A Framework for a Simple and Effective Assessment and CQI Process

Dr. William Lasher, Dr. Russell L. Warley, Mr. Jonathan Meckley
School of Engineering
Penn State Behrend
Erie, PA
lasher@psu.edu

***Abstract*—This paper presents a proposed structural framework for an assessment and CQI process that is both simple and effective. The process is based on the ideas that assessment should be done where it makes sense, it should not be necessarily redundant, and it should provide useful information for improving the program. Elements of the proposed process have been piloted in two separate programs, including direct and indirect summative assessment of program outcomes and direct formative assessment. Preliminary results suggest that the process will be simple and effective.**

Keywords—assessment; continuous quality improvement

I. Introduction

The School of Engineering at Penn State Behrend consists of five engineering programs, three engineering technology programs, and computer science, all of which are ABET accredited or are in the process of becoming accredited. ABET requires that all accredited programs have a process for assessing student outcomes at the time of graduation, and a continuous quality improvement process that uses the results of this assessment and other data as input. The assessment process used at Behrend, which is recommended by ABET and followed at many other schools, is to collect examples of student work (instruments) throughout the curriculum (commonly referred to as embedded assessment) and relate that work to one or more outcomes. If the students perform satisfactorily on each instrument it is presumed that the outcome has been met.

The process described above is time consuming and, in the experience of the authors, not very effective. It takes time to remind faculty to collect the various instruments, organize them, and evaluate what the results mean. A common recommended action at Behrend from an assessment is to refine the instrument because it did not seem to measure what was intended. Additionally, there are outcomes such as student knowledge of mathematics that, according to our processes, are being met satisfactorily; however, the faculty perceive that our students are weak in applying math to engineering courses (although this could conceivably be influenced by the performance of a small percentage of students). A likely reason for this contradiction is that what students can do or know in a course and what they can do or know following the course can be significantly different. For example, in a study of medical education students retained less than half of what they had learned after two years [1], and in a consumer behavior course most of the knowledge was lost within two years [2]. Finally, this process is designed to determine whether outcomes have been met, leaving interpretation and corrective action as a follow up task.

We are proposing a process that we believe will be simpler and more effective than our current process. A fundamental difference in the proposed process is philosophical - to change the focus from "assessing outcomes to satisfy ABET" to "assessing student learning to improve the program while satisfying the ABET requirement" – a viewpoint which is supported by Penn State's Office of Planning and Assessment. The guiding principles for this process are:

- Assess attainment of outcomes only once as close to the end of the program as possible.
- When an outcome has not been met at the desired level, obtain information that will help determine the potential reasons for the outcome not being met.
- Minimize the effort needed to assess any outcomes that are being met at a high level and focus attention on outcomes where action is indicated.
- Obtain assessment information that will improve the program, regardless of whether it is useful as a measure of whether an outcome has been met.

II. Overview of Process

The proposed process is summarized in Fig. 1 and explained below:

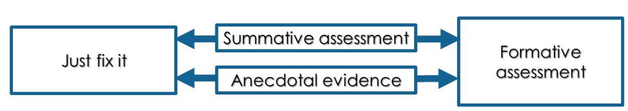

Fig. 1. Flowchart of proposed process

The process consists of both summative and formative assessment. The summative assessment determines whether the outcomes have been met through both direct and indirect assessment. The direct assessment is primarily (although not completely) done at the final presentations for the senior project. The indirect assessment consists of a focus group of graduating seniors. The formative assessment will be initiated based on problems identified by either the summative assessment or "anecdotal evidence", which basically means that there is reason to suspect a problem. Formative assessment will be used when it is not clear what the cause of the problem is or how to resolve it. It will be initiated and discontinued as needed; once an issue has been resolved there is no need to continue to assess it at the formative level. Additionally, problems that are identified with an easy fix will simply be fixed. Parts of the proposed process have been piloted, and a summary of these is provided below.

III. Summative Assessment of Program Outcomes

A. *Direct assessment*

Our objective is to perform summative assessments of the outcomes as close to graduation as possible. Most of the outcomes can be assessed in the capstone design projects course, while some of them are better assessed in other junior or senior-level courses. Table I summarizes which of the EAC ABET a-k outcomes can be assessed either entirely or partially in the Mechanical Engineering (ME) capstone projects course.

TABLE I. Summary of Outcomes that can be Assessed in Capstone projects

Outcome/Outcome text	
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 multidisciplinary 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
i	a recognition of the need for, and an ability to engage in life-long learning
k	an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice

Due to limited space we cannot describe how each of these outcomes is assessed in the capstone course; some of them (c, g and k) are obvious, whereas others require explanation. For example, with outcome i faculty evaluation committees will be asked to rate (on a scale of 1-5) the ability of the team to:

1. Identify something needed for their project that they didn't know from previous classes and explain how they learned it.
2. Explain what the implications of this are for their careers once they graduate.

The first item demonstrates an ability to engage in life-long learning; for the second item the faculty will be looking for evidence that the students recognize the need for life-long learning (for example, stating that they will need to continue to take courses or learn things on their own).

We also believe we will be able to assess outcome f, at least in part, by using the following performance criteria:

1. Develop a safe design.
2. Meet the needs of the sponsor.
3. Be honest about problems they encountered and acknowledgement of weaknesses in their solution.
4. Appropriately maintain confidential information from their sponsor.
5. Give appropriate credit to others.

Additional assessment of outcome f will be done during the in-class part of the course.

Outcome a will be assessed through prerequisite quizzes for the machine design and heat transfer courses, which are taken at the end of the junior year or beginning of the senior year. Outcome b will be assessed in the Measurements course which is taken during the junior or senior year. We have not yet determined where outcomes h and j will be assessed; this will be done over the summer or early fall and we hope to include it in the presentation along with some results of the assessments from the capstone project presentations.

B. *Student interviews*

Students provide a different perspective on the outcomes, and their opinions on whether an outcome has been met should be a part of the assessment. Traditionally this is done through student surveys; however, follow-up discussion is not possible during a typical survey and it may be difficult to figure out what the problem is from a survey response. We are proposing to use focus groups, where a sample of students will rate attainment of performance criteria for each outcome, and if they rate something low, the facilitator will be able to lead a follow-up discussion in real time to better understand the issue and see if the students have suggestions for addressing it.

In December of 2015, the Plastics Engineering Technology (PLET) program piloted their first focus group of the graduating class to determine the strengths and weaknesses of the program. Fifteen students were randomly chosen to participate. To provide balance, half of the students were selected from the top of the class and half from the bottom. The invitation explained why the focus group was being performed and asked for their feedback to help make the PLET program stronger. It also instructed the students to keep their participation undisclosed from the faculty. This was done to allow the students to feel they could freely share their opinions.

Of the fifteen students invited, fourteen students attended. The Director for the School of Engineering served as a moderator and an administrative assistant was present to take notes. The moderator set the ground rules by explaining this was a discussion on the curriculum of the program and not a forum to discuss individual faculty. The grading scale to be used was associated with the five buttons marked A through E on the survey clickers that were used. "A" meant that the student felt strongly that they were taught and learned that particular knowledge or skill, "C" meant they felt it was taught by needed to be reinforced, and "E" meant that they felt it was never taught or could not be applied.

The moderator followed a set of instructions to evaluate whether the question needed follow-up discussion based on an analysis of real-time responses visible only to the moderator. If the responses were all A's and B's or at most two C's, then there was no need to ask follow-up questions. If there were several C's or any D's or E's, then follow-up questions were to be asked. The survey was constructed around the ETAC-ABET a-k outcomes. If there was more than one performance criteria in an outcome, questions were asked on the individual criteria rather than on the overall outcome.

The same questions were posed to the faculty separately. The evaluation was not based on how well the faculty felt they taught the outcome or criteria, but how well they felt the students met the outcome or criteria. This was based on their opinions of student's performance on homework, exams, labs, and project work. Even though all of the faculty were involved in the construction of the survey, several of the performance criteria and a few of the outcomes were noted as confusing when the faculty took the survey. These outcomes and performance criteria will need to be reworded or redefined in future surveys.

For the most part, the survey results indicate the outcomes are being met. There were a few items that generated discussion. Most of this centered around a for-credit course on plastics processing and statistical methods. Statistics are used to evaluate the results of the experiments run in the lab. The students felt that there was too much material covered in the four credits and that they still cannot apply the statistics one and a half years after taking the course. They also felt the course seemed rushed. However, when asked if the course should be broken into two courses, the students said it was acceptable the way it was. While there were mostly comments on what was wrong with the course, one practical suggestion was provided by the students. It was suggested that some of the statistics content and analysis software could be moved earlier in the curriculum to the computer applications course in the freshman year, which is being considered.

Another discussion item concerned several support courses - CAD, Statics, Stress Analysis (FEA), and Strength of Materials. The students noted that GD&T is taught in the CAD course but is not used in the PLET curriculum. The students noted that they felt weak in application of FEA to part design. Currently, there are several new projects being discussed that will use stress analysis. The faculty also expressed concern that the students cannot apply the knowledge they learned in Strength of Materials. There is ongoing discussion about using the content from Strength of Materials earlier in the curriculum since there is typically a year gap.

Perhaps the most revealing comments concerned the use of math. The survey response to the math question in the focus group generated follow up questions. Surprisingly, the students stated that they have no problems using algebra and trigonometry, and their negative survey response was related to not knowing how to select the correct equations to solve problems, notably in Thermodynamics and Heat Transfer. From homework, exams, and lab reports, the faculty noted that some students struggle to do simple algebra and trigonometry.

At the end of the student focus group, there was an open discussion on any other topic in the PLET program which gravitated to the senior project course. The PLET program is in the process of making changes to the project process and many of the comments concerned confusion over due dates and requirements. Since students are given the guidelines well before the due dates for the papers and presentations this seems to be a transient issue.

The PLET faculty feel strongly that the results of the survey were valuable and give the department some areas to develop and strengthen. It was apparent that the next iteration of the survey should move away from some of the outcomes and performance criteria that were met in the survey and a separate Senior Exit Survey. The focus should instead be on areas where there are perceived weaknesses, such as where faculty see issues or where the Senior Exit Survey indicates problems.

IV. Formative Assessment

If the summative assessment indicates a problem, or if there is anecdotal evidence that a problem exists that was not identified through the summative assessment, one of two actions will be taken. If the solution to the problem is obvious then the problem will be solved. If there is a need for additional assessment, then a formative assessment process will be developed and implemented. The purpose of the formative assessment is to further identify the root of the problem and also directly assess whether programmatic changes have had the desired effect. This formative assessment will continue until the problem is resolved, at which point it will end.

Two formative assessments were implemented in the ME program during the spring 2016 semester, both based on anecdotal evidence. Faculty believed that our students were weak in math and statics in spite of the previous assessment process suggesting there was not a problem in these two areas. In both cases a prerequisite assessment was developed and implemented during the first week of classes. The math assessment was done in the statics course. All students in statics have passed the first calculus course and most of them had passed the second calculus course. A statics assessment was done in the strength of materials course which has statics as a prerequisite.

In both cases the assessments did not count towards the student's grade in the course. The advantage of grading the assessments is that the students could be more likely to take them seriously; however, the disadvantages are that it might encourage students to cheat, and we were more interested in finding out what the students actually knew than what they could memorize for an assessment. Furthermore, graded quizzes can induce enough anxiety that students focus on their anxiety rather than doing well on the quiz [3], and a study comparing the effect of grading vs. not grading homework showed that exam performance in economics was improved for freshman but not for other groups of students [4]. Both of these studies suggest that the presumed incentive advantage of grading prerequisite assessment may not be significant. Finally, the use of a prerequisite assessment has the advantage of providing the instructor with information about the students' prior knowledge, which can help the instructor tailor the course to the students' background [5].

The math assessment was completed by 81 students and consisted of 29 questions covering algebra, geometry, trigonometry, and calculus. It was done on the first day of classes and no electronics, notes or books were allowed. The average scores on the four parts were: algebra – 83%; geometry – 39%; trigonometry – 60%; calculus – 55%. The questions were straight-forward and focused on understanding basic concepts as opposed to memorization of formulas. Each problem was scored as correct/incorrect with no partial credit.

The high score in algebra is encouraging and indicates that the students took the assessment seriously. The low scores on the other three subjects suggest serious problems. For example, in trigonometry, only 27% of the students could identify the law of sines or the law of cosines as the way to solve a triangle. A surprising fraction of students could not correctly answer the most trivial questions. Only 82% could determine the length of one side of a triangle that was similar to another triangle but 1.5 times larger, and only 79% could determine the length of the adjacent side of a right triangle given the hypotenuse and included angle. Only 12% of the students could determine the area of a sector of a circle, only 15% knew the limit definition of a derivative, and only 31% could find the location of the local minimum of a simple cubic polynomial.

The results of this assessment were shared with the mathematics faculty, who found them both shocking yet not totally surprising. They are very interested in trying to improve the assessment results and will be discussing ways to do this. Even if their attempts are not effective, it is helpful for the engineering faculty to know where the students are weak so these areas can be reinforced. Interestingly, one of the statics faculty members asked his students why they thought they did so poorly, and they responded that they were surprised because they thought they knew the material. The disconnect between what students know and what students think they know is significant, which raises questions about the way in which the focus group results should be interpreted.

The statics assessment consisted of 14 simple problems and was administered to 135 students. Overall the results were similar to the results of the math assessment. These students did reasonably well (scoring above 74%) on some problems (such as calculating the 2-D moment due to a force or drawing the free body diagram of a truss section or joint), however, only 17% of the students could determine the 3-D moment of a force about a point, only 25% could determine the centroid of a simple 2-D shape, and only 30% could draw a free body diagram of a circle resting on a V with frictionless surfaces.

The results of the statics assessment were compared with the final course grade and the final exam grade for a sample of 32 students who took the statics course the previous semester. The average of the assessment scores was 21% lower than the average of the final course grades and 23% lower than the average of the final exam grades. The correlation between the statics assessment and the final course grades in the prerequisite statics course was low ($r^2 = 0.17$), and there was essentially no correlation between the assessment scores and the final exam scores ($r^2 = 0.021$). The reasons that students did so poorly on so many simple statics questions yet were able to pass the statics course are complex; however, one of the potential reasons is that the exams in statics are focused exclusively on the student's ability to solve problems. We believe they can get very good at solving a specific type of problem by memorizing solutions, and yet will forget the solutions as soon as they finish the course. One of the changes made in statics as a result of this assessment is to try to assess the students' understanding of concepts more specifically in regularly assigned problems. Continued administration of the assessment will allow us to determine whether this change (and other changes made in the future) had the desired effect.

V. Strengths and Weaknesses of Proposed Process

The weakness of the proposed process is that the outcomes are assessed only once at the end of the program. This creates two potential problems. Multiple (redundant) assessments throughout the program are more likely to detect an issue with attainment of outcomes. In addition, if there is an issue early in the program it may not be detected until three years later. The first potential problem is mitigated by the fact that our current process did not reveal an issue in math and statics that was revealed by the ad-hoc process. The second problem is mitigated by the fact that we will be assessing every outcome every year, as opposed to the current process where the outcomes were assessed on a three-year cycle (which is typical).

In our opinion the strengths of the proposed process outweigh the weaknesses and inefficiencies of the current process. The most significant strengths are:

- A much simpler process that is easier to implement and maintain
- Faculty buy-in, because the purpose of the process is to improve education, not just satisfy ABET
- An ability to identify problems that are not identified through assessment of the outcomes
- Information regarding the potential reasons for not meeting an outcome through student focus group interviews
- Assessment of every outcome every year
- Minimizing the effort required to assess an outcome when the outcome is already being met at a high level

VI. Summary

We believe the proposed process will be a significant improvement over the current process. It is more efficient and, we believe, more effective. The results of the process have already engaged the mathematics faculty in a way that the previous process could not. The engineering faculty are excited about the ad-hoc assessment because it directly measures their students' learning and allows them to assess changes they make in their courses. The faculty were also engaged in the development of both the student focus group questions and the performance criteria for the assessment through the capstone design course. Preliminary results indicate areas for improvement that would not have come out of the previous process, and in the end we believe we will get more and better data to help improve the program. In the end, that is what assessment should be about.

REFERENCES

[1] E.J.F.M. Custers, “Long-term retention of basic science knowledge: a review study,” Advances in Health Sciences Education, vol. 15, No. 1, pp. 109-128, February 2008.

[2] D.R. Bacon and K.A. Stewart, “How fast do students forget what they learn in consumer behavior? A longitudinal study,” J. Marketing Education, vol. 28, No. 3, pp. 181-192, December 2006.

[3] M.M. Khanna,”Ungraded pop quizzes: test-enhanced learning without all the anxiety,” Teaching of Psychology, vol. 42, No. 2, pp. 174-178, 2015.

[4] W.A. Grove and T. Wasserman, “Incentives and student learning: a natural experiment with economics problems sets,” The American Economic Review, Vol. 96, No. 2, pp. 447-452, May, 2006.

[5] S.A. Ambrose, , M.W. Bridges, M. Dipetro, M.C. Lovett and M.K. Norman, How learning works: seven research-based principles for smart teaching, San Francisco: Jossey-Bass, 2010, pp. 10-39.