Quantifying ABET Assessment of Outcomes

Martin R. Parker¹ and William A. Stapleton²

Abstract

The essence of the relatively new ABET EC 2000 assessment criteria is a requirement that Programs under evaluation by the accreditation board clearly demonstrate that, at the very least, 11 separate learning objectives (outcomes) are being achieved. Such evidence can only be provided by a comprehensive assessment of the effectiveness of instruction related to these outcomes. Methods of assessment of outcomes have been widely discussed by academics in recent times. Several of the more high-profile approaches, (student grades, retention rates, FE exam results, graduate school admissions, employment statistics, etc.), seem to us to be somewhat limited in their usefulness. The approach advocated here is founded on the belief that one key to better measurement of outcomes as well as to a better process for the achievement of learning objectives lies primarily with student feedback. This paper discusses how student-based outcomes assessment can be used to quantify the achievement of learning objectives, and, thereby, the quality of instruction, in a meaningful way.

Introduction

The recently constituted set of criteria known as Engineering Criteria 2000 (EC2000), adopted by the Accreditation Board for Engineering and Technology (ABET) to deal with the matter of accreditation of engineering programs, has reflected a growing disillusion in the engineering community with the more traditional yardsticks for evaluation of the quality of the final product [1]. The latter have included sample test papers, student grades, SAT and GRE scores, graduate employment data, etc. The thrust of these new ABET criteria, however, is clearly aimed just as much at the effectiveness of the learning process as it is at the rigor of the program. Of course, once a program has been deemed to possess sufficient rigor, individual student grades may contribute to confirming the learning process of the individual engineering student. Such traditional methods, nevertheless, fail, in most instances, to provide any clear markers of the effectiveness of the learning process for the student body taken as a whole. In the present context, the corporate student body may be considered in units as small as a single group attending a particular class. Alternatively, it may be considered to be the entire student body in a given program.

Our findings appear to verify that non-traditional assessment can be used to validate that effective group learning has (or has not) occurred. Intriguingly, such assessment tools can also be used evaluate the quality of the instruction provided.

Assessment Details

The present authors have, in cooperation with other faculty, experimented with a variety of non-traditional assessment procedures over a period of several semesters. The scope and variety of such potential assessment tools is almost limitless. All of the methods used by the present authors, however, which previously have been described

---

¹ Martin R. Parker, Department of Electrical and Computer Engineering, University of South Alabama, Mobile, AL 36688.
² William A. Stapleton, Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL 35487.
by us in detail [2, 3] have one common feature: they involve the assignment of numbers (on a scale of 1 to 10). Our data indicates a surprisingly high level of consistency between the various assessment tools, as indicated by a master matrix of numerical scores for the desired program outcomes judged to best address the eleven learning objectives of ABET Criterion 3. That having been said, one particular assessment tool has been seen to correlate better with the numerical scores for the overall evaluation of the Programs than any other. Not surprisingly, the most valuable of all approaches is an end-of-semester questionnaire responded to by all students in every (engineering) class offered in a given semester. A typical form for a typical class is illustrated in Figure 1.

In this form, the student, on the last meeting of the class, evaluates his or her perceived level of understanding of (usually 5 to 10) key learning objectives for the course in question. The student completes the form in the absence of the instructor and assigns a number, on a scale of 1 to 10, in response to the various questions. The student is encouraged, but not compelled, to provide a name (or student ID number). The instructor does not see these responses until after grades have been posted. Once individual student grades have been determined, the responses of each student to the key issues of each course are averaged. These can then be correlated with the letter grades of those students, (also on a 10-scale, with (say) 9-10 for an A, 8-9 for a B, etc.). Future plans include similar correlations to embedded course assessments such as individual test questions or homework assignments in addition to the overall course grade to provide a finer granularity of assessment [4].

Detailed correlation is, of course, crucially dependent on the willingness of individual students to provide a personal identification on his or her response to the questionnaire of Figure 1. However, even in the absence of such student cooperation, a limited, but still valuable correlation is often still possible.

![Figure 1: Typical End of Semester Questionnaire](image-url)
Before addressing the issue of correlation, the value of the assessment tool of Figure 1 is now explained. Each course offered in a given semester is considered to be a metric for at least one of the (17) desired outcomes of the (Electrical/Computer Engineering) Programs. The desired outcomes for Computer Engineering, for example, are shown in Figure 2, where the relevance of those outcomes to the 11 objectives of ABET Criterion 3 is also indicated.

The average score of each student for each class is determined. That score is assigned to one or more of the 17 desired outcomes addressed by that class. This exercise is repeated for every class offered in a given semester. In this way multiple scores (1 → 10) are assigned to each of these outcomes from all of the classes, each semester. In this way, a current record of attainment of the desired outcomes is always at hand.

---

**Computer Engineering Program Outcomes**

Students completing the degree requirements in the Program of Computer Engineering should have:

1. A broad understanding of the theory of mathematics, probability and statistics, including that relating to differential and integral calculus, discrete mathematics, differential equations, and complex variables.
2. An understanding of basic scientific theory and methods in the areas of chemistry and calculus-based physics.
3. An understanding of engineering fundamentals.
4. An ability to apply mathematical, scientific, and engineering principles in the field of Computer Engineering. [Criterion 3(a)]
5. An ability to plan and conduct experiments in engineering science and engineering design. [Criterion 3(b)]
6. Design skills sufficient for the successful completion of a process, component, or system. [Criterion 3(c)]
7. An ability to function on multi-disciplinary teams. [Criterion 3(d)]
8. A capacity for problem identification and formulation, background research, solution generation, and decision making. [Criterion 3(e)]
9. An ability to identify global, societal, legal and other key issues in arriving at ethical decisions in professional life. [Criterion 3(f) and 3(h)]
10. A capacity for effective written and oral communication. [Criterion 3(g)]
11. In-depth education in the hardware and software subdisciplines of Computer Engineering.
12. Recognition of the need for, and an ability to engage in life-long learning. [Criterion 3(i)]
13. Knowledge of contemporary issues and an awareness of the changing technological environment. [Criterion 3(j)]
14. An ability to use modern engineering techniques, skills, instruments, and software tools necessary for effective participation in the Computer Engineering Profession. [Criterion 3(k)]
15. An appreciation of the unique concerns regarding safety required when working with electrical systems.
16. A general preparedness for graduate school or professional licensure.
17. A consideration of factors including economic, environmental, sustainability, manufacturability, reliability, reusability, serviceability, health and safety, social, and political in the design of digital systems.

---

**Figure 2: Computer Engineering Program Outcomes**
Assessment of Instruction

Traditional assessment tools offer little insight into the quality of instruction. Two unrelated engineering classes may have nearly identical near-Gaussian grade distributions but totally dissimilar results in terms of student opinions of the teaching involved. This is most crucially evident in response to the summative question: “Did the instructor do an overall effective job in teaching the course?” In the aftermath of all of this, the key question then arises: is the student criticism valid or unreasonable?

To respond to this we offer the following rubric, shown in Figure 3, as a guide to determining effective instruction. Basically the rubric divides the data into four domains, each of which relates to two axes: the other to student levels of understanding of course fundamentals/core issues. Only one of the domains is desirable and only two are wholly acceptable. The desirable domain, of course, is that in which good scores and good grades correlate strongly. The polar opposite of this domain, in which poor grades and poor levels of understanding correlate well, is undesirable from a group learning aspect but, at least, logically consistent. The other two domains are, of course, unacceptable but to varying degrees. Poor grades, correlated with high levels of perceived comprehension, are possibly indicative of unreasonably difficult testing. In isolated individual cases, it may indicate unrealistically high self-assessment. By far the most worrisome domain, fundamentally and, most significantly from the point of view of satisfying Criterion 3, is that in which good grades correlate with poor levels of understanding. This is the classic case of “grade inflation.”

![Figure 3: Assessment Rubric](image)

An example of good correlation between student comprehension and student grades is seen in Figure 4. This is an example of a class in which a relatively high level of personal identification was provided in the student responses. Here, the shaded area indicates the desirable domain in which students’ grades and perceptions of learning are at least minimally acceptable. (In the grading paradigm used here, the domain lies between 6 and 10 on both axes). Note that a few of the data points lie in the second domain, mentioned earlier, in which poor levels of perception are recorded but with logically consistent grades; in the current paradigm, the domain would lie between 1 and 6 on both axes.
Using Figures 3 and 4, it would appear to the present authors that a desirable standard for effective group learning is that 90% of students receiving pass grades (i.e. D or better), should lie in the $6 \rightarrow 10$ domain (or the equivalent thereof in other grading paradigms).

Even with the best will in the world, sometimes, for whatever reason, students may be reluctant to self-identify. Under those circumstances one may still glean valuable information regarding the quality of instruction. Even in the absence of personal identification, our experience has shown that histograms of letter grades versus student perceptions of learning ought to correlate reasonably well. An example appears in Figure 5. Very poor correlation of the dual histograms should raise a red flag.

In this initial implementation, assessments of student perceptions of learning have been correlated with student grades as a measure of performance. Some bias may occur to this correlation as students generally have a reasonable awareness of the grade they will receive before they complete their responses. In the future, the authors plan to explore the efficacy of utilizing embedded assessments which are much more finely-grained such as individual test questions, homework, or laboratory assignments each designed to assess a specific desired outcome. At the other extreme of granularity, some other existing broad assessment tools such as the FE exam can also be considered.

![Figure 4: Example of Strong Correlation Between Grades and Student Perception](image-url)
One of the more traditional yardsticks of quality assessment, the FE examination results, can also be used to evaluate the effectiveness of instruction. The overall FE exam results (i.e. pass rates), while interesting, invariably offer little insight into the quality of a particular engineering program. In some programs, students are required to take the exam; in others this is not the case. In some programs, a high percentage of the student body comprises foreign students who have little motivation to pass the exam. A further skew on the results is that a program may choose, for perfectly legitimate reasons, not to include in the engineering curriculum key sections of the FE exam, such as fluid mechanics, dynamics, and the like. Using pass rates to evaluate a program may, accordingly, be akin to comparing apples and oranges. That being said, discipline-specific pass rates (for circuits, electromagnetics, controls, etc.), may offer valuable insight into the quality of the program but, more importantly, the quality of instruction. An example from The University of South Alabama is shown in Figure 6. A discontinuous jump (either rise or fall) in pass rates might provide important information concerning the quality of instruction. A temporal record of this sort is, accordingly, an important assessment tool. One example of a potential problem that was noted from the data of Figure 6 appeared in the first group on the chart, “Analog Electronic Circuits” where a performance delta between semesters 4 and 5 was noted and continued into semester 6. Upon review it was noted that both the instructor and the textbook had changed and that adjustments needed to be made to the course content to ensure complete topical coverage. For many of the subject areas too few points have been charted to indicate a clear pattern. Continued collection of assessment data should more fully define the patterns.
Conclusion

In the above, we have attempted to demonstrate that that a variety of simple numerical assessment tools can provide useful insight into engineering programs, not only in terms of metrics indicating achievement of educational objectives for accreditation agencies but also internally in the evaluation of the quality of instruction being provided. As a step in the continuing process of refining the tools used to assess the efficacy of instruction, correlation between reported student perception and grades has been shown to provide useful information. The student survey served to identify a service course at The University of Alabama that was not providing coverage of a key prerequisite topic when taught by a particular instructor. This led to an improvement in the enumeration of expected outcomes on a course-by-course basis throughout the program. The authors are in the process of incorporating embedded assessments throughout their courses to improve on the granularity of the data gathered to provide more complete assessment than the single course grade can provide.

References


ASEE Southeast Section Conference 2004

**Martin R. Parker**

Martin R. Parker received his B.S. degree from The University of Glasgow and his Ph.D. from The University of Salford, U.K. His research interests have been mainly in the field of magnetics, including magneto-optics, magnetic separation, and magnetoresistance. From 1990 until 1996 he was the M.I.N.T. Professor at The University of Alabama. He is a Professor and former Chair of Electrical and Computer Engineering at The University of South Alabama. In 1985, he became both a Senior Member of the IEEE Magnetics Society and Fellow of the Institute of Physics, U.K.

**William A. Stapleton**

William A. Stapleton has a Ph.D. in Electrical Engineering from The University of Alabama. He began his teaching career at The University of South Alabama where he became well-acquainted with his co-author Martin Parker. Dr. Stapleton is currently serving as an Assistant Professor at The University of Alabama. His areas of specialization and interest are in the computer architecture and digital systems fields. Dr. Stapleton’s topical interests include image compression, parallel and distributed computing, and embedded systems design.