The change to administering the NSSE every three years has allowed CSM to track trends in student engagement over time, but does not require the investment of time and expense needed for a yearly survey. Participating every three years also allows the institution to more carefully evaluate the rich data collected from each administration of the survey and thus to plan curricular and co-curricular changes more deliberately and carefully.

CSM noticed that many of the conclusions reported in the *Engineering Change* study were reflected on our campus and that the mechanisms are in place on the campus to lead to continuous improvement. Although the changes at CSM do not include 100 percent of the courses and faculty, clearly "engineering programs and faculty members have modified their curricula, teaching methods, professional development practices, program assessment and decision making" (Lattuca, Terenzini & Volkwein, p. 123). These changes will be discussed more fully in the Lessons Learned section of Phase IV.

Phase IV - The Future

Those who believed twenty years ago that assessment was just another higher education fad have been proven wrong. In fact, assessment and accreditation in higher education have recently been in the national spotlight as U.S. Secretary of Education Margaret Spellings convened a commission to examine higher education in this country.

The resulting report, *A Test of Leadership: Charting the Future of U.S. Higher Education* (2006), arrives at some harsh conclusions regarding higher education accreditation:

> Accreditation, the large and complex public-private system of federal, state and private regulators, has significant shortcomings. Accreditation agencies play a gatekeeper role in determining the eligibility of institutions and programs to received federal and state grants and loans. However, despite increased attention by accreditors to learning assessments, they continue to play largely an internal role. Accreditation reviews are typically kept private, and those that are made public still focus on process reviews more than bottom-line results for learning or costs. The growing public demand for increased accountability, quality and transparency coupled with the changing structure and globalization of higher education requires a transformation of accreditation (p. 15).

Since the Commission made its recommendations, Secretary Spellings has convened an Accreditation Forum (2006) to look at "ways to streamline and improve the accreditation process to support innovation, promote consistency in accreditation standards, increase accountability, and become more transparent to the public."

The Forum has struggled with the issues articulated here and no solution has yet been agreed upon by all parties, though the Department of Education appears to be moving forward with its agenda.

At the same time, the *Engineering Change* study has confirmed the theory that assessment is a much more complex activity than it might initially appear:

> The evidence indicates that, whether direct or indirect, the impact of any single factor, by itself, tends to be small, statistically significant, perhaps, but often substantively unimpressive. More impressive, and perhaps daunting, are the patterns of interconnections that this study has only begun to map. One implication of this finding is that a programmatic or policy focus on only one or a handful of revisions or interventions is likely to be effective only at the margins. Any single "silver bullet" or "best practice" may have its impact on some aspect of student learning, and such enhancements are not to be discussed. The effects of multiple programmatic or policy revisions, of increased opportunities for students to link and extend their in- and out-of-class experiences, however, are more likely to bring about significant and substantial change (Lattuca, Terenzini & Volkwein, 2006, p. 126).

<u>Assessment at CSM at the Present and in the Future</u>

CSM, like other engineering schools, is responding to assessment mandates from multiple sources including the State of Colorado, the Higher Learning Commission of the North Central regional accreditation organization, and ABET. Thus, also like other institutions, it is trying to develop assessment processes that are as efficient and effective as possible. Some promising recent events provide evidence that CSM is using the continuous improvement model to improve its assessment processes and learning outcomes as well as to become a learning organization.

- The Assessment Committee was made an official University Committee in 2007. Thus, it has transitioned from being a long-standing ad hoc group to a fully institutionalized entity. As a University Committee it is responsible for helping to guide CSM in matters pertaining to assessment of the program educational objectives and program outcomes of its undergraduate and graduate programs. Among the tasks delegated to the University Assessment Committee are to:

1. Review, on an annual basis, undergraduate and graduate assessment plans provided by each academic unit as required by the administration.
2. Review, on an annual basis, documentation provided by each academic unit, which indicates how the unit has carried out its assessment plan, and what changes it has made to its academic programs as a result.
3. Recommend additional actions academic units could take to enhance their assessment efforts.
4. Help to implement CSM's assessment plan for its core undergraduate and graduate requirements, evaluate this plan and outcomes annually, and recommend actions based on these, and
5. Oversee the production of an annual institutional assessment report led by the Office of Educational Innovation.

Thus, for the first time, the Assessment Committee has a clear mandate with some "teeth." It can use that mandate to move the assessment efforts on campus in a positive direction

- CSM is beginning to assess graduate student outcomes in a systematic way, partly as a result of regional accreditation pressures, but primarily because CSM believes it is the right thing to do and an important next step in assessment activity. Although ABET accredits graduate programs at the master's level, according to the ABET/EAC policy II.B.8.A, a program may be accredited at either the bachelor or masters level, but not both, and thus most programs, like those at Mines, choose the baccalaureate level. However, the School's experience with ABET has been invaluable in setting the stage for graduate assessment. In developing a process for graduate assessment, the Assessment Committee held a faculty workshop and developed a template for annual reports on graduate assessment to complement the already-existing undergraduate template (see Tables 2 and 3). The committee has followed the advice articulated by Lovitts for "making the implicit explicit" (2006, p. 163). Lovitts' study at nine doctorate-granting institutions asked focus groups of faculty to characterize dissertations at four different quality levels—outstanding, very good, acceptable, and unacceptable. The result was a compilation of characteristics of dissertations at each level, a way of making implicit faculty assumptions about dissertations explicit and therefore open to discussion. CSM's Assessment Committee plans to conduct a similar cross-campus focus group and to use the results to build a rubric for evaluating the quality of dissertations produced at the School.

In addition, the Assessment Committee has adopted the recommendations of the Council of Graduate Schools (2005) in developing and implementing its graduate assessment program. According to the CGS, outcomes-based assessment "is a process of determining the indicators of an effective graduate program,

Table 2
Undergraduate Program(s)
Instructional Unit Assessment Report—Due Annually on October 1

Instructional Unit Information

Department/Division/Program:

Degree programs to which report applies:

Documentation of Assessment and Evaluation Process

(Attach copy or provide URL of current Assessment and Evaluation Plan)

List names of faculty charged with evaluation of assessment materials for instructional unit:

Briefly describe changes instituted to Assessment and Evaluation Plan since submission of last report:

Briefly describe assessment data collected since submission of the last report:

Documentation of Program Improvement Efforts

Briefly describe results of evaluation of assessment activities completed since submission of last report:

What actions/changes has the instructional unit made based on these evaluations?

Table 3
Graduate Program(s)
Instructional Unit Assessment Report—Due Annually on October 1

Instructional Unit Information
Department/Division/Program: Degree programs to which report applies:

Documentation of Assessment and Evaluation Process (Attach copy or provide URL of current Assessment and Evaluation Plan)
List names of faculty charged with evaluation of assessment materials for instructional unit: Briefly describe changes instituted to Assessment and Evaluation Plan since submission of last report: Briefly describe assessment data collected since submission of the last report:

Documentation of Program Improvement Efforts
Briefly describe results of evaluation of assessment activities completed since submission of last report: What actions/changes has the instructional unit made based on these evaluations?

using those indicators as criteria for evaluating the program, and applying the results of this evaluation toward the ongoing and continuous improvement of the program" (p. 23).

- CSM has put in place the Trefny Institute for Educational Innovation. The Institute is comprised of 1) the Center for Engineering Education (CEE), 2) the Office of Innovation in Learning and Teaching, and 3) the Center for the Assessment of Science, Technology, Engineering, and Mathematics (CA-STEM). By pulling these programs under one umbrella, the Institute provides participating faculty with the opportunity to integrate their findings in a "cycle of innovation." Thus education research findings will feed into improving teaching, learning, and assessment. Faculty who initially focus on curriculum and pedagogy improvement will have opportunities to become education researchers. The assessment center will highlight current exemplars in curriculum design informed by education research and also suggest those areas in the curriculum that might be better informed by what has been learned from education research.

In summary, we have learned a great deal over the past twenty years about both the process of conducting assessment and the value of it. Not everything has gone smoothly, not every year has been a banner year, but, like many engineering schools, we have persevered, learned, and developed an assessment program that works for us.

References

ASEE Task Force (David, E. E. Jr., Chair). (1987). *A national action agenda for engineering education*. Washington, DC: American Society for Engineering Education.

Baker, M. J., revised and updated by King, M. F., Larick, D.K., & Carter M. P. (2005). *Assessment and review of graduate programs: A policy statement*. Washington, DC: Council of Graduate Schools.

Colorado School of Mines. (2006). *Undergraduate Bulletin*, 5-6. Retrieved June 18, 2007 from http://www.mines.edu/index_js.shtml

David, E. E., Jr. (1987). *A national action agenda for engineering education, report of an ASEE task force (E. E. David, Jr., Chair)*. Washington, D.C.: American Society for Engineering Education.

Johnson, R. (1992). *South Carolina higher education assessment project final report*. Retrieved June 18, 2007, from http://eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED359862&ERICExtSearch_SearchType_0=eric_accno&accno=ED359862

Lattuca, L. R., Terenzini, P. T., &Volkwein, J. F. (2006). *Engineering change: A study of the impact of EC2000*. Baltimore, MD: ABET, Inc.

Lovitts, B. E. (2006). Making the implicit explicit. In P. L. Maki & N. A. Borkowski (Eds.), *The assessment of doctoral education: Emerging criteria and new models for improving outcomes* (pp. 163-187). Sterling, VA: Stylus Publishing.

McMasters, J. H. (2004). Influencing engineering education: One (aerospace) industry perspective. *International Journal of Engineering Education*, *20*, (3), 353-371.

National Academy Press. (1985). *Engineering education and practice in the United States*. (1985). Washington, D.C.: author.

National survey of student engagement. Accessed March 14, 2008 from www.nsse.iub.edu.

Olds, B. M., & Miller, R. L. (1998). An assessment matrix for evaluating engineering programs. *Journal of Engineering Education*, *87* (2), 173-178.

Olds, B. M. & Pavelich, M. J. (1996, June). *A portfolio-based assessment program*. Proceedings of the ASEE Annual Conference, Washington, D.C.

Prados, J. W., Peterson, G. D., & Lattuca, L. R. (2005). Quality assurance of engineering education through accreditation: The impact of Engineering Criteria 2000 and its global influence. *Journal of Engineering Education*, *94* (1), 165-184.

Prus, J. & Johnson R. (1994, Winter). Assessment and testing, myths and realities. *New Directions for Community Colleges*, 88.

Secretary of Education's Commission on the Future of Higher Education. (2006). *A test of leadership: Charting the future of U. S. higher education*. Retrieved June 19, 2007 at http://www.ed.gov/about/bdscomm/list/hiedfuture/index.html.

Secretary Spellings convenes accreditation forum in Washington, D.C., with key stakeholders (November 29, 2006). Retrieved June 19, 2007 from http://www.ed.gov/news/pressreleases/2006/11/11292006.html.

CHAPTER 10

ASSESSMENT IN ENGINEERING EDUCATION

Victoria E. Robson, Vinod K. Lohani, and John A. Muffo
Vanderbilt University
Virginia Polytechnic Institute and State University
Ohio Board of Regents

Introduction

A new Department of Engineering Education (EngE) was created within the College of Engineering (COE) at Virginia Tech (VT) in May 2004 to improve engineering pedagogy within the COE and to initiate engineering education research activities. Engineering freshmen (~1500 every year) at COE are required to complete a General Engineering (GE) (also called freshman engineering) program before they can advance on to one of eleven degree-granting departments. The EngE faculty is responsible for conducting the GE program. Another primary mission of the EngE department is to carry out rigorous research in the area of engineering education and to support the research agenda as described in a recent article titled "The Research Agenda for the New Discipline of Engineering Education" that appeared in the October 2006 issue of the *Journal of Engineering Education*. Such rigorous research efforts in engineering education require collaboration between engineering and education faculty within and outside the university.

A National Science Foundation (NSF)-supported planning grant "Bridges for Engineering Education—Virginia Tech (BEEVT)" laid the foundation for such an engineering-education collaborative at Virginia Tech in 2003 (Lohani, Sanders, et al., 2005). One objective of BEEVT was to create a contemporary framework for undergraduate engineering pedagogy, beginning with freshman engineering experiences. Accordingly, BEEVT investigators (engineering and education faculty) proposed to reformulate the engineering curriculum in one of the engineering departments, namely, Biological Systems Engineering, using a spiral curriculum approach.

The twentieth century psychologist, Jerome Bruner, proposed the concept of the spiral curriculum in his classic work, *The Process of Education* (Bruner, 1960). Bruner advocates that a curriculum, as it develops, should revisit basic ideas repeatedly, building upon them until the student has grasped the full formal apparatus that goes with them. This approach has been adopted at VT in a 4-year (2004–2008) implementation grant under a Department-Level Reform (DLR) program of the NSF, hereafter referred to as DLR project (Lohani, Sanders, et al., 2005).

As part of the DLR project, a number of EngE faculty members are collaborating with faculty from other engineering departments and the School of Education to reformulate the freshman engineering program within the EngE and the bioprocess program within the Biological Systems Engineering (BSE) department using a theme-based spiral curriculum approach (Bruner, 1960). One of the objectives of the DLR project is to develop a continual assessment plan to measure the impact of the reformulated curricula, faculty improvement activities, and student learning. To accomplish this, the DLR investigators are collaborating with the VT Office of Academic Assessment to develop and implement various formative and summative assessment tools.

This chapter provides a description of the assessment tools and procedures developed to initiate assessment activities in collaboration with assessment experts in the GE program. Further, analysis of assessment data is presented to identify and highlight predictors of success in the engineering program and demonstrate the impact of assessment activities in improving a freshman engineering course. The authors have considered two measures of success (grade-point average and retention of students after completing ~60 credits worth of course work) and considered predictors of success like course grade in an introductory freshman engineering course. A number of studies have documented that experiences of students in the first semester or year is crucial to their performance in subsequent semesters (Budny, LeBold, & Bjedov, 1998; Heywood, 2005).

Freshman Engineering Course Engineering Exploration EngE1024

The EngE offers a common one-year GE program for initial preparation of approximately 1500 incoming engineering freshmen. All engineering freshmen are required to take a two-credit Engineering Exploration (EngE1024) course during their first semester of enrollment. This is the only common course all engineering undergraduates take within the COE. The course primarily focuses on developing problem solving, critical thinking, and engineering design skills. The key learning objectives of this course are to (a) impart early design experiences along with teamwork and communication skills; (b) graph numeric data collected in hands-on experiments describing basic engineering experiences and derive simple empirical functions; (c) demonstrate an understanding of professional ethics and application to real-life situations; and (d) develop and implement algorithms that focus on object-oriented approaches.

The DLR investigators have implemented a number of activities in this course in recent years to make it learner-friendly, contemporary, and research-driven (Lo, Lohani, & Griffin, 2006; Lohani, Sanders, et al., 2005). Some examples include:

- Use of classroom response system (i.e., clickers) to obtain students' feedback (Lohani & Lo, 2006);
- Introduction to sustainability (Mullin, Lohani, & Lo, 2006);
- Use of ethics skits to instruct engineering ethics (Mullin, Lohani, & Lo, 2006);
- Introduction of international activities (Lohani, Mullin, Lo, & Griffin, 2006);
- Use of electronic portfolio (e-portfolio) for instruction (Knott, M. et al., 2005; Knott, T. W. et al., 2004);
- Use of multiple models of a problem to instruct different aspects of the course (Connor, Lohani, Mallikarjunan, Loganathan, & Lo, 2006);
- Introduction to systems concepts (Mallikarjunan, Laksmikanth, Wolfe, Lohani, & Connor, 2007); and

- Use of mechatronics to introduce multidisciplinary design to engineering freshmen (Lohani, Kachroo, et al., 2006).

A number of these activities (i.e., sustainability, systems concepts, e-portfolio, ethics instruction, international activities) are chosen so that they can be linked to activities in upper- level learning modules, particularly in the bioprocess engineering program, which is the focus of the DLR project. The scope of these activities has been kept fairly general so that they remain equally applicable to all engineering curricula. In the 2006–07 academic year, Tablet PC-based instructions were introduced in this course (Lohani, Castles, Lo, & Griffin, 2007).

Method

Assessment Measures

A number of assessment (formative and summative) activities are being implemented in EngE1024 as part of the DLR project (Muffo, Lohani, Mullin, Backert, & Griffin, 2005; Robson, Lohani, & Bateman, 2007; Robson, Muffo, & Lohani, 2006) to evaluate the learning experiences of engineering freshmen. This chapter focuses on results of assessment activities in EngE1024, and assessment activities to evaluate learning experiences in upper-level courses in the bioprocess program are being developed at the time of this writing.

Engineering Education New Student Survey

A New Student Survey was developed by the DLR investigators, representing several years of teaching experience, to assess both the students' academic background prior to enrolling in college as well as the kinds of experiences that they have had involving computers, mechanics, and automobiles. An example survey question is, "Have you ever installed software on a personal computer?" This survey has been implemented in EngE1024 since fall 2004 and has undergone several revisions, specifically, eliminating items with no variance (e.g., "Have you completed four years of mathematics?"). The survey's main objective is to examine the role of prior experiences in academic success and retention of students in engineering. Although student background is not a specific part of the spiral concept incorporated in the DLR project, the data being collected have potential to answer questions like: "What type of prior experiences lead to choice of the bioprocess engineering major?"

Computer Attitudes Survey

The Computer Attitudes Survey assesses students' attitudes towards computers and related technology. An example item from this survey is, "Computers make me feel uneasy." This survey was developed by combining items from other published questionnaires concerning attitudes toward technology.

Learning Styles

This online survey, developed by Felder and Solomon (2006), measures learning style (Felder & Spurlin, 2005) preferences on four dimensions: (a) active vs. reflective, (b) sensing vs. intuitive, (c) visual vs. verbal, and (d) sequential vs. global.

The survey includes 44 pairs of statements, and students are asked to choose from each pair the statement that best describes them.

Programming Knowledge (Pre- and Post-Test)

Starting in fall 2004, a programming language called Alice (www.alice.org) was introduced in EngE1024 to teach introductory concepts of object-oriented programming. Prior to fall 2004, MATLAB software was used to cover introductory programming concepts. A primary reason for adopting Alice was to bring "object-oriented programming"-related content into the EngE1024 course to partially satisfy the need of the Computer Science (CS) department that became part of the College of Engineering beginning in fall 2004. Since then, all CS-bound students have been required to take EngE1024. There was no data basis for making this change from MATLAB to Alice, and DLR is the first major effort that's helping EngE faculty to develop an assessment database for examining learning experiences of engineering freshmen. The Alice programming environment uses the drag-and-drop method to develop a virtual world (i.e., a computer program) and has a built-in collection of three-dimensional objects that are utilized to develop and run computer programs. The Alice programming environment allows the user to construct very simple to fairly complex computer programs that can be used to solve engineering or other real-world problems. Further, the user interface of the Alice programming environment is designed to make computer programming more accessible to people without a significant amount of programming experience. A Programming Concepts multiple-choice test with 19 questions was adopted from the Alice collaborators (S. Cooper, personal communication, 2004) to provide a pre- and post-test measure for assessing the effectiveness of Alice instruction. Based on faculty experiences and students' feedback, the pre-and post-test was revised in fall 2005. The revised measure included only 13 questions. This test was used to collect pre-and post-test data until fall 2006, when it was decided to replace Alice with LabVIEW programming. The summary of data analysis section in this chapter discusses specific reasons for this change. Table 1 includes the descriptive statistics for the programming test (i.e., Alice).

Course Exit Survey

This end-of-semester survey was designed by DLR investigators to measures perceptions of learning outcomes in different areas including engineering design, logical thinking skills, teamwork, etc. An example is, "Have your problems solving and logical thinking skills improved as a result of this course?"

Focus Groups

Focus groups were conducted to gather in-depth feedback on specific aspects of the course (e.g., hands-on activities, programming instruction).

Table 2 presents the period of data availability for various assessment tools implemented in the EngE1024 course.

Procedure and Timeline

Fall 2004

At the beginning of the fall 2004 semester, students in Engineering Exploration (EngE1024) were instructed to complete three surveys including the New Student Survey, Computer Attitudes Survey, and the Learning Styles Survey online as part of a course

Table 1

Descriptive Data (in proportions correct) for Programming Knowledge Tests by Semester

Semester	*n*	Pre-test	Post-test	Gain Scores (Post-test – Pre-test)
Fall 2004	888	.47 (.14)	.58 (.16)	.11 (.16)
Spring 2005	198	.45 (.14)	.58 (.15)	.12 (.15)
Fall 2005	673	.42 (.18)	.68 (.18)	.26 (.19)
Spring 2006	111	.44 (.16)	.62 (.17)	.26 (.18)

Note. Numbers in parentheses represent standard deviations.

Table 2

Timetable of Assessment Measures for EngE1024

	Fall 2004	Spring 2005	Fall 2005	Spring 2006
New Student Survey	*	*	*	*
Computer Attitudes Survey	*			
Learning Styles	*	*	*	*
Programming data	*	*	*	*
Course Exit Survey		*	*	*
Focus Groups		*		*

assignment. The Web server of the Office of Academic Assessment hosted the New Student Survey and the Computer Attitudes Survey. Students completed the Learning Styles Survey online from a website hosted by the creators of the survey. After approximately 10 weeks of instruction, students completed the programming concepts pre-test in class. Students took the same test (i.e., the post-test) after programming instruction concluded. Approximately 800 students participated in these surveys. A detailed description of Alice programming activities is provided by Snook et al. (2006).

Spring 2005

At the beginning of the spring 2005 semester, students enrolled in EngE1024 completed only[1] the New Student Survey and the Learning Styles Survey online as part of a course assignment. The procedure for administering the programming concept tests was the same as in the fall 2004 semester. During the last few weeks of the semester, two focus groups were conducted by the third author. This was the first time that focus

[1]Since there was little variance on the responses from the Computer Attitudes Survey, this survey was eliminated. An example of an item with no variance is, "Learning about computers is a waste of time."

group interviews were conducted to assess the learning outcomes in this course. In addition, students were asked to complete a course exit survey.

Fall 2005 and Spring 2006

The assessment procedures were identical to the spring 2005 semester except that the programming pre- and post-test was modified (i.e., included 13 questions instead of 19).

Results

Demographic Data

Based on analysis of data of two engineering classes (i.e., class of 2008 and 2009), it is observed that the majority of the freshman engineering class is male (85%) and white (80%). In regards to prior background experiences, approximately 50% of the class has an engineer in the family and also has prior programming experience. In addition, the majority of students (75%) did not take any pre-engineering courses in high school. When asked to indicate all[2] of the engineering majors they were interested in, students selected mechanical engineering (53%), followed by aerospace and ocean engineering (40%), electrical or computer engineering (37%), and civil or environmental engineering (33%). In regards to learning styles, the majority of the students are active, sensing, visual, and sequential learners (see Figure 1). The pattern of learning styles for engineering students at VT is comparable to the pattern reported by Felder and Brent (2005) for engineering students from Iowa State University.

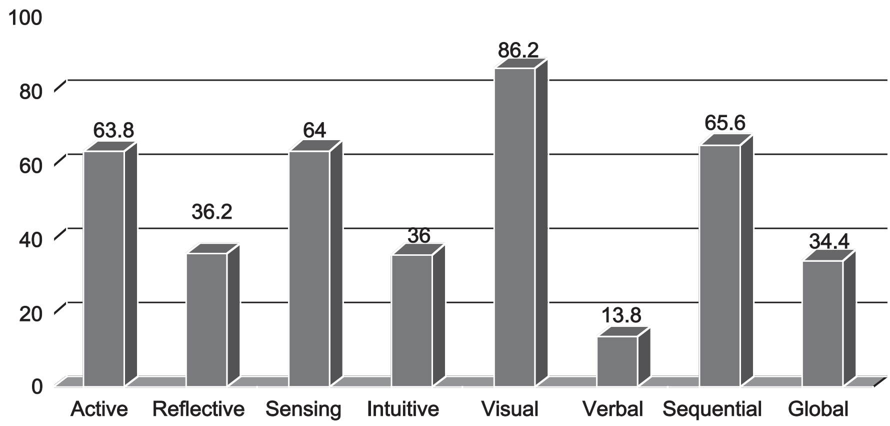

Figure 1.
Learning style preferences—engineering freshmen at Virginia Tech.

[2]Students were allowed to select multiple possible majors.

New Student Survey

Factor analysis, using varimax rotation, was conducted to determine if there were survey items that correlated with each other and, as such, reflected underlying constructs. The survey was constructed to measure multiple constructs such as computer knowledge, mechanical knowledge, prior course work, etc. Based on the observed patterns, four underlying scales were computed (by averaging responses across the items) from the New Student Survey items. Table 3 displays the factor loadings for the survey. Factor one (α = .76) comprised items related to computer knowledge (e.g., "Do you know what open source software is?"). Factor two (α = .70) contained items related to mechanical/automobile knowledge (e.g., "Have you worked on the engine or transmission of an automobile?"). Factor three (α = .72) included items regarding engineering-related high school course work (e.g., "Did you take a mechanical drawing or drafting class in high school?"). Finally, factor four (α = .52) was composed of items related to time spent studying and reading (e.g., "During your senior year in high school, how many hours per day did you spend studying?").

Table 3
Factor Loadings for the New Student Survey

Items	1	2	3	4
Do you know what open-source software is?	.73			
Do you know to write a simple computer program in any one of the following computer languages: Basic, Fortran, Pascal, Visual Basic, C, C++, Java?	.71			
Have you taken a formal course in computer programming?	.67			
Do you know how to use DOS?	.64			
Have you installed a hardware component inside a personal computer?	.62			
Have you used Linux or Unix?	.54			
Do you have a personal website that you developed yourself?	.52			
Have you used a personal computer for anything other than word processing, games, email, or web browsing?	.51			
Have you installed games on your personal computer?	.44			
Have you changed the oil and oil filter of an automobile?		.63		
Have you replaced a fuse or reset a breaker in your home?		.60		
Have you changed an automobile tire?		.60		
Have you worked on the engine or transmission of an automobile?		.59		
Have you used a ratchet and socket set?		.57		
Have you built or help to build a house, garage, or large shed?		.54		
Have you repaired an electronic device other than a computer?		.47		
Do you know where the fuse panel or breaker panel is in your home?		.45		
Did you take a mechanical drawing or drafting class in high school?			.76	
Did you take a computer-aided design class during high school?			.76	
Were you enrolled in any pre-engineering course in high school?			.66	
During your senior year in high school, how many hours per day did you spend studying?				.57
During your freshman year, how many hours do you expect to study?				.52
During your senior year in high school, how many hours per day did you spend reading for pleasure?				.51

Note: *n* = 830

Predicting Engineering Success

The analyses relating to predicting engineering success comes from the sample of first-year engineering (~ 900) students who started in the fall 2004 semester. Cumulative grade-point average (GPA) and retention data were collected (from the VT Office of Institutional Research) at the end of spring 2006 for the same fall 2004 cohort. For the purpose of this analysis, success in the engineering program for the fall 2004 cohort is defined using three factors: (a) cumulative GPA after four semesters (i.e., at the end of spring 2006 semester), (b) course grade in the first engineering course (i.e., EngE1024), and (c) retention in the engineering program. The retention factor is interpreted as the continuation of a student in the engineering program at the end of the 4th semester (i.e., spring 2006). The data were analyzed using SPSS 11 for Mac OS X. Forced entry multiple-regression analyses were conducted using the recommended procedure in Pedhauzer (1997).

Cumulative GPA

Regression analysis was performed to predict cumulative GPA at the end of the 4th semester using SAT scores and course grade in EngE1024 as predictors. Together, SAT scores and course grade in EngE1024 (taken in fall 2004) explained approximately 55% of variance in cumulative GPA at the end of spring 2006. That is, higher GPAs at the end of 4th semester were associated with higher SAT scores and course grades in EngE1024 in their 1st semester. Course grade alone accounted for 47% of the variance in GPA (see also Table 4 and Figure 2). The introductory EngE1024 course is a 2-credit course, and the average number of credits completed at the end of the 4th semester is approximately 60 credits.

Table 4
Summary of Variables Predicting Cumulative GPA after 4 Semesters

Variable	Beta	Sig.	R-Square
Step 1			
SAT	0.004	0.870	0.080
Step 2			
Course Grade in EngE1024	0.741	0.000	0.552

Note: n = 878

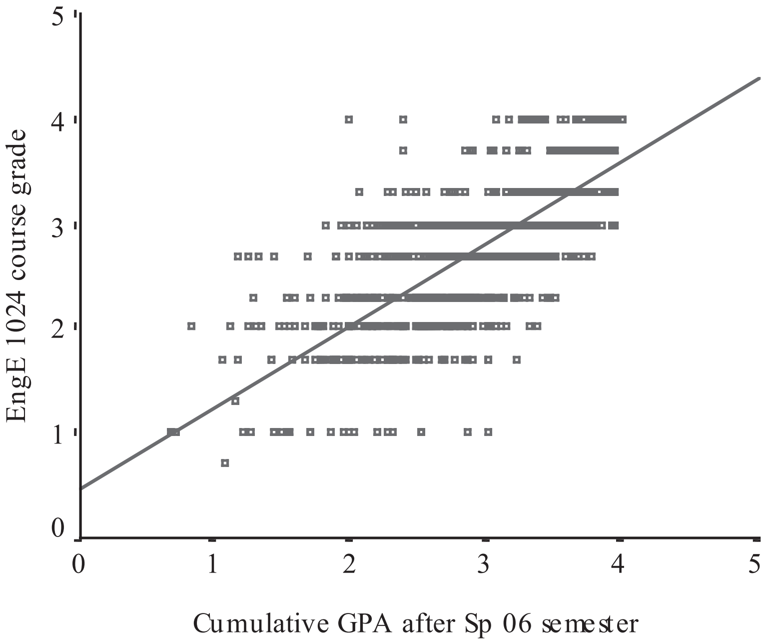

Figure 2.
Comparison between EngE1024 course grade and cumulative GPA.

Course Grade

Since EngE1024 grade was a strong predictor of cumulative GPA, a series of regression analyses were conducted to explore what factors including prior experiences predicted grade in EngE1024. First, regression models that included demographic variables, prior experiences, and learning preferences (see Table 5 for descriptive statistics) were tested. Gender and race (whites and minorities) were entered into block one, followed by learning style preferences (block two), and, finally, the New Student Survey factors (block three). Gender and race together as a set were significant predictors of course grade. This prediction was primarily a function of race. White students performed better in EngE1024 than the minorities. As a set, the learning style variables were marginally significant predictors of course grade above and beyond race and gender. This was a function of the active – reflective dimension. Specifically, reflective learners had a significantly higher course grade than active learners. Reflective learners prefer to think about information first, whereas, active learners learn best by doing something active with the information. Lastly, the four factors from the New Student Survey were also marginally significant predictors of course grade. The only New Student factor to approach statistical significance was the factor pertaining to computer knowledge. Taken together, gender, race, learning style preferences, and prior experiences as measured by the New Student Survey only explained approximately five percent of the variance in course grade (see Table 6).

Table 5
Descriptive Statistics for Course Grade by Demographic Variables

Variable	*n*	Mean	S.D.
Gender			
Male	617	2.767	0.684
Female	115	2.872	0.591
Race			
White	447	2.858	0.655
Minority	84	2.664	0.652
Learning Preferences			
Active	569	2.701	0.730
Reflective	270	2.830	0.687

Note[1]: Only demographic variables that resulted in significant differences are included in this table.
Note[2]: Course grade was converted into a 4-point scale analogous to GPA.

Table 6
Summary of Demographic and Background Variables Predicting Course Grade

Variable	Beta	Sig.	R-Square
Step 1			
Gender	0.066	0.206	
Race	0.129	0.009*	0.018
Step 2			
Active – Reflective	0.111	0.019*	
Sensing – Intuiting	0.026	0.590	
Visual – Verbal	0.033	0.487	
Sequential – Global	-0.018	0.699	0.036
Step 3			
Factor 1: Computer Knowledge	0.086	0.089	
Factor 2: Mechanical Knowledge	-0.069	0.184	
Factor 3: Prior Course Work	-0.072	0.134	
Factor 4: Time	-0.065	0.191	0.054

Note: *n* = 443; * Statistically significant

A second set of regression models was tested to examine how well programming knowledge and SAT scores predicted course grade. SAT scores were entered into block one followed by programming pre- and post-test scores. Results are shown in Table 7. It can be observed that SAT scores were significant predictors of EngE1024 grade. Programming scores also explained a significant amount of variance beyond SAT scores. This was primarily a function of post-test scores. Students with higher programming post-test scores were more likely to do better in the course than students with lower post-test scores. Taken together, SAT and programming scores explained approximately 24% of the variance in course grade (see Table 7).

Table 7
Summary of Quantitative Variables Predicting Course Grade

Variable	Beta	Sig.	R-Square
Step 1			
SAT	0.359	.000*	0.202
Step 2			
Programming Pre-Test Score	0.053	0.162	
Programming Post-Test Score	0.194	.000*	0.243

Note: n = 689; * Statistically significant

Engineering Retention

Additional analyses were conducted to examine what factors influenced engineering retention. That is, are there any specific background factors or prior collegiate experiences that related to engineering retention rates? Again, for the purpose of this analysis, a student is considered as "retained" in engineering if she/he was enrolled in the engineering program after the 4th semester of enrollment. Approximately 9% of the 1170 freshman students who completed EngE1024 in the fall 2004 semester were no longer considered engineering majors by the end of the spring 2006 semester. Due to the small number of students who switched majors (approximately 100) and the significant amount of missing data on the New Student Survey, discriminate function analysis or logistics regression could not be conducted. Therefore, chi-square analyses were conducted on many of the individual items on the New Student Survey, discussed below, that were hypothesized to influence retention.

The strongest background predictor of engineering retention after four semesters was experience repairing an electronic device other than a computer. Specifically, students who indicated in their New Student Survey that they had not repaired an electronic device other than a computer prior to college were more likely to drop out of the engineering program than students who had such experiences. Some prior experiences with computers also influenced retention rates. In particular, students who had no experience with DOS and did not know how to write computer programs were more likely to leave the engineering field than students who had worked with DOS and were able to write computer programs. Also, students who had worked with an engineer were less likely to drop out of engineering than students without such engineering-related work experience. A list of the background variables that resulted in significant differences can be found in Table 8. There were also three quantitative variables that predicted engineering retention. SAT scores, programming post-test scores, and EngE1024 course grades were significant predictors of retention. Specifically, students with higher SAT scores ($t\,[1128] = 2.84$, $p < .01$), programming post-test scores ($t\,[838] = 2.19$, $p < .01$) and EngE1024 course grades ($t\,[952] = 4.85$, $p < .01$) were more likely than students with lower SAT scores, programming post-test scores, and course grades to still be an engineering major.

Table 8
Background Variables Influencing Retention

Variables from New Student Survey	Pearson Chi-Square	Sig.
Knowledge of writing computer programs	4.227	0.049
Repairing electronic device other than a computer	12.569	0.001
Replaced a fuse or reset a breaker	6.749	0.014
Worked on engine or transmission	4.813	0.031
Worked for an engineer	4.217	0.040
Math Level: Calculus I or higher	9.768	0.006

Summary of Data Analysis

Course grade in the first engineering class was the strongest predictor of cumulative GPA for engineering students after four semesters. As can be recalled, course grade explained nearly 50% of the variance in cumulative GPA four semesters later. The regression analyses pertaining to course grade suggested that background factors and learning styles have only a minimal influence on success in an introductory engineering course. Programming post-test scores and SAT were much stronger predictors of course grade than any of the background or demographic variables. This information suggests that students can be successful in engineering despite a lack of prior relevant experiences (e.g., programming knowledge, engineering family member, developing a website, pre-engineering courses) before college.

The retention results reached a different conclusion regarding the influence of the background variables. Here, background experiences related to computers, mechanics, and prior course work (specifically math) did account for differences between students who decided to stay and leave engineering. In addition, quantitative variables also had a strong influence on engineering retention rates. Specifically, SAT scores, programming knowledge, and course grade in the introductory course all impacted student decisions to stay in engineering.

Exit Survey Results

Students in EngE1024 completed an exit survey at the end of the semester. Across the semesters, there was a significant difference in students' responses to the following question, "Have your problem-solving and logical-thinking skills improved as a result of this course?" Specifically, each semester (i.e., spring 2005 [n = 186]; fall 2005 [n = 894]; spring 2006 [n = 113]) of students indicated that they had improved in problem-solving and logical-thinking skills significantly more than the prior semester's students. This finding is supported by the fact that more in-workshop problem solving activities have been introduced in the course over the last 3–4 semesters.

Focus Group Results

As mentioned previously, focus groups (~1-hour session) were conducted at the end of two semesters (spring 2005 and spring 2006). Two random samples of approximately ten[3] students per semester were solicited to participate in a focus group concerning specific aspects of the course. Students were asked questions about hands-on experiences, teamwork, professional ethics, and their problem-solving and logical-thinking skills (specific questions can be seen in Table 9). Some interesting findings emerged from the focus groups.

Table 9
Focus Group Questions

Questions
What do you understand about the design process?
How will teamwork experience in this course affect your performance in team projects in future courses?
Was the hands-on experience important/valuable in learning the design process?
Has your perspective on professional ethics changed, and if so how?
How did your problem-solving/logical-thinking skills evolve as a result of this course?
[1]What have you learned in this class that you think will be useful?
[2]What is your understanding of sustainability?

Note[1]: This question was only asked in the spring 2005 semester.
Note[2]: This question was only asked in the spring 2006 semester.

Hands-on experiences

The students indicated that they would have preferred more hands-on experiences and projects in the class. It may be noted that hands-on activities in EngE1024 have increased significantly over the last two years (Lo et al., 2006). One of the hands-on activities involved building a 2-wheel robot as part of a mechatronics workshop (Lohani,

[3]Researchers (e.g., Ary, Jacobs, & Razavieh, 1996) recommend focus groups contain between five and ten people.

Kachroo, et al., 2006). While students enjoyed this hands-on activity—as was evident by their responses to in-class clicker based assessment questions (Lohani, Kachroo, et al., 2006)—students made interesting observations in the focus group sessions. For example, students commented that they wanted more freedom and creativity in the mechatronics hands-on project. The investigators currently are developing an improved version of the mechatronics workshop.

Professional ethics

A variety of measures were employed to cover engineering ethics instruction in EngE1024. Some examples are assigned reading of engineering case studies followed by in-class discussion, in-class presentations on contemporary ethics-related issues, and use of an ethics video similar to *Incident at Morales*. Based on student responses during focus group sessions, students did not perceive a difference between personal and professional ethics. Nearly all of the students said that ethics couldn't be taught in the classroom. They commented that, by their age, their morals are already formed. They also said that most people know when they are doing something wrong. When queried as to how the ethics instruction could be improved, the students suggested a series of debates on ambiguous cases. This format would provide legal information and lend itself to a discussion of the consequences of unethical behavior. A pre- and post-test for assessing the effectiveness of ethics instruction has been designed as a result of these findings from focus group sessions.

Teamwork

Overall, the students said they had a positive experience working in teams throughout the semester. Many students liked that their teams were chosen at random, as this gave them the chance to meet new people. Additionally, many students said they did not learn anything new regarding working in teams. As with ethics, the students felt as though they had already learned everything about teamwork. More activities related to teamwork are being introduced in the course, and formal instruction to improve teamwork skills are planned also.

Problem-solving and logical-thinking skills

The students reported in the focus groups that their problem-solving and logical-thinking skills did not evolve as a result of this course. Most students said this was because they were never presented with a problem and instructed to figure out a solution. This finding appears contradictory to exit survey results. More in-class and in-workshop problems are being introduced to further emphasize the key concepts of engineering problem solving. In addition, with the introduction of a Tablet PC in the course in fall 2006, collaborative features of Tablet software like Microsoft OneNote are being explored to emphasize in-workshop collaborative problem solving and design work (Lohani, Castles, et al., 2007).

Programming experience

Based on the data collected as part of the New Student Survey, about half of the engineering freshmen have a programming background. The Alice programming environment is primarily targeted to students with no prior programming experience.

Students, particularly those with prior programming experience, didn't enjoy the programming environment Alice offers for teaching basic concepts of object-oriented programming. Those students with Java or C++ programming experience, therefore, did not appreciate the drag and drop approach for developing a program in the Alice environment and didn't consider Alice a "true" programming software. The authors, however, note that several first-time programmers admitted liking the programming approach in the Alice environment, and pre-and post-test results showed positive learning gains (see Table 1). Although students did a number of programming exercises on engineering applications using Alice (e.g., simulation of a sine wave, simulation of a circular motion of an object, simulation of a motion of a pendulum), several students did not foresee direct engineering applications of Alice in future engineering courses. The authors are not aware of any upper-level courses that directly use Alice programming language, although use of object-oriented programming concepts occurs in several courses.

EngE1024 Course Changes and Improvements

Based on the results from the assessment activities (i.e., focus groups, exit surveys) the following specific changes/improvements have been made in course content and delivery style.

Hands-on Activities

A number of hands-on activities reflecting various engineering disciplines have been introduced. For example, a mechatronics workshop involving introductory engineering concepts from mechanical, electrical, computer engineering, and science disciplines was included. In addition, engineering faculty from various degree-granting departments are given opportunity to make a brief presentation (5–10 minutes) about their departments in an EngE1024 lecture. PowerPoint slides describing specific research, teaching, study abroad opportunities, and so forth, in various engineering departments are shared with students. These changes address students' consistent comments expressed in EngE1024 exits surveys and focus groups to learn more about engineering departments.

Programming

Alice programming, although successful in conveying preliminary object-oriented concepts as judged by the learning gains results presented in Table 1, has been replaced by LabVIEW (www.NI.com) programming beginning in spring 2007. LabVIEW (**Lab**oratory **V**irtual **I**nstrumentation **E**ngineering **W**orkbench) uses a graphical programming language, known as G programming language, to create programs relying on graphical symbols to describe a problem-solving activity or programming action (Bishop, 2007). LabVIEW provides a powerful interface for collecting and processing data from engineering experiments. The latest version of LabVIEW (i.e., version 8.2) includes object-oriented programming features. In addition, the authors are aware of upper-level courses that utilize LabVIEW software for engineering problem solving. MATLAB wasn't considered for EngE1024, since it is now covered in a spring semester engineering course.

Problem-solving activities

Additional problem-solving activities have been introduced in the workshop section of EngE1024 in response to students' requests made as part of focus group and exit survey results.

Ethics

A variety of activities including ethics videos, in-class skits on realistic ethics case studies, use of actual ethics case studies from sources like the monthly "Newsletter of the American Society of Civil Engineers (ASCE)" for in-class discussion, etc., have been used to cover professional engineering ethics in the course. A pre- and post-test has been developed to assess learning gains.

Hands-on Experiments

Students used to perform book problems for analyzing experimental data and fitting empirical functions. Students now participate in hands-on experiments (e.g., water tower experiment, world population activity) to generate the data before analyzing them.

Design Projects

In order to emphasize societal importance of engineering activities, freshman design projects now focus on analyzing sustainability-related issues in the world. A sustainable energy design project has been implemented to expose engineering freshmen to contemporary engineering challenges in the world (Mullin, Jinsoo, Lohani, & Lo, 2007). In the past, design projects in this course have focused on design of cars, boats, ping-pong launchers, etc. In addition, students are introduced to contemporary engineering issues by discussing recent publications like *The Engineer of 2020* (National Academy of Engineering, 2004) and *The World is Flat* (Friedman, 2005).

Study Abroad

Lastly, study abroad presentations have been introduced in the course starting in spring 2005 to bring global engineering education issues into the classroom. For example, in fall 2005, five engineering seniors with study abroad experiences in various countries including South Africa, Ireland, New Zealand, France, and Spain were invited to share their experiences with all 1200 engineering freshmen in EngE1024. Their short presentations (~30 minutes) were very informative and motivating, as indicated by in-class clicker-based questions/answers. These clicker-based responses are considered in conjunction with results from other assessment instruments including prior experiences, focus group results, exit surveys, learning styles for implementing curriculum changes.

Future Plans

EngE1024 has become a research- and assessment-driven course. A number of scholarly publications have been generated based on various innovative activities implemented in the course. The course is providing an excellent venue for COE graduate students to gain classroom teaching experience. About 30 graduate students, mostly Ph.D. students, have been involved in conducting EngE1024 workshops between fall

2004 and fall 2006. A Ph.D. student is currently assessing creativity in freshman design projects using EngE1024 data as part of her dissertation. Another Ph.D. student will use fundamental engineering concepts covered in this course to develop data structures for formal knowledge representation. A nanotechnology module is proposed to be introduced into the course during year 2007–2008.

Acknowledgements

The authors would like to acknowledge the support of all DLR investigators, EngE1024 faculty, and workshop instructors for their help in developing and implementing various assessment measures. We would like to sincerely thank Dr. Hayden Griffin for his strong support in implementing various changes in EngE1024. We are also grateful to Dr. Jenny Lo and Ms. Jennifer Mullin who provided immense help in implementing various changes in the course. We would also like to sincerely acknowledge the funding support provided by the National Science Foundation (DLR grant # 0431779). The opinions, findings, and conclusions or recommendations expressed in this chapter are those of the authors and do not necessarily reflect the views of the National Science Foundation. Lastly, we would like to dedicate this chapter to our dear DLR project investigator, Dr. G. V. Loganathan, who was one of the victims of the horrific crime that took place on the VT campus on April 16, 2007. Dr. Loganathan, an outstanding engineering educator, was one of the first faculty members who joined the now strong engineering-education collaborative at VT and made significant contribution to the success of the BEEVT and DLR projects. The engineering education community, in general, and our collaborative, in particular, will miss him greatly.

References

Ary, D., Jacobs, L. C., & Razavieh, A. (1996). *Introduction to research in education*. Fort Worth, TX: Harcourt Brace College Publishers.

Bishop, R. H. (2007). *LabVIEW 8 student edition* [Computer software]. Upper Saddle River, NJ: Pearson Prentice Hall.

Bruner, J. (1960). *The process of education*. Cambridge, MA: Harvard University Press.

Budny, D., LeBold, W., & Bjedov, G. (1998). Assessment of the impact of freshman engineering courses. *Journal of Engineering Education*, *87*(4), 405–411.

Connor, J., Lohani, V. K., Mallikarjunan, K., Loganathan, G. V., & Lo, J. L. (2006, June). *Multiple models of a freshman engineering experiment*. Proceedings of the ASEE Annual Conference, Chicago, IL.

Felder, R. M., & Brent, R. (2005). Understanding student differences. *Journal of Engineering Education, 94*(1), 57–72.

Felder, R. M., & Solomon, B. A. *Index of learning styles*. Retrieved April 2, 2008, from http://www.ncsu.edu/felder-public/ILSpage.html

Felder, R. M., & Spurlin, J. E. (2005). Applications, reliability, and validity of the Index of Learning Styles. *International Journal of Engineering Education, 21*(1), 103–112.

Friedman, T. L. (2005). *The world is flat: A brief history of the 21st century*. New York: Picador.

Heywood, J. (2005). *Engineering education research and development in curriculum and instruction*. Piscataway, NJ: IEEE Press.

Journal of Engineering Education. (2006). The research agenda for the new discipline of engineering education. *Journal of Engineering Education*, *95*(4), 259–262.

Knott, M., Lohani, V. K., Griffin, O.H., Jr., Loganathan, G. V., Adel, G., Paretti, M., et al. (2005, June). *Using e-portfolios in a large engineering program.* Proceedings of the ASEE Annual Conference, Portland, OR.

Knott, T. W., Lohani, V. K., Griffin, O.H., Jr., Loganathan, G. V., Adel, G., & Wildman, T. (2004, June). *Bridges for engineering education: Exploring e-portfolios in engineering education at Virginia Tech*. Proceedings of the ASEE Annual Conference, Salt Lake City, UT.

Lo, J. L., Lohani, V. K. & Griffin, O. H., Jr., (2006, June). *Full implementation of a new format for freshmen engineering course.* Proceedings of the ASEE Annual Conference, Chicago, IL.

Lohani, V. K., Castles, R., Lo, J. L., & Griffin, O. H., Jr., (2007, June). *Tablet PC applications in a large engineering program*. Proceedings of the ASEE Annual Conference, Honolulu, HI.

Lohani, V. K., Kachroo, P., Chandrachood, S., Zephirin, T. Loganathan, G. V., & Lo, J. L. (2006, July). *Mechatronics experiment in a freshman year course.* Proceedings of the International Conference on Engineering Education (iCEE06), San Juan, PR.

Lohani, V. K., & Lo, J. L. (2006, May 4). *Clicker-based assessment in an engineering course.* Virginia Tech Center for Excellence in Undergraduate Teaching Year-End Poster Session, Blacksburg, VA.

Lohani, V. K., Mullin, J. S., Lo, J. L., & Griffin, O. H., Jr. (2006, October). *Implementation of international activities in a freshman engineering course.* Proceedings of the ASEE Global Colloquium on Engineering Education, Rio de Janeiro, Brazil.

Lohani, V. K., Sanders, M., Wildman, T., Connor, J., Mallikarjunan, K., Dillaha, T., et al. (2005, June). *From BEEVT to DLR - NSF supported engineering education projects at Virginia Tech.* Proceedings of the ASEE Annual Conference, Portland, OR.

Lohani, V. K., Wildman, T., Connor, J., Mallikarjunan, K., Wolfe, M. L., Muffo, J. A., et al. (2005, October). *Work in progress: Spiral curriculum approach to reformulate engineering curriculum.* Proceedings of the Annual Frontiers in Education Conference, Indianapolis, IN.

Mallikarjunan, K., Laksmikanth, A., Wolfe, M. L., Lohani, V. K., & Connor, J. (2007, September). *Introduction to systems approach using bio-energy resources as a tool for freshman engineering education.* Proceedings of the International Conference on Engineering Education (iCEE07), Coimbra, Portugal.

Muffo, J. A., Lohani, V. K., Mullin, J. S., Backert, R., & Griffin, O. H., Jr. (2005, March). *From engineering fundamentals to engineering education—What's in a name?* Proceedings of the International Conference on Engineering Education & Research (iCEER05), Tainan, Taiwan.

Mullin, J. S., Jinsoo, K., Lohani, V. K., & Lo, J. L., (2007, June). *Sustainable energy design projects for engineering freshman*. Proceedings of the ASEE Annual Conference, Honolulu, HI.

Mullin, J. S., Lohani, V. K., & Lo, J. L. (2006, October). *WIP: Introduction to engineering ethics through student skits in the freshman engineering program at Virginia Tech*. Proceedings of the Annual Frontiers in Education Conference, San Diego, CA.

Mullin, J. S., Lohani, V. K., Lo, J. L., & Griffin Jr., O.H. (2006, June). *Sustainable development design projects for engineering freshmen*. Proceedings of the ASEE Annual Conference, Chicago, IL.

National Academy of Engineering. (2004). *The engineer of 2020: Visions of engineering in the new century*. Retrieved April 2, 2008, from http://books.nap.edu/catalog.php?record_id=10999#toc

Pedhauzer, E. J. (1997). *Multiple-regression in behavioral research*. Belmont, CA: Wadsworth.

Robson, V., Lohani, V. K. & Bateman, T. (2007, June). *Foundational predictors of success in the collegiate engineering program*. Proceedings of the ASEE Annual Conference, Honolulu, HI.

Robson, V., Muffo, J. A., & Lohani, V. K. (2006, May). *The little engineer that could: Background, attitudinal, and academic predictors of student success*. Poster session presented at the Association for Institutional Research Annual Forum, Chicago, IL.

Snook, J., Lohani, V. K., Lo, J. L., Sirvole, K., Mullins, J. S., Kaeli, J., et al. (2006). Incorporation of a 3-D interactive graphics programming language into an introductory engineering course. *Computers in Education Journal*, *16*(3), 63–72.

CHAPTER 11

EMPLOYER ASSESSMENTS OF EC2000 GRADUATES AND THE ABET CRITERION 3 (A–K) OUTCOMES[1]

J. Fredericks Volkwein
Lisa R. Lattuca
Patrick T. Terenzini
Center for the Study of Higher Education
Pennsylvania State University

Introduction

In 1996, the ABET Board of Directors approved a new set of accreditation standards, called *Engineering Criteria 2000: Criteria for Accrediting Programs in Engineering in the United States* (ABET, 1997). Known to the engineering education community simply as "EC2000," the new criteria shifted the basis for accreditation toward student learning outcomes and away from resources, curricular requirements, faculty, and facilities. The new criteria specify 11 learning outcomes and require engineering programs to assess and demonstrate their students' achievement in each of those areas. EC2000 retains earlier accreditation standards' emphasis on the development of students' mathematical, scientific, and technical knowledge, but also emphasizes other professional skills, such as solving unstructured problems, communicating effectively, working in teams, and developing ethical and contextual considerations in engineering (see Table 1).

In 2001, ABET concluded that, in the spirit of continuous improvement that it encourages in programs and schools, it should ascertain the impact of the EC2000 criteria on the preparation of engineering graduates. The agency issued a request for proposals for an evaluation to assess student learning outcomes, and employer satisfaction with graduates' preparation. The Center for the Study of Higher Education at Penn State University (University Park) is among those who responded with proposals, and in 2002, ABET commissioned the Center's team to undertake the study.

Thus, in 2002, with funds from ABET and NSF, the Center at Penn State began a four-year study to assess whether the new EC2000 learning criteria are having the intended effects. To find the answers, our research team collected data from graduates, faculty members, employers, deans and program chairs in over 200 engineering programs and seven engineering fields on 40 campuses. This chapter discusses and describes the employer component of this research, so that other assessments of engineering

[1]This study was supported by a grant from the Accrediting Board of Engineering and Engineering Technology (ABET). The opinions expressed here do not necessarily reflect the opinions or policies of ABET, and no official endorsement should be inferred.

outcomes might benefit from our work. A copy of the complete EC2000 study may be purchased from ABET, Inc., and an Executive Summary is available at http://www.abet.org/papers.shtml.

Table 1

EC2000 Criterion 3 Learning Outcomes

a. An ability to apply knowledge of mathematics, science, and engineering
b. An ability to design and conduct experiments, as well as to analyze and interpret data
c. An ability to design a system, component, or process to meet desired needs
d. An ability to function on multi-disciplinary teams
e. An ability to identify, formulate, and solve engineering problems
f. An understanding of professional and ethical responsibility
g. An ability to communicate effectively
h. The broad education necessary to understand the impact of engineering solutions in a global and societal context
i. A recognition of the need for, and an ability to engage in life-long learning
j. A knowledge of contemporary issues
k. An ability to use the techniques, skills, and modern engineering tools necessary for engineering practice
Source: ABET (1997). *Engineering Criteria 2000*.

Evaluation Framework

To provide a data-based evaluation of the status and impact of its reforms, ABET engaged the Center for the Study of Higher Education at Penn State University to answer the question: "Are engineers who graduated from programs since implementation of the EC2000 standards better prepared for careers in engineering than their counterparts who graduated before introduction of the criteria?" Our research team designed a three-year evaluation, entitled *Engineering Change: A Study of the Impact of EC2000*. The study targets 203 programs in seven engineering disciplines at 40 institutions carefully selected for their representativeness. We selected four engineering fields (chemical, civil, electrical, and mechanical engineering) that produce the vast majority of engineering graduates in any one year, and another three (aerospace, computer, and industrial engineering) were selected for their strong ties to specific employment sectors.

The framework for this study assumes that if EC2000 has been effective, evidence of change in ABET-accredited programs will be linked to changes both in engineering programs and in student outcomes. The conceptual framework for this project, shown in Figure 1, posits that the altered EC2000 accreditation standards influence curricular modifications, instructional practices, institutional policies, program reorganization, faculty development activity, and related faculty culture. These changes in engineering education influence student experiences both inside and outside the classroom, and these subsequently impact student outcomes. Thus, the EC2000 processes and criteria have a direct impact on the organizational changes that, in turn, impact student experiences and learning outcomes. These outcomes are visible not only to faculty, but also to employers. This framework posits that employer feedback, as well as other information about student learning outcomes, provides the basis for further improvements in curriculum and instruction, as well as in educational and organizational policies and practices. Thus, employer assessments offer a crucial continuous improvement component to the framework.

The assumptions that are embedded in our conceptual framework are generally supported by the research literature. Several studies have already documented industry and EC2000 impact on faculty and curricula—institutions are developing educational goals and objectives, measurable learning outcomes, and outcomes assessment processes. Moreover, quality assurance, quality control, and improvement is common practice in every engineering field; see Volkwein, Lattuca, Terenzini, Strauss, and Sukhbaatar (2004) for some references to this literature.

Engineering Change:
Studying the Impact of EC2000

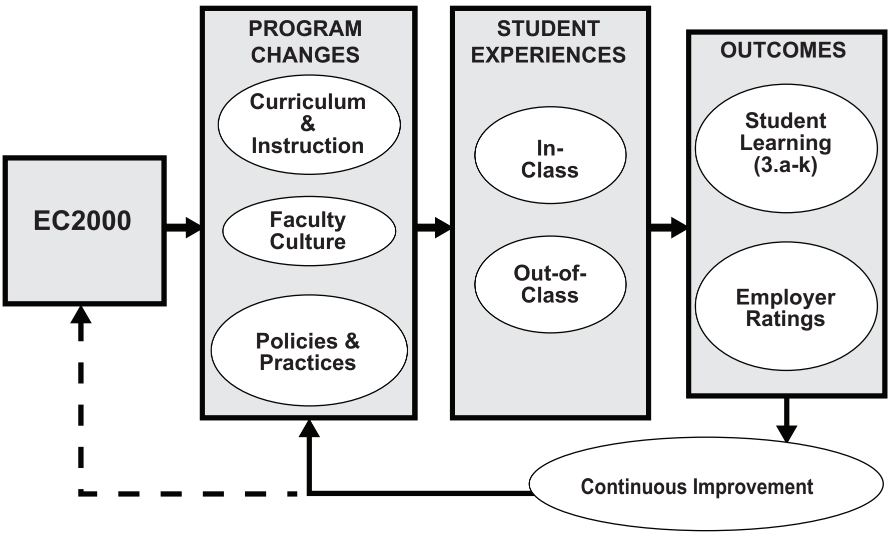

Figure 1. Conceptual framework.

The Purpose of the Employer Component of the EC2000 Study

The design of our EC2000 study assesses engineering student learning outcomes from multiple lenses, with the employer survey constituting an important ingredient. Employer surveys have long been considered an important assessment component, especially in fields like engineering (Hoey & Gardner, 1996). While preparing for the study, the research team discussed the purpose and proposed methodology with several engineering stakeholder groups who suggested that we examine not only employers' perceptions about the adequacy of engineering preparation, but also their judgments about the importance of the 11 Criterion 3 (a–k) outcomes. Engineering programs may be meeting EC2000 objectives articulated by ABET, but if a disconnect exists between these objectives and the outcomes needed by employers of entry-level engineers, then all of these educational efforts may still fail to produce a workforce ready to meet the challenges of engineering practice in the real world.

While the main objective of the Engineering Change study is to assess the pre- and post-EC2000 student learning Criterion 3 (a–k), the employer ratings of importance of these outcomes have been infrequently assessed since the formative years of EC2000 in the early 1990s. When assessments have occurred, they have been limited to either a single institution or a single discipline, like mechanical engineering.

Several studies within institutions assess employer perceptions of student ability and the importance of skills consistent with the (a–k) outcomes. Hoey, Marr, and Gardner (2002) surveyed on-campus recruiters and cooperative education supervisors. With a relatively large number of respondents (1,022) they asked employers to rate student ability and the importance of skills consistent with the (a–k) outcomes. Employer ratings of student knowledge of contemporary issues were higher than ratings of importance for contemporary issues. This study also reports a match between ratings of importance and student ability to design a system, life-long learning, and conducting experiments; however, employers noted students were not performing as well on the remainder of the Criterion 3 (a–k) outcomes. This is an excellent study, but unfortunately limited to one campus.

Koen and Kohli (1998) ask similar questions, but used a slightly different methodology, surveying employers of recent graduates for three cohorts of students. They found that the most important attributes for new hires were problem solving, ability to design and conduct experiments, life-long learning, professional ethics, and teamwork. While this is another example of a sound study and very useful to the specific institution, using students from only one institution and employers of students at one institution limits the generalizability of the results.

A second type of study are those conducted by the engineering societies. The American Society of Mechanical Engineers (ASME) conducted a study of 66 individuals at 33 companies who responded to their survey (ASME/NSF, 1995). ASME also surveyed all 243 accredited mechanical engineering programs, of which 42 responded. The study found that the five most important skills for graduating mechanical engineers are teamwork, communication, design for manufacture, CAD systems, and professional ethics. Employers not only rated recent graduates on these skills, but also identified specific areas for improvement. Another strong study reviewed for this paper is the SPINE Benchmarking Study (Bodmer, Leu, Mira, & Rutter, 2002), which includes multiple employers and multiple institutions across all engineering areas. While the primary purpose of this study is to compare U.S. institutions to European institutions, it also includes employer feedback from 1,372 engineers and 14 managers. These employers rated practical engineering experience and widely applicable skills (problem solving, analysis) as most important. These employers were not asked to rate employees on the adequacy of their preparation.

While many of these studies are useful to particular institutions or engineering societies, few provide a comprehensive assessment of employer perceptions of student engineering skills. Because the objective of our Engineering Change study is to determine the preparation of new engineering graduates across the nation in seven subfields, we designed our research to reach a national sample of employers who are familiar not only with the qualifications of recent engineering graduates, but also with those in the pre-EC2000 1990s.

Research Methods and Data Collection Procedures

This section summarizes the methods we used at Penn State to gather information about the pre-EC2000 and post-EC2000 changes that employers may have observed. The employer study design, like that of all others in this study, was an *ex post facto,* cross-sectional survey that we developed in consultation with both local and national employers and advisory bodies.

Development of the employer questionnaire followed the same procedures as those used to construct the other survey instruments: literature reviews; item-bank development; design of the content areas of the survey; and an extended process of revising, augmenting, and reviewing potential items that were then pilot tested. After conversations with the six engineering societies that endorsed the study and with members of ABET's Industry Advisory Council, a draft of the survey instrument was shared with the EC2000 study National Advisory Board and with selected engineering employers in focus groups at one meeting with local employers and at two national engineering conferences.

The final survey instrument gathers information on respondent's personal characteristics and then addresses three basic issues: (a) the adequacy of preparation among recent engineering graduates; (b) whether graduates' abilities have increased, stayed the same, or decreased since the mid-1990s; and (c) the importance employers attach to each of the 11 EC2000 learning outcomes (Penn State, 2006).

For the first two questions, the 11 EC2000 Criterion 3 (a–k) outcomes were collapsed *a priori* into five clusters based upon discussions with an employer focus group and a preliminary factor analysis of items that were pilot tested on senior engineering students. The five dimensions are:

1) using engineering, math, science, and technical skills;
2) applying problem-solving skills;
3) communicating and working in teams;
4) understanding the organizational, cultural, and environmental contexts and constraints of engineering practice, design, and research, and
5) continuing to learn, grow, and adapt as technology and social conditions evolve in unpredictable directions.

Employers responded to the following stem: "Think about the field [with whose graduates you are most familiar]. In this field, how prepared are recent graduates?" Employers used a three-point scale, where 1 = *inadequately prepared,* 2 = *adequately prepared,* and 3 = *well prepared*. Respondents also had a *don't know* option. Employers evaluated the extent to which "the abilities of graduates [have] increased or decreased compared to engineering graduates of 7–10 years ago." Respondents used on a three-point scale, where 1 = *decreased,* 2 = *about the same,* and 3 = *increased*. Again, respondents could choose *don't know* when appropriate.

The population of employers was defined to include individuals who had been: (a) involved in hiring or evaluating bachelor's degree-level engineering graduates from any of the seven target disciplines, and (b) involved in that process for seven or more years. To reach a mix of employers, the six professional societies associated with the targeted disciplines and that had endorsed the study (AIAA, AIChE, ASCE, ASME, IEEE, IIE) sent a broadcast e-mail to their members inviting those who met the population definition criteria to participate in the study. Engineering society staff assisted in identifying their members whose job titles suggested that they might be involved in the evaluation of new engineers. The broadcast e-mail contained a URL to which they could go to complete the survey online. In addition, the end of the 1994 graduate survey contained an invitation to respondents meeting the study's employer population specifications to participate and provided an appropriate URL to complete the survey online. Thus, we used multiple strategies to reach the national population of employers.

The data collection strategies yielded responses from over 2,065 employers, 1,622 of whom met the twin criteria for inclusion in our analysis and who responded to a minimum of 80 percent of the survey items. Because no information is available on the characteristics of the employer population as defined for this study, no assessment is possible of the representativeness of respondents.

Variable and Scale Development

Based on advice from employer focus group members and the study's National Advisory Board, we concluded that few employers would take the time to complete a lengthy questionnaire. Consequently, we put a premium on brevity in the employer survey and favored single-item questions over sets of items that could later be aggregated into scales—as were used in the graduate, faculty, and program chair surveys.(A small pilot test indicated the final instrument took 7–10 minutes to complete.) The survey instrument is available at http://www.ed.psu.edu/cshe/abet/instruments.html

Only one scale was created. To differentiate small local employers from large national ones, we combined three variables into one that combines company size (number employees in 5 range categories, where 1 = *less than 50* and 5 = *more than 10,000*), span of employee recruitment (1 = *local,* 2 = *regional,* and 3 = *national*), and average number of new engineers evaluated each year (1 = *1 or less,* 2 = *2–5*, 3 = *6–10,* and 4 = *more than 10*). Responses were summed, and employers with less than 5 points on this scale were characterized as *local* and those with 8 or more points are characterized as *national.*

Results

The employer component of our Engineering Change study collects and summarizes the judgments of 1,622 employers who have been involved in the evaluation of new engineers since the mid-1990s or earlier. The paper and web survey collected information on respondent demographics, their ratings of recent engineering graduates, a comparison of these recent graduates to pre-EC2000 graduates, and a rating of the importance of each Criterion 3 (a–k) student learning outcome.

These employer respondents are highly diverse in their geographic locations, industry type, company size, educational attainment, and experience in evaluating engineers. For example, we received responses from every state and U.S. territory and from every industry sector of the Engineering Workforce Commission classification that we used. Over half the employer respondents (54%) have earned degrees beyond the bachelor's, and 12 percent have doctoral or first professional degrees. The vast majority report that they are either senior level engineers or mid-level managers, and 60 percent of them indicated that they have *both* management and practicing engineer responsibilities in their organizations. Approximately half of the respondents work in companies engaged in manufacturing or providing scientific and technical services. Moreover, the respondents represent a good range of company sizes:

Less than 50 employees	25%
50–499 employees	39%
500–3000 employees	24%
Over 3000 employees	12%

The engineering society and alumni survey invitations yielded a healthy number of employer responses in each discipline. Four of the fields are roughly proportional to their numbers in the workforce (aerospace, chemical, computer, and industrial), but the civil engineering employers are over-represented, while the electrical and mechanical are under-represented. This may be of less concern because we found few differences in employer perceptions by engineering field.

Fields	**N**	**%**
Aerospace Engineering	42	3
Chemical Engineering	267	16
Civil Engineering	661	41
Computer Engineering	135	8
Electrical Engineering	205	13
Industrial Engineering	118	7
Mechanical Engineering	194	12
Total	**1622**	**100.0**

The majority of respondents (54%) evaluate 2–5 new engineers per year, and 30 percent evaluate more. Over half of these respondents (55%) recruit their new engineers at a handful of engineering schools (3–5) within their state or region, but about one-third of them recruit nationally. The rest recruit locally from one or two engineering schools.

In the survey we first asked the responding employers to rate the preparation of recent engineering bachelor's degree graduates, then to indicate whether this level of preparation has increased or decreased since the mid 1990s, and, finally, to rate the importance of each of the 11 Criterion 3 (a–k) abilities and skills. Their responses are summarized in Figures 2–7.

The bar graph in Figure 2 shows the employer ratings of engineer preparation for each of the five clusters of skills and abilities emphasized in EC2000. In their assessments of new engineers, over 90% of employers rate them as adequately prepared or well prepared to use math, science and technical skills, and about 8 out of 10 employers give them passing marks for problem solving and for learning, growing, and adapting. Teamwork and communication skills are assessed as at least adequate by 3 out of 4 employers. On the other hand, barely half of the employers give an adequate rating to new engineers' understanding of the organizational, cultural, and environmental contexts and constraints of their work. This appears to relate most clearly to EC2000 Criterion 3(h) and 3(j).

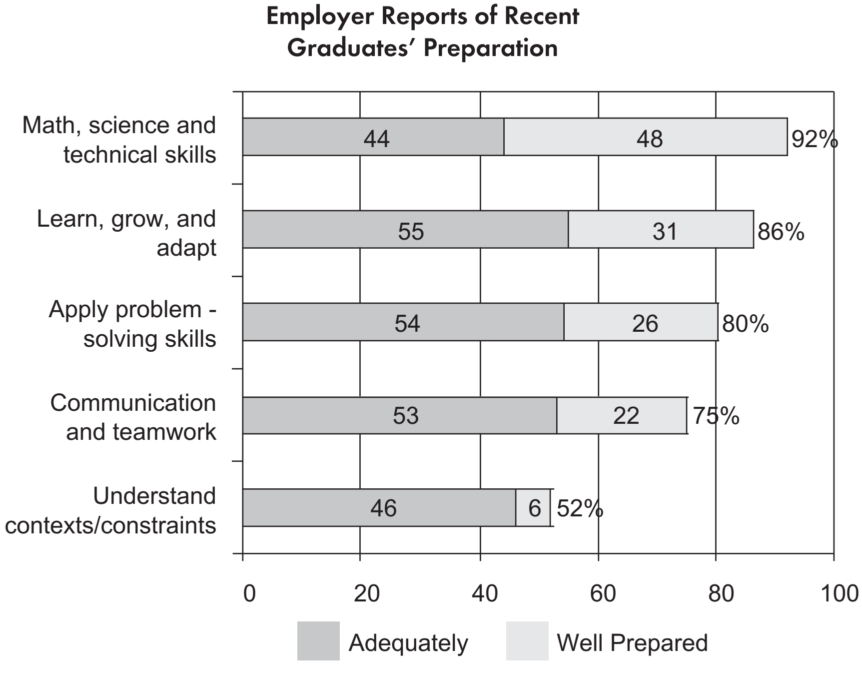

Figure 2. Employer evaluations of preparation of recent graduates.

Figure 3 shows the distribution of responses on the increase/decrease scale. Since the introduction of EC2000, communication skills and teamwork appear to have improved modestly, along with learning to grow and adapt to changing technology and society. Technical skills in math and science appear unchanged in the past decade, but problem-solving skills and understanding cultural and environmental contexts and constraints appear to have declined modestly. Not only is understanding contexts and constraints the lowest rated skill by employers, but this appears to be an area of some decrease over the past decade. Although the erosion in problem solving is a concern, 80 percent of employers still rate students highly in this area.

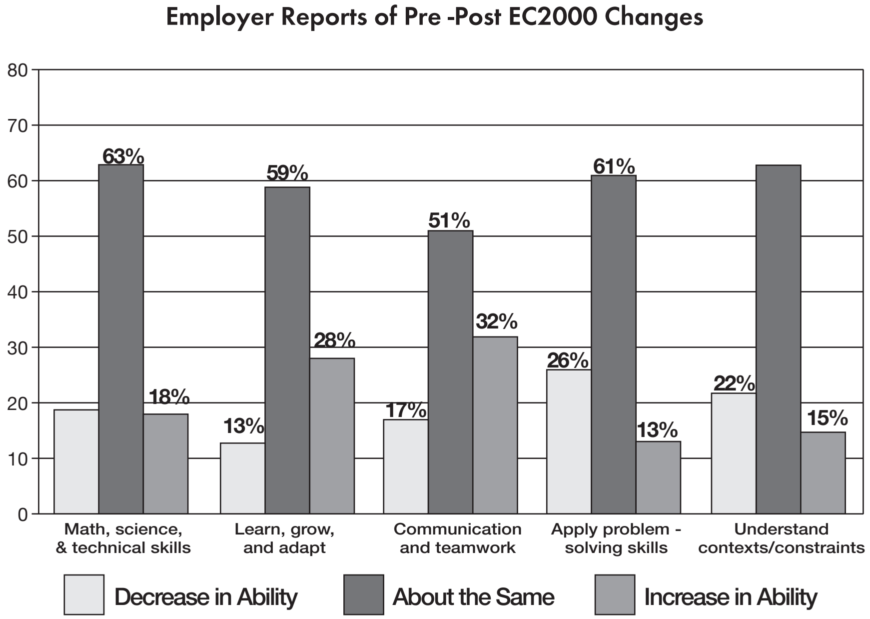

Figure 3. Employer reports of change.

Finally, we asked the employers to rate the importance of each of the 11 (a–k) items on a 5-point scale, and the results are displayed in Figure 4. The bars in Figure 4 show the percents of all employers rating each (a–k) skill as moderately important or highly important/essential (3, 4, 5 on the scale). Clearly, the (a–k) list as a whole is substantially in sync with the qualities that employers value in new engineering bachelor's degree recipients. Effective communication, engineering problem solving skills, math and science application, using engineering tools, teamwork, and professional ethics rate as the most universally important engineering skills. In fact, all but two (a–k) criteria are rated as at least moderately important by more than 85 out of 100 employers. The two rated lower are knowledge of contemporary issues and understanding the impact of engineering solutions in a global and societal context. These (along with life-long learning) are the most liberal education oriented of the (a–k) items and the least dependent on the engineering curriculum. Even so, more than 7 out of 10 employers rate them as at least moderately important job qualities among new hires.

Employers' Ratings of Importance:
a-k Criteria for New Hires

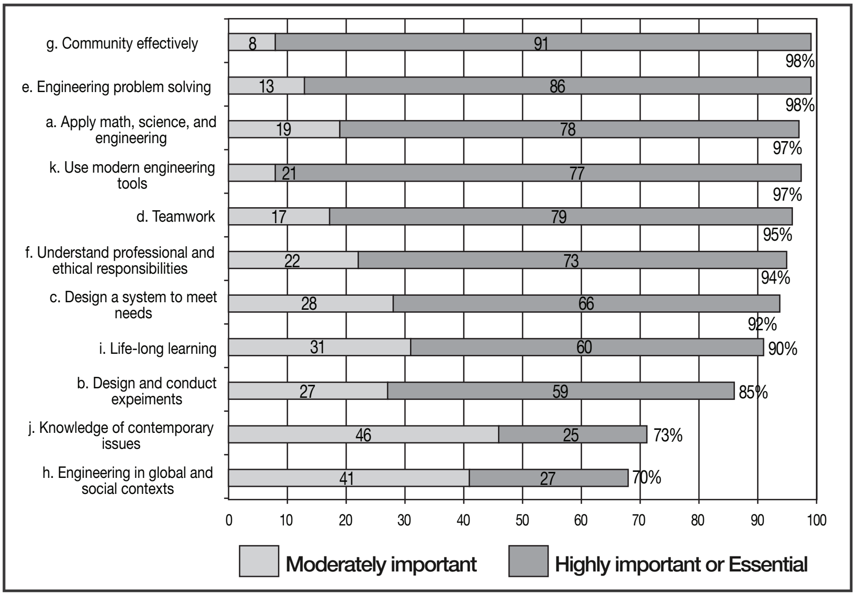

Figure 4. Employer evaluation of relative importance of Outcomes (a–k).

We also encouraged employers to indicate other important skills that may not be represented in Criterion 3 (a–k), and we received entries from over 600 respondents, or about 37% of the total. Our team reviewed the entries and judged 60% of them to be substantially like the existing (a–k) categories, perhaps phrased with greater emphasis or sharper language. However, there are 239 "other" abilities and skills that we judged to fall outside the general (a–k) concepts, so we aggregated them into separate clusters as follows:

Leadership/Management skills = 49 entries
Accounting/financial/business skills = 33 entries
Social/political skills = 17 entries
Creativity = 15 entries
Organizational loyalty/devotion = 12 entries
Co-ops/internships/hands-on experiences = 11 entries
Common sense = 10 entries
Miscellaneous others = 93 entries

Since ABET accreditation is program-specific, we wanted to see if employer ratings differ significantly by engineering field. Despite the geographic and company diversity and variable backgrounds and experiences of the employer respondents, we found relatively high agreement about the preparation of new engineers, about the amount of change in engineer preparation since the mid-1990s, and about the importance of the (a–k) abilities and skills, even across engineering disciplines. These results are displayed in the Appendix.

Finding few differences by engineering field, we also examined the other employer characteristics to see if there are strong associations between employer profiles and employer ratings. We found remarkably consistent patterns across industry type, geographic location, organizational level, highest degree, as well as engineering field. However, we did find that the large national employers, who recruit and evaluate the greatest numbers of engineers, rate current engineering graduates in the aggregate more highly and also report more improvement since the mid-1990s. Figure 5 compares the average ratings of those employers reporting that they are primarily local in orientation with those reporting that they are primarily national in orientation. While the differences are not dramatic, they are statistically significant. The national employers give higher rating both to current preparation and to improvement over time. The importance of this finding is that the national employers recruit and evaluate the most engineers. The majority of the employers classified as local report that they evaluate only one engineer per year, and none evaluate as many as six. On the other hand, 65% of the national employers report that they evaluate six or more engineers per year. Thus, the national employers clearly recruit and evaluate the greatest number of new engineers each year; hence, their assessment of EC2000 graduates suggests a more positive impact than we otherwise might conclude.

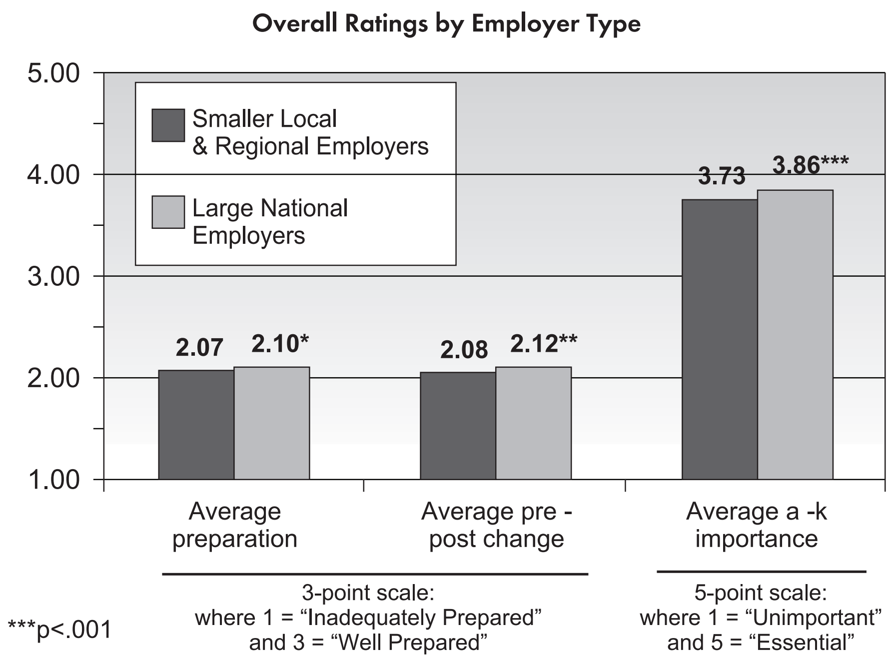

Figure 5. Employer overall rating.
Note: * * * $p < .001$

Comparing Employer Responses with Faculty and Chairs

For the five areas of engineering preparation, we compared employer and faculty perceptions of recent graduates, and the results are summarized in Figure 6. In each of the five areas, faculty see more increase over time than do employers, and the differences are statistically significant. The gap between employer and faculty perceptions is least for math, science and technical skills, where faculty see a deterioration that employers do not see. The greatest gap between the two perceptions exists for teamwork and communication skills. More than twice as many faculty (68%) perceive an increase in ability, compared to 32% of employers. Similar divergence in the perceptions of employers and faculty are seen in problem-solving skills (13% of employers see an increase versus 27% of faculty) and in understanding organizational and environmental contexts (15% of employers see an increase versus 43% of faculty). There are several possible explanations for these employer-versus-faculty gaps in perception. Although the faculty may be seeing students through rose-colored glasses, we know that there is a lag of a couple of years between when faculty have students in class and when employers evaluate them after a year or so on the job. Moreover, when we collected our data in 2004, EC2000 was still not fully implemented. In addition, it may also be possible that the faculty see the top performing graduate-school-bound students, and that the workforce-bound students are less impacted by the EC2000 changes.

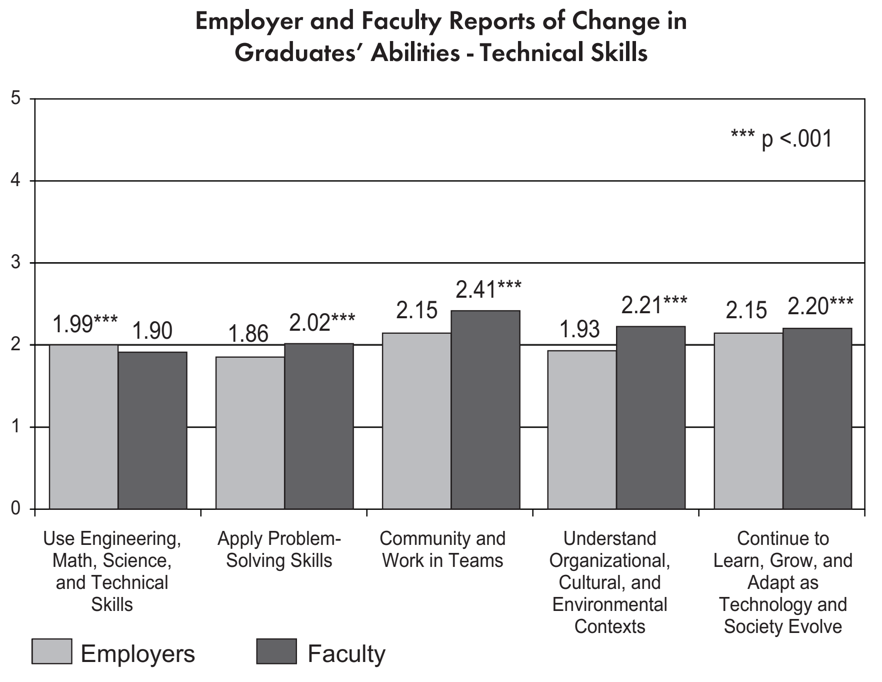

Figure 6. Employer and faculty reports of change in graduates' abilities and technical skills.

Note: *** $p < .001$

In our study, we also examined the connections, if any, between employer ratings and changes in engineering programs. To do this, we compared the employer ratings of recent graduates to the program chair reports of changes in curricular emphases. Specifically, we compared employer ratings of inadequate preparation and the chairs' reports of increases in program emphasis on the five clusters of engineering skills associated with EC2000.

Figure 7 reveals that the curricular clusters reported by chairs as receiving the most and least emphasis align roughly with the areas that employers view as weaknesses in the preparation of new hires. Most employers rate new hires as well-prepared in technical and problem-solving skills, and the reports by both program chairs (and faculty as well) indicate those as the two areas of least change. Program chairs and faculty report the greatest increases in emphasis in the communication and teamwork cluster, closely followed by contexts and constraints. A significant number of employers perceive these to be areas that continue to need improvement. Thus the actions by faculty and chairs appear to be largely in harmony both with the goals of EC2000 and with the perceptions of employers.

If there is a misalignment in Figure 7, it appears in the responses to adapting, learning and growing "as technology and social conditions evolve in unpredictable directions." Few employers see weakness here, but most chairs report this to be an area of increased emphasis. Naturally, that increased emphasis may still be quite modest, since chair- and faculty-reported increases in attention are relative rather than absolute. Curricular topics that received little or no attention before EC2000 may be the object of significant attention as a result of EC2000, but the degree to which they are emphasized in the curriculum overall may still be modest.

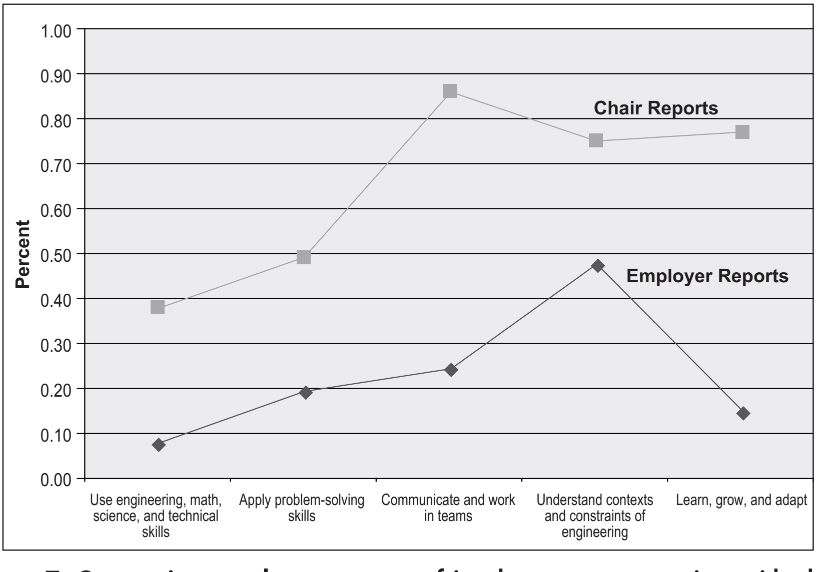

Figure 7. Comparing employer reports of inadequate preparation with chair reports of change in program emphases.

Conclusions

We analyzed 1,622 responses from a diverse array of employers from seven engineering fields (aerospace, chemical, civil, computer, electrical, industrial, and mechanical). The responding employers represent richly variable geographic locations, industry types, company sizes, organizational ranks, educational preparation, and evaluation experience. Despite their diversity, employers are in substantial agreement about the importance of the Criterion 3 (a–k) outcomes and the preparation of new engineers, even across engineering disciplines.

We conclude that the ABET emphasis on the (a–k) criteria is in harmony with the views of employers, and thus, engineering schools appear to be generally on the right track. For example, all 11 of the (a–k) criteria are rated as at least moderately important by 7 out of 10 employers, and nine items are rated as *highly important or essential* by at least 6 out of 10. Moreover, the faculty portion of our study suggests that engineering programs and faculty are giving a curricular priority to most of the areas that are rated highest by employers—effective communication, teamwork, modern engineering tools, and design. Moreover these findings are consistent with the few other published employer studies (Bodmer, 2002; Koen & Kohli, 1998). Although not much has been published about what's least important, it appears that employers agree about what's most important. Those who hire new engineers are looking for graduates with problem-solving and technical skills, along with the ability to communicate effectively and work in teams.

In their assessments of new engineers, over 90% of employers rate them as adequately prepared or well prepared to use math, science and technical skills, and about 8 out of 10 employers give them passing marks for problem solving and for learning, growing, and adapting. Teamwork and communication skills are assessed as at least adequate by 3 out of 4 employers. Since the advent of EC2000, the areas of modest increase are in teamwork, communication skills, and ability to learn, grow and adapt—all rated highly important by employers. The areas of modest decrease are in problem-solving skills and understanding cultural and environmental contexts and constraints. Although the erosion in problem solving is a concern, 80 percent of employers still rate students highly in this area. The area of least preparation (understanding contexts and constraints) is also the area of least importance to employers.

However, the overall impact of the (a–k) criteria in the eyes of employers generally has been quite modest thus far. In every skill area, more than half the employers report no increase or decrease in the abilities of new engineers over the 7–10 years preceding our data collection in 2004. In fact, when we compare employer and faculty perceptions, we see that employers report significantly less change than faculty in every area. There may be several factors contributing to the gap. One explanation may be the pipeline lag because faculty interact with the students well before they show up in the workforce and are evaluated by employers. Another explanation may be the lag in EC2000 accreditation. Some institutions underwent review earlier than others, and some institutions had not yet been reviewed under EC2000 when we collected our data in 2004. Finally, our analysis suggests that many of the most talented engineers go directly into graduate or professional school, rather than into the workforce. Those entering the workforce directly might be the ones least changed by their undergraduate engineering education, and—in any case—employers are not seeing the entire top of the talent pool.

Any interpretation of these findings needs to consider the limitations of the survey. We asked employers to respond within a fairly narrow context—namely, to evaluate the qualifications of newly hired bachelor's degree recipients and to assess the importance of the (a–k) criteria to their entry-level job performance. Areas of greatest employer importance appear to be the ones where preparedness is high and/or modest improvement is occurring. We cannot conclude that these are the ONLY qualities necessary for overall professional development and growth. Over the long haul, knowledge of contemporary global and social issues may prove to be as valuable and as enduring as the more technical and scientific skills.

However, the program and course changes alike appear generally in sync with the goals of EC2000. For their part, employers (despite their diversity in terms of engineering field, industry sector, geographic location, company size, and organizational rank) are in substantial agreement about the preparation of new engineers and, thus far, report only modest change. They believe that new engineers possess adequate competency in foundational and technical skills—and program chairs and faculty report little change in their courses and program curricula in these areas. In the EC2000 areas where employers perceive the need for more attention to skill-building, such as communication, teamwork and use of modern engineering tools, faculty and chairs report the greatest increases in curricular emphasis.

The results of this analysis suggest that engineering programs generally are heeding the requests of their industry partners to update and broaden the education they provide to engineering students, and ABET is assisting in that process by focusing their accreditation process on the development of student knowledge and skills consistent with the needs of industry. The rough alignment between changes in curricular emphases reported by program chairs and faculty versus the assessment of new hires by employers suggests the success of this coordinated effort to produce effective engineers for the 21st century. On the whole, engineering programs are increasing the emphasis on those areas of knowledge and skill that employers judge to be the least developed in new engineering hires. This increased emphasis on the many professional skills required for good practice has been accomplished without substantive decreases in attention to the development of necessary technical skills. Employer judgments about levels of preparation in basic math and science have not declined as a result of the implementation of EC2000. Moreover, the large employers that recruit nationally appear to see more EC2000-like changes than do the smaller employers that recruit locally.

Acknowledgements

The authors acknowledge and thank the other members of the research team who participated in the development of the design, instruments, and databases for this project: Dr. Linda C. Strauss, senior project associate; Suzanne Bienart, project assistant; and graduate research assistants Vicki L. Baker, Robert J. Domingo, Betty J. Harper, Amber D. Lambert, and Javzan Sukhbaatar.

References

ABET. (1997). *Engineering criteria 2000 [electronic version]*. Baltimore, MD: author.

American Society of Mechanical Engineers (ASME)/National Science Foundation (NSF). (1995). *Mechanical engineering curriculum development initiative: Integrating the Product Realization Process (PRP) into the undergraduate curriculum*. New York: American Society of Mechanical Engineers.

Bodmer, C., Leu, A., Mira, L., & Rutter, H. (2002). *SPINE: Successful practices in international engineering education. Benchmarking study*. Retrieved April 2, 2008, from http://www.ingch.ch/pdfs/spinereport.pdf

Hoey, J. J., & Gardner, D. C. (1996). Using surveys of alumni and their employers to improve an institution. In J. Pettit & L. H. Litten (Eds.), *New directions in institutional research: No. 101. A new era of alumni research: Improving institutional performance and better serving alumni* (pp. 43-59). San Francisco: Jossey-Bass.

Hoey, J. J., Marr, J., & Gardner, D. C. (2002, October). *Multiple vantage points for employment-related feedback: Some results*. Paper presented at the 2nd National Conference on Outcomes Assessment, Pittsburgh, PA. Retrieved April 2, 2008, from http://www.ineer.org/Events/ICEE2002/Proceedings/Papers/Index/Posters/P03.pdf

Koen, P. A., & Kohli, P. (1998, June). *ABET 2000: What are the most important criteria to the supervisors of new engineering undergraduates?* ASEE Annual Conference, Seattle, WA.

Penn State. (2006). *Engineering Change EC2000 Survey Instruments*. Retrieved April 2, 2008, from http://www.ed.psu.edu/cshe/abet/instruments.html

Volkwein, J. F., Lattuca, L. R., Terenzini, P. T., Strauss, L. C. & Sukhbaatar, J. (2004). Engineering change: A study of the impact of EC2000. *International Journal of Engineering Education*, *20*(3), 318–328.

Appendix

ANOVA Tests for significant differences among Employer Responses by Engineering Field

Since ABET accreditation is program-specific, we wanted to see if employer ratings differ significantly by engineering discipline. We analyzed 1,622 responses from a diverse array of employers from seven engineering fields (aerospace, chemical, civil, computer, electrical, industrial, and mechanical). Despite their diversity in geographic location, industry type, company size, organizational rank, educational preparation, and evaluation experience, employers are in substantial agreement about the preparation of new engineers and the importance of (a–k) criteria, even across engineering disciplines.

Figure Q-5 shows the relatively high agreement about the preparation of new engineers across engineering disciplines. For each of the five clusters of skills and abilities, the shaded boxes in Figure Q-5 highlight the highest and lowest employer ratings. In each case, the differences among engineering disciplines are small, and only three of the 105 ANOVA tests among the means in the table are significant (the ones

in bold for 5a, 5c, 5d). The highest employer ratings in all five skill areas are given by employers of industrial engineers, and the lowest ratings in the three professional areas (5c,5d, 5e) are given by employers of computing engineers.

How well prepared are recent Engineering Grads?	Aerospace	Chemical	Civil	Computer	Electrical	Industrial	Mechanical	Overall
To use engineering, math, science, and technical skills? (5a)	**2.26**	2.46	2.38	2.39	2.32	**2.56**	2.45	2.41
To apply problem-solving skills? (5b)	2.05	2.11	2.05	2.08	2.00	2.14	2.04	2.06
To communicate and work in teams? (5c)	1.95	2.02	1.96	**1.79**	2.01	**2.16**	1.93	1.97
To understand the organizational, cultural, and environmental contexts and constraints of engineering practice, design, and research? (5d)	1.64	1.60	1.56	**1.47**	1.57	**1.73**	1.64	1.58
To continue to learn, grow, & adapt as technology & social conditions evolve in unpredictable directions? (5e)	2.15	2.12	2.18	2.10	2.14	2.21	2.17	2.16

Yellow = lowest value in each row, Green = highest value in each row

Figure Q-5. Question 5: Employer ratings of engineering graduates preparation.
Note: Numbers **in bold s**how statistically significant differences between the high and low scores.

The survey also solicited employer assessments of changes in the preparation of new engineer hires since the mid-1990s. Once again, there is high agreement across engineering specialties and skill areas, and a majority of the employers in each fie ld report no increase or decrease in the abilities of new engineers since the mid-1990s. Figure Q-6 shows that in three areas (6a, 6c, 6e), industrial engineers received the highest improvement scores by their employers. Aerospace engineers and civil engineers received the lowest improvement ratings on two dimensions each. Nevertheless, the differences by field are so slight that only three of 105 ANOVA tests are statistically significant (the industrial engineer employers see significantly more improvement in the ability to communicate and work in teams than do the aerospace, civil, and electrical engineer employers).

Increase or Decrease over 10 years?	Aerospace	Chemical	Civil	Computer	Electrical	Industrial	Mechanical	Overall
To use engineering, math, science, and technical skills? (6a)	1.98	2.00	1.97	2.02	1.98	2.06	2.01	1.99
To apply problem-solving skills? (6b)	1.80	1.88	1.82	1.98	1.85	1.97	1.88	1.86
To communicate and work in teams? (6c)	**2.05**	2.28	**2.08**	2.12	**2.08**	**2.37**	2.18	2.15
To understand the organizational, cultural, and environmental contexts and constraints of engineering practice, design, and research? (6d)	2.00	1.95	1.87	2.04	1.94	1.96	1.99	1.93
To continue to learn, grow, & adapt as technology & social conditions evolve in unpredictable directions? (6e)	2.11	2.12	2.14	2.14	2.10	2.25	2.22	2.15

Yellow = lowest value in each row, Green = highest value in each row

Figure Q-6. Question 6: Employer ratings of change in ability over the past decade.
Note: Numbers **in bold** show statistically significant differences between the high and low scores.

Finally, we asked employers to rate the importance of each (a–k) outcome on a 5-point response scale. As with the other parts of the survey, we searched for differences by engineering field, and the results are displayed in Figure Q-7. Again, the highest and lowest means on the 5-point importance scale are highlighted, and the significant differences across disciplines are bolded. Only 20 of the 231 ANOVA tests among the means in Figure Q-7 are statistically significant. For example, perhaps reflecting the nature of their work, aerospace engineer employers place a higher importance on math and science skills than their counterparts in industrial engineering. Likewise, computer and industrial engineers have opposite views about the importance of teamwork and the global and societal context. For another example, the employers of civil engineers place a lower value on designing experiments and analyzing data than most other groups of engineering employers.

ABET (a-k) Criteria	Aerospace	Chemical	Civil	Computer	Electrical	Industrial	Mechanical	Overall
Ability to apply knowledge of mathematics, science and engineering (7a)	**4.24**	**4.19**	4.17	**3.87**	4.05	**3.82**	4.04	4.10
Ability to design and conduct experiments, as well as to analyze and interpret data (7b)	**3.98**	**3.86**	**3.30**	**3.68**	**3.93**	**3.81**	**3.69**	3.60
Ability to design a system, component, or process to meet desired needs (7c)	**3.76**	**3.76**	**3.73**	**4.13**	3.84	**3.74**	3.84	3.80
Ability to function on multi-disciplinary teams (7d)	4.24	4.25	**4.01**	**3.87**	4.10	**4.36**	4.14	4.10
Ability to identify, formulate, and solve engineering problems (7e)	4.10	4.18	4.28	4.13	4.20	4.09	4.16	4.21
An understanding of professional and ethical responsibilities (7f)	4.07	3.97	**4.09**	**3.53**	3.95	3.97	**3.75**	3.95
Ability to communicate effectively (7g)	4.36	4.39	4.42	4.27	4.32	4.52	4.41	4.39
Understand the impact of engineering solutions in a global and societal context (7h)	2.83	2.94	2.98	**2.66**	2.90	**3.12**	2.88	2.93

Figure Q-7. Question 7: Employer ratings of importance by engineering field.
(Continued on next page)

Recognize the need for, and an ability to engage in life-long learning (7i)	3.74	3.70	3.68	3.84	3.70	3.63	3.56	3.68
A knowledge of contemporary issues (7j)	**2.74**	2.85	3.07	2.76	2.79	**3.08**	2.86	2.94
Ability to use the techniques, skills, and modern engineering tools necessary for engineering practice (7k)	4.02	3.90	4.11	3.88	4.00	3.96	3.89	4.00

Yellow = lowest value in each row, Green = highest value in each row

Figure Q-7. Question 7: Employer ratings of importance by engineering field. Note: Numbers **in bold** show statistically significant differences between the high and low scores.

The *Assessment in the Disciplines* volumes have been designed to provide assistance to both faculty who have taken on responsibility for assessing their academic programs, as well as institutional researchers who are often asked to support student learning assessment activities across their campuses. We hope that the discussions presented in this series will contribute to the development of assessment strategies that will result in improved student learning on our college and university campuses.

Association for Institutional Research