
Avoiding the Pratfalls of Program Assessment

C. M. Chewar, K. L. Huggins, and Jean R. S. Blair

Department of Electrical Engineering & Computer Science
 United States Military Academy
 West Point, New York 10996 USA
 {christa.chewar, kevin.huggins, jean.blair}@usma.edu

Abstract

Most agree that assessment is an important step in improving an academic program. However, implementation often falls short of expectations due to common missteps. In this paper, we identify and discuss common program assessment pratfalls. We also describe the assessment program and the manner in which we have taken steps to avoid each of these pratfalls.

Keywords: Outcomes, assessment, accreditation

1. Introduction

Many of us have undoubtedly experienced some form of job enlargement through program assessment initiatives. Curricular accreditation boards, such as ABET/CSAB, require collection and documentation of data relating to program outcomes through some sort of an assessment effort. Those that extol the virtues of assessment efforts (or any quality improvement process for that matter) might argue that quality results from a continuous critical review process that drives development and revision of program strategy. However, a common perception of such programs is that they are largely a waste of time and effort.

For years, we have been talking about likely assessment program pitfalls [1]. We argue that a “pitfall” can only be thought as such for so long before it should be considered a “pratfall.”¹ For the record, we do have confidence that a well-designed and well-led assessment effort can yield many positive results within a curricular program.

In this paper, we probe forces we believe often underlie the negative feelings directed toward assessment programs. We discuss these forces as *pratfalls*, and present the tenets of our assessment program that should help us achieve our program’s potential. This material should be of use to others interested in casting a critical eye toward their own curricular assessment initiatives.

2. Background

Our department consists of three degree producing programs, but the primary concern of our investigation is

the assessment process for the computer science program. Our previous assessment process served the program very well, developed for our last accreditation cycle and used before the last ABET visit. However, we intuitively sensed that the process was not perfect and could always be improved

The next section presents a frank discussion of the potential pratfalls we identified during reflection on our previous assessment experiences. We began our reflection with the question: “*When does an assessment effort feel like a waste of time?*” and we looked for specific symptoms within our own assessment efforts and others that we were familiar with. Curricular assessment initiatives can have a lot in common with military training management, in which commanders at multiple echelons assess (and develop goals to enhance) the state of a unit’s training level. Therefore, we also drew from our experiences with military training management to infer problems that could arise in curricular assessment efforts.

3. Potential Pratfalls

Each potential pratfall is presented through a description of the pratfall itself, discussion of some symptoms stemming from the pratfall, and actions we have taken or strategies employed to avoid any negative effects. A complete overview of our assessment program is presented in Section 4.0.

3.1 Pratfall #1: Feeding a Zero-Defects Mentality

Senior leaders within a program might perceive an assessment effort *only* as a highly visible opportunity to showcase their program. Knowing that the fruit of their labor will be exposed to external critical review, they may only be interested in presenting their program in the best

¹ *Roget’s New Millennium Thesaurus* (2006) notes that “a *pitfall* is an unforeseen or unexpected difficulty or disaster - or a trap in the form of a concealed hole; a *pratfall* is an embarrassing or humiliating mistake, blunder, or mishap.”

light possible. Some program leaders may even relish the opportunity to receive the accolades of an accreditation committee, so their “assessment program” will tolerate “zero-defects.” In other words, no real problems will ever be seen by an external reviewer. The focus will be on ensuring that the program appears to exemplify excellence from every angle.

One obvious symptom of the zero-defect mentality is that no serious problems or limitations of the program are ever discussed or revealed in reviewable materials. Making programs look perfect takes a lot of work, and ironically, actual improvement efforts may be viewed as counter-productive. Another symptom of a zero-defects mentality are “dog and pony shows,” characterized by the highly-resented expectation that everything shows off perfectly. A final clue is when junior leaders within the program (such as course directors) adopt highly defensive postures about their accomplishments—the zero-defects mentality can be contagious within a program.

To ensure that our assessment process would stay clear of the zero-defects mentality, we have emphasized several points of our broader philosophy when introducing the assessment process to our faculty. The program is not perfect, and should not appear perfect. We do not expect courses or our students to be perfect either. The ability of Course Directors to analyze performance indicator data and draw meaningful, actionable conclusions is usually valued by the Program Director more heavily than achieving consistently high performance ratings. Such ratings could be indicative of material that is not challenging enough to students.

3.2 *Pratfall #2: Over-Assessing*

Each performance indicator (PI) is important, and it is natural to feel like the more efforts toward it we can show, the better off we will be. This tendency of over-accessing can result in key nuggets becoming lost in a sea of irrelevant data and questioning the cost of our efforts (see *Pratfall #6*). Others also warn against the dangers of over-assessment, such as [4].

One initial sign that over-assessment is imminent is that course directors claim to support an excessive number of PIs, potentially leading to multiple courses collecting data on the same PI. While a small level of redundant collection could be valuable, such efforts should only be planned to increase an assessor’s ability to draw meaningful conclusions. For a single course, there could also be too many planned collection events for a single PI.

To avoid this pratfall, our program scaffolds PIs and courses into groups that correspond to Bloom’s cognitive domain categories [2] (see Section 4.0 for details). Secondly, we use Course Monitoring Teams (CMTs) that are responsible for each group of PIs/courses. Prior to each semester, course directors make *claims* about which PIs should/will be measured during course execution and

provide a tentative data collection plan for each claim. The CMTs screen these plans from courses in their group (about 1-2 each semester) and adjust to ensure the most efficient data collection.

3.3 *Pratfall #3: Restricting Academic Freedom /Micromanaging*

Some may view assessment as overly prescriptive. A highly respected faculty member with many years experience teaching a certain course may view the assessment process as a restriction of academic freedom, forcing programs and courses to teach specific topics using certain techniques. Shouldn’t you be able to explore new concepts without jeopardizing your accreditation? Why should this professor be put under inspection?

What are the signs of pratfall #3? Seasoned instructors may be frustrated and non-supportive. Evaluation events may have little to do with what actually goes on in class. Students may be frustrated, claiming unfair grading. There may be a lack of change in the program over time; courses may not normally be updated to reflect current literature, technologies and practices.

In our program, we attempt to avoid this pratfall by adopting PIs contributed by input from all faculty, and then mapping the PIs to ABET/CSAB criteria for outcomes. We also allow the course director for each course offering to select the PIs they want to evaluate based on their coverage of the topic and pedagogy. Thus, we are measuring the knowledge, skills, and behaviors that *we* value. This enhances faculty buy-in as they are playing a major role in defining the assessment process.

3.4 *Pratfall #4: Perceiving Low Return on Investment (ROI)*

Regardless of whether or not an assessment program is useful at the program level, what does it really do for the average faculty member that is not involved in program administration? If we’ve successfully avoided pratfalls #1 and #2, then ROI will have nothing to do with showing off or receiving buy-in for their accomplishments. So can we convince our faculty that their time spent assessing is worthwhile?

When a strong assessment program loses steam over time, this is a clue that ROI is perceived as low. Faculty members who do not administer the program might devote less of their time. Perhaps a larger number of faculty members are complaining about the time taking by assessment activities.

We see our assessment program as an opportunity for junior faculty to receive advice and for senior faculty to be mentors. Every member of the faculty wants to grow and to promote growth as instructors, so we push our assessment process as a framework developed to provide mentoring relationships. The assessment program has multiple tiers that provide opportunities for increased

responsibility (e.g. a course director vs. course director in a data collection semester; a CMT member vs. the CMT lead). This potential for ‘upper mobility’ could be inspirational to new faculty and provide a recurring opportunity for relief to the seasoned faculty. We also select CMT members in a way that leverages a faculty member’s feelings as a stakeholder. Finally, assessment activities provide an opportunity to become familiar with other parts of the curriculum—perhaps even insights on future classes to teach—and collaborate with new faculty.

3.5 *Pratfall #5: Breeding Intellectual Incest*

Some folks see an assessment program as a process that makes us all teach mostly the same things, the same way. If we learn to communicate, compromise, and design complementary learning events, we may produce a better program on paper. However, will we be exposing our graduates to a wide variety of ideas that will prepare them for an ever-changing world? Will we be stamping out great new ideas in favor of the ones we are all more comfortable with? Will we even be capable of producing great new ideas?

This pratfall may be present within a program where assessment works like a well-oiled machine. In the early years, there may be none of the common discouraging symptoms—all faculty might be enthusiastic participants and meaningful changes might be made obvious from data evaluation efforts. But after a few years, there is very little change in the curriculum because we already ‘got it right.’ This may breed a feeling of disconnectedness with constituents or external colleagues working in the same area. Finally, the students’ perception of the program’s relevancy may be low.

We attempt to counter this pratfall several ways. First, we maintain strong connections with our constituents, especially through our research programs. In addition, we take our Advisory Board’s recommendations seriously in our decisions for program change. Also, we organize tough external reviews for our capstone events (coupled with internal ones). At West Point, we have about a third of our faculty rotate in/out every year from graduate school or from an Army unit (often from jobs that demand technology application and problem-solving skills). We include some of these people on the CMTs and in course director positions, while other CMT members and course directors are civilian professors or non-rotating military faculty that provide long-term stability. All of these methods for countering Intellectual Incest are underscored by a departmental culture that values diversity.

3.6 *Pratfall #6: Not Seeing the Forest through the Trees*

Program leaders can become besieged by the information overload involved in collecting and analyzing data. They may be unable to sense the need for programmatic changes,

since they are immersed in all the details of their assessment program.

What are some signs that you are being blinded at the strategic level by the details of the assessment operation? There may or may not be lots of good recommendations for changes in individual courses and adjustments in pedagogy. In fact, many of these recommendations might be coming from the Program Director, or other program administrators who are actively involved in the assessment process. There may be many other hopeful benefits of assessment actually observed—increased awareness, recognition of excellence, feelings of legitimate mentorship, etc. However, very few, if any, programmatic changes are ever proposed, discussed, or implemented.

In our program, we make a concerted effort to keep the Program Director removed from the CMT-level actions (except as an individual contributor). We place responsibility on the CMTs to sort out operational issues while the Program Director provides oversight. Finally, we devote two reports and at least one meeting each year exclusively to strategic analysis. For the purpose of this meeting, a CS Steering Committee is formed by the leaders of each CMT, the Program Director, and the Assessment Coordinator. Although each individual has a slice of operational experience within the assessment program, the Steering Committee focuses on analyzing outcomes that have (several) low-rated PIs to produce options for programmatic improvement.

4. Putting It All Together

Although we have discussed many isolated aspects of our assessment program, it may be helpful to the reader to review its full design. Like programs that assess agreement with ABET criteria, our program has *objectives* (11), which should be achieved by graduates within 5-7 years, and *outcomes* (9), which should be observable by graduation.

4.1 *Performance Indicators*

We observe/measure/evaluate student knowledge, skills, and behaviors (KSB) with performance indicators (PIs). A PI relates to a single outcome, but describes a concrete KSB at a particular learning level, 1-4². The set of our PIs are truly the central object within our assessment process, intended to fulfill several different goals. The PIs were generated during several working meetings with all faculty members invited, and later reordered and slightly revised by the CS Steering Committee. In concept, all PIs must be directly measurable, and the PIs for an outcome implement a scaffold (in the constructivist learning theory sense [3]) toward achieving the outcome (see Figure 1, Appendix A).

² Based on Bloom’s 6 cognitive domain categories (knowledge, comprehension, application, analysis, synthesis, and evaluation), where categories 1-3 are combined into our level 1.

Although the outcome is expected to be achieved by students before graduation, we can evaluate evidence of the PI scaffold at various points throughout the program. This practice ensures that contributions from introductory courses, as well as capstone courses, are valued during the assessment process. Furthermore, if an outcome is *not achieved* by students before graduation, our arrangement of PIs helps us realize the learning achievements they did reach and to focus our improvement efforts.

4.2 Course Groups

In total, we have 36 PIs. Each of the nine outcomes are measured with 2-6 PIs, which occur at a minimum of two learning levels. This complete set could be daunting to manage, so we have organized the courses in our program into learning levels as well. We refer to the arrangement of courses by learning level as a *course group* (see Figure 2, Appendix A).

4.3 Course Monitoring Teams (CMT)

Each course group has a CMT assigned to it, charged with evaluating the PIs that should be observable at their learning level. A CMT has a designated lead (a senior faculty member) and three other members that typically teach courses in the next higher course group. The CMT works with the course directors for all courses within the group, but only one or two courses undergo a data collection effort each semester. Each course provides data once during every two-year rating cycle.

4.4 Course Proposals & PI Support Claims

Prior to a given semester, the course director of each course offering submits a Course Proposal to their CMT. This document describes the vision for the course, and indicates which PIs could be measured during course execution (*support claims*). A tentative schedule of evaluation events for each support claim is also included. With this, the CMT helps focus a course director on creating a set of strong support claims to provide a complete set of PI-related data over the two-year rating cycle, while minimizing collective effect as much as possible. Course directors are expected to select about 2-5 of the set of 7-11 PIs associated with their course group.

Following a course execution (during which a course director collects data to validate his/her support claims), the course director and the CMT draw conclusions about whether those PIs were actually supported by that course offering. The course director and the CMT's ratings might disagree, but both are archived. The CMT then generates a three-level rating and provides program, outcome, course group, or course level suggestions.

4.5 Outcomes Assessing & CS Steering Committee

After each academic year, the four CMTs will have collected data to support ratings on at least half of the program's PIs. Each PI will provide a snapshot of student progress toward one of the outcomes, and some of the outcomes will have several of these snapshots at multiple learning levels. As part of the Annual Outcomes Review, the CS Steering Committee considers the CMT ratings for each PI evaluated that year rates each outcome to reflect the highest level of learning actually measured (see Figure 3, Appendix A).

While we have not settled on a "target level" for each outcome, our general thought is that if most outcomes were achieved at Level 4, a few Level 3 outcomes could be acceptable.

Beyond rating each outcome at the Annual Outcomes Review, the CS Steering Committee also seeks to provide recommendations for programmatic improvements, identify and address any PI evaluation abnormalities, and refocus the PI evaluation process for the upcoming academic year. Since the CS Steering Committee is composed of the leader of each CMT and Program Director, recommendations and decisions can be directly addressed—a continuous process to improve our program.

It is important to note the concepts for this assessment program have only recently been adopted. We began the identification process for PIs about a year ago, and stood up the CS Steering Committee to define and review the major aspects of the assessment program during the past six months. In the past month, we have stood up the complete CMTs, discussed the new process with the faculty at large, and begun our first semester of data collection under the new system. We are eager to share more of our reflections and findings with others.

References

- [1] Blaha, K. D. and Murphy, L. C. (2001). Targeting Assessment: How to Hit the Bull's Eye. In *Proceedings of the Consortium for Computing in Small Colleges: Northwestern Conference*, pp. 112-121.
- [2] Bloom B. S. (1956). *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain*. New York: David McKay Co Inc.
- [3] Driscoll, Marcy P. *Psychology of Learning for Instruction*, Allyn & Bacon, 2000.
- [4] Rogers, G. Death by Assessment: How Much Data Are Too Much? *Communications Link*. Spring 2002. <<http://www.abet.org/Linked%20Documents-UPDATE/Newsletters/Link/Spring2002.pdf>> (24 Aug 2006).

Appendix A

Outcome #5: An ability to use current techniques, skills, and tools necessary for computing practice.

Performance Indicators:

- 5.1 Solve a problem by integrating COTS hardware and software systems (*level = 1*)
- 5.2 Diagnose the cause of an unexpected result (*level = 1*)
- 5.3 Solve a problem using the specification of an unfamiliar software package (*level = 2*)
- 5.4 Develop a plausible explanation of observed system behavior (*level = 3*)
- 5.5 Solve a problem using an unfamiliar language or system (*level = 4*)
- 5.6 Creatively integrate COTS hardware and software systems (*level = 4*)

Figure 1. Performance indicators provide a learning level scaffold for reaching an outcome.

Course Groups

Foundations (Level 1)

Fundamentals of Computer Science
Database Design and Implementation
Digital Computer Logic
Introduction to Computer Architecture
Computer Science Seminar

Analysis Core (Level2)

Data Structures
Design and Analysis of Algorithms
Object-Oriented Concepts
Discrete Math

Grad School Prep (Level 3)

Operating Systems
Programming Languages
Computer Networks
Fundamentals of Computer Theory

Capstone (Level 4)

Software Systems Design I and II
Projects Day
Cyber-Defense Exercise
Summer internships

Figure 2. Courses arranged in Course Groups, also according to learning level.

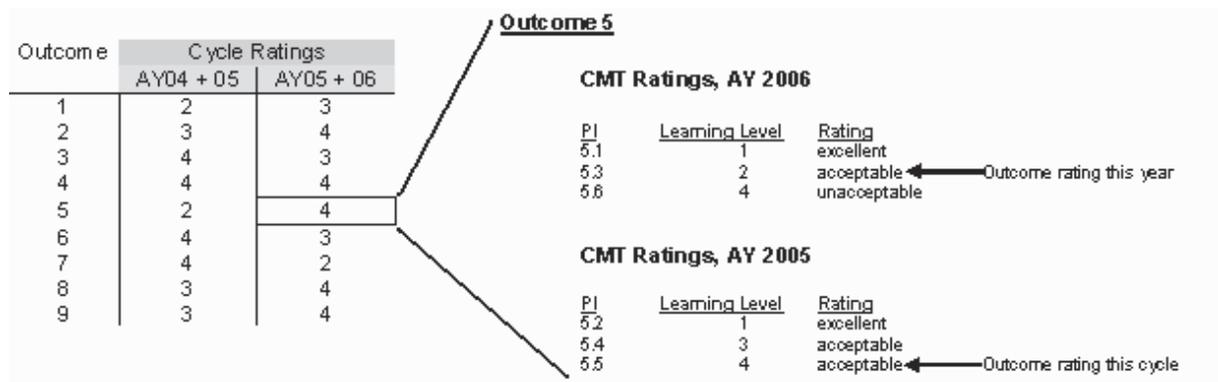


Figure 3. Outcome 5 is a Level 2 outcome in AY 2006, but a Level 4 outcome in the two-year rating cycle.