Negotiating the COAPRT Learning Outcomes Transition Using Quality Management Tools

A Case Study of the COAPRT Beta Test Site

Gary D. Ellis
Texas A & M University

Abstract

This paper is a case study. It tells the story of the process that the Council on Accreditation for Parks, Recreation, Tourism and Related Professions beta test site created its learning outcomes assessment program. A planning process was used that has evolved from quality management philosophy and practice: DMADV. Use of DMADV required precise definition of the problem, measurement of stakeholder needs, analysis of design options, design of the program, and verification of the program's efficiency and effectiveness. The resulting program was the basis for successful reaccreditation and it has met and exceeded expectations of the institution and the regional accrediting body.

Keywords: *learning outcomes assessment; DMADV; accreditation*

Gary D. Ellis is the head of the Department of Recreation, Park and Tourism Sciences, Texas A&M University. The author would like to express his appreciate to his friend and mentor, Dr. Peter A. Witt, who provided critical review of this manuscript and Dr. Erik Rosegard for his thorough and thoughtful editorial review. Please send correspondence to Gary D. Ellis, Department of Recreation, Park, and Tourism Sciences, Texas A&M University, 600 John Kimbrough Blvd, Room 417, 2261 TAMU , College Station, TX 77843-2261, Gellis@ag.tamu.edu

This paper provides a case study of a project directed at development of a learning outcomes assessment program. The learning outcomes assessment program that resulted meets institutional expectations that follow from the accreditation standards of regional accrediting bodies. The program was also successfully used in the beta test phase of development of the Council on Accreditation for Parks, Recreation and Tourism and Related Professions (COAPRT) 2013 accreditation standards. Project development followed a Lean Six Sigma design-for-innovation strategy known as DMADV (define, measure, analyze, design, verify). The DMADV planning model yielded an approach to learning outcomes assessment program that meets the needs of faculty, students, administrators, professionals and accrediting bodies for efficiency and effectiveness.

Context

The COAPRT transformation to learning outcomes that occurred between 2005 and 2013 was the most substantial revision of the COAPRT accreditation process since its inception in the 1970s. This change in focus from "what is taught" (curriculum content) to "what is learned" (learning outcomes) reduced the number of standards related to curriculum and instruction in parks recreation and tourism from over eighty to only four. Evaluation of programs for the purpose of COAPRT accreditation no longer requires reviewers to sift through extensive stacks of course syllabi to identify evidence of content required by one of the eighty-plus curriculum content standards. Instead, under the 2013 program, reviewers engage in thoughtful triangulation of evidence that students are learning the core content that is expected by the profession and by the professionals who hire graduates. This new COAPRT accreditation program is fully consistent with contemporary requirements and expectations of regional accrediting bodies and colleges and higher education deans, provosts, and presidents. The transformation is a success story.

But the success has not occurred without casualties, and substantial angst continues. Academic departments continue to struggle with designing and implementing learning outcomes assessment programs. At a macro level, these tensions are a function of the magnitude of the change and of the lack of academic preparation and training of park and recreation faculty in the disciplines most closely aligned with learning outcome assessment concepts and processes. Although virtually all park, recreation, and tourism professionals are prepared as social scientists, few have extensive experience or training in the specifics of designing and implementing learning outcomes assessment programs. To most faculty members, learning outcomes assessment is a foreign concept and is burdened with negative connotations. As Hutchings (2010, p. 9) points out, learning outcomes assessment "… echoes of… accounting, testing, evaluation, measurement, benchmarking, and so forth—language from business and education…it is striking how quickly assessment can be seen as part of the 'management culture' rather than as a process at the heart of faculty's work and interactions with students."

Indeed, many faculty and administrators continue to be suspicious of the increasingly pervasive "management culture." This culture of compliance, documentation, metrics, and accountability consumes increasing quantities of faculty and administrator time with each passing year. Learning outcomes assessment has quickly woven itself into the fabric of management culture, while continuing to be relatively foreign to most faculty and department chairs. Given their limited background in learning outcomes assessment, faculty and administrators have struggled with such questions as the following:

- What are learning outcomes? At what level of detail should they be written? How many learning outcomes should be written for a given academic program? How can we know if the set of learning outcomes we identify is sufficient to meet the three COAPRT standards (Foundations [knowledge of the scope and nature of the industry and essential history, philosophy, and social science], provision of services and experience opportunities, and management/administration) and other institutional requirements for assessment?
- What measures and processes do we need to put in place to assess the identified learning outcomes? How can we know if these measurement tools and processes are of sufficient quality to keep us in compliance with COAPRT standards and the expectations of our universities and colleges? Can data from these instruments actually inform quality improvement?
- How do we decide the acceptable levels of performance required to demonstrate that learning has occurred?
- What is the difference between direct and indirect measures of learning outcomes? How many of each type do we need to meet COAPRT and institutional requirements?
- What content must be included in if written assessment plans?
- What does it mean to "close the loop" in the assessment process? Are we obligated to make changes on every occasion that one of our measures falls outside a target level of learning?
- How can we make the benefits of the assessment program we create at least commensurate with the cost of creating, implementing, and maintaining that system?
- What ethical issues are associated with learning outcomes assessment? For example, can we evaluate faculty in terms of student learning? Can we use the same measures that we use to evaluate learning outcomes of the program to evaluate the performance of students in our classes?

As the COAPRT beta test site, the Texas A&M University Department of Recreation, Park and Tourism Sciences, addressed these questions using the Lean Six Sigma DMADV strategy (Define, Measure, Analyze, Design, and Verify). The following sections describe the process in detail.

DMADV Innovation Development Process

DMADV is planning process that has evolved from quality management philosophy and practice (e.g., Evans & Lindsay, 2008, 2009; Pyzdek, 2003; Taylor, 2009) DMADV is appropriate for situations in which no current process is in place for a given function, and the planner is committed to creating a product that addresses very specific needs of internal and external customers or other stakeholders. DAMADV can be used to create both intangible resources (such as processes) as well as tangible, "brick and mortar" resources. An example of use of DMADV for creation of an intangible resource is an employee performance evaluation system (Thatte, 2012). A "brick and mortar" example is the use of DMADV to design a new residence hall at the Pennsylvania State University (The Pennsylvania State University Office of Planning and Institutional Assessment, 2008).

Each letter in DMADV defines a step in a sequence: Define, Measure, Analyze, Design, and Verify. Each of the phases carries its unique objectives, and specific landmarks must be established within each phase to signal completion of that phase. Table 1 provides details

on the objective of each phase. The "Define" phase, for example, requires formal definition of the precise problem to be resolved, and establishment of goals for the overall project.

Table 1

DMADV Sequence

Phase	Full Name	Goal
D	Define	Establish a *precise* definition of the problem. Specify the purpose and goals, develop schedule and guidelines, analyze risk
M	Measure	Determine specifications; characteristics of the product that are "critical to stakeholders" and the relative importance of these
A	Analyze	Develop and evaluate design options
D	Design	Develop a detailed components and processes, optimize the design
V	Verify	Implement the program and evaluate its efficiency and efficacy

Define Phase: Learning Outcomes Assessment Project

Quality management philosophy and practice consistently emphasizes the pivotal importance of formal and precise definition of problems. Problem definitions should be highly detailed descriptions of circumstances. They should be data-based to the greatest extent possible and should include attention to specific costs and risks associated with the issue being addressed. In the case of the pilot test site of the learning outcomes assessment project, our problem definition was as follows:

> Our University and our professional accrediting bodies require us to create and maintain a program for ongoing assessment of learning outcomes and annual reporting of results. This program must be in place no later than June of 2010. The program must include an assessment plan, at least two direct measures of learning outcomes, and at least one indirect measure of learning outcomes. A written assessment plan is also required. Data must be collected and analyzed annually, and results must be entered into the electronic data base used by our university to document compliance, WEAVEonline. Finally, we must "close the loop" by documenting our use of assessment data to inform quality improvement of our curriculum and instruction. Failure to fully implement this program by June 30, 2010 will result in our department being out of compliance with our University (and Southern Association for Colleges and Schools) standards and out of compliance with the standards of our professional accrediting body, COAPRT. We are at risk of sanctions by the University and termination of accreditation by COAPRT.

The goal of the learning outcomes assessment program was as follows:

> To develop and fully implement a learning outcomes assessment program that meets the requirements of Texas A&M University and the COAPRT by June of 2010. That program will include a written assessment plan, instrumentation for assessment of learning outcomes, data from the first year of implementation, a report of assessment results, and a record of results in the University's electronic data base, WEAVEonline.

An action plan was also developed as part of the "define" phase. The plan listed the major tasks to be accomplished to achieve the goal, the target date for completing the tasks, and the designated person responsible for leadership on each task. An example of one of the tasks is "*Finalize selection of direct and indirect measures to be used in the assessment program*." Completion of the problem definition, the goal, and the action plan signaled the conclusion of the Define phase.

Measure Phase

The goal of the measure phase of DMADV is to craft a set of specifications for the product or process being developed. Measure refers to measurement or discovery of the needs of the stakeholders or "customers" who will be served by and benefit from the innovation. These needs become the product specifications. Lean Six Sigma stresses the vital importance of collecting and analyzing data to develop a firm understanding of these needs. Depending on the magnitude of the project, any of a variety of data acquisition and analysis techniques might be used. Among these are surveys, focus groups, check-sheets, and analyzing measured process outcomes and outputs through Pareto charts and the like. Lean Six Sigma practitioners also strongly endorse "going to the gemba" as an effective way of developing product specifications based on the voice of the customer. "Gemba" is a Japanese word that translates to, "the real place." Going to the gemba, then, refers to spending time on-site with people who will actually be using the innovation or will otherwise be impacted. Going to the gemba requires becoming a detective and ethnographer; observing, participating, asking questions, thinking, listening, analyzing, synthesizing, evaluating, and confirming.

Lean Six Sigma endorses a variety of tools for measuring the relative importance of specifications identified through these "voice of the customer" strategies. Analytic Hierarchical Process (Pyzdek, 2003), for example, involves evaluating each criterion against the other. Ratings that result are entered into algorithms that produce coefficients that reveal the relative importance of each criterion. Less sophisticated procedures may involve simply assigning different "points" to different options, perhaps while using one of the options as a reference point. A valuable source describing different methods for weighting criteria is Edwards and Newman's (1983) Multiattribute evaluation.

If we were to evaluate the efficacy of our implementation of DMADV in the learning outcomes assessment project, we would almost certainly conclude that our execution of the "measure" phase was incomplete. Our data sources were limited to members of the learning outcomes assessment task force (all were educators), a COAPRT Board member who happened to be a member of our faculty, and an authority on assessment from our University's Office of Institutional Assessment with whom we periodically consulted. In a more complete application of DMADV, we would have also listened to the voices of our

students, and we could have consulted with advisory group members, administrators of our college, and perhaps experience industry professionals. We did not evaluate the relative importance of our specifications. Through a series of meetings, the team shared opinions on the needs of these various stakeholders. Conversations were focused on the needs of students, faculty, the COAPRT, University administrators, and industry professionals. The five specifications listed in Table 2 resulted from these conversations.

Table 2

Assessment Program Produce Specifications

Our Learning Outcomes Assessment program must…	Abbreviated Name
1. …be *cost efficient* in data acquisition and analysis	Efficient
2. …yield *tangible curriculum quality improvement benefits* that are commensurate with costs	Commensurate benefits
3. … be *minimally intrusive* on the right of faculty to design and implement preferred approaches to curriculum and instruction that they believe are best for their students	Minimally intrusive
4. …be *compelling to COAPRT, University, and SACS*	Compelling
5. …not violate the criterion of *consequential evidence of validity*	Consequential Validity

Specification 1: Cost-efficient. The first specification is that the learning outcomes assessment program must be cost-efficient. Costs are the "elephant in the room" in most discussions about learning outcomes assessment. Few administrators want to talk about the costs associated with administrative mandates. Hutchings (2010) though, boldly unveiled the elephant, pointing out what few in the "management culture" of contemporary higher education have dared to mention: Despite the enormous investments that universities and accrediting bodies have made in learning outcomes assessment over the past decade, "… faculty have not yet seen sufficient evidence that assessment makes a difference… [their typical decision to not invest time, resources, and energy into] assessment may be a rational decision."

I agree with Hutchings (2010). I also agree with Churchill and Iacobucci (2010), who advocate rigorous evaluation of costs before embarking on expensive research. Will the information gained from the evaluation lead to vastly greater benefits to program success than if the program proceeds absent that information? Will the benefits be worth the costs? I shudder when I think of how learning outcomes assessment would fare if subjected to rigorous cost-efficiency evaluation. Although direct financial costs may be manageable, opportunity costs are staggering. Even after the enormously costly development process is complete, learning outcomes assessment continues to takes its toll. Data must be collected and, depending on the form of assessment, class time that might otherwise be devoted to more engaging or productive learning activities is likely lost. Data must be cleaned, entered, analyzed, and interpreted. Reports must be written. Meetings must be held in which results are discussed. Objections of faculty who are troubled by the data must be appropriately negotiated. Documentation of results and how they are being used must be entered into

potentially tedious electronic data bases, such as WEAVEonline. Hours of expensive faculty and staff time are required. And, at best, modest adjustments to curriculum and instruction may be the result. Many of those revisions would likely be made in the absence of learning outcomes assessment, based on the ongoing commitment of faculty to provide quality educational experiences for their students. Even when no assessment programs are in place, faculty members are very much "at the gemba." They are, in fact, "at the gemba" every time they deliver a lecture, lead a discussion, supervise a learning lab, provide leadership to a service learning project, or hold a private meeting with a student. Cost-efficiency was thus a key specification for our learning outcomes assessment program.

Specification 2: Tangible curriculum quality improvement benefits. The second specification was that the assessment program must yield tangible curriculum quality improvement benefits that are commensurate with costs. The phrase, "commensurate with costs" establishes the second specification as also being about efficiency (efficiency is defined as a ratio of outputs per unit of input). The intent of Criterion 2, though, was to focus on the potential for the program to generate information that can actually inform curriculum quality improvement decisions. As an example of what we wished to avoid, consider the ubiquitous student ratings that are routinely administered following each course. For the most part, these devices consistently yield high scores with little variation, regardless of what questions are asked. They are thus useful for ongoing monitoring of students' affinity toward their teachers and classes. The high scores tend to please faculty, affirm good teaching, and they provide a convenient, cost-efficient basis for evaluating teachers. But these instruments are limited in the extent to which they inform faculty of specific ways to improve their classes. The occasional lower scores give rise to multiple alternative explanations. A host of factors beyond the quality of the teaching itself may influence affinity ratings. Among these are class size, subject matter content, time of day of the class meetings, time with respect to the lunch hour, rigor of the course, experiential vs. didactic delivery, charisma of the faculty member, and number of minutes of meeting time per session. In contrast to student affinity assessment, our criterion that the assessment program must yield tangible benefits indicates that we sought to produce measurement processes that yield data from which we can make valid inferences about how our curriculum impacts our students' learning, at levels commensurate with our investment of time and other resources.

Specification 3: Minimally Intrusive. The third criterion is that the assessment program should be minimally intrusive on the ability of the faculty to design and implement preferred approaches to curriculum and instruction. We wanted to minimize use of assessment techniques that would impose requirements on how specific classes are taught. In particular, evaluation of portfolios and class-embedded assignments are sometimes touted as top choices for assessment programs.The portfolios approach to assessment would require faculty in several designated classes to require specific projects or assignments. As students progressed through their degree plans, these assignments would be graded, and then assembled into portfolios. Perhaps, for example, students would enter an essay on leisure theory from their "foundational knowledge" class, a marketing plan from a services marketing class, a budget from a management class, an event plan from a class on programming or event management, and responses to select case studies from classes on diversity, management, foundational knowledge, and experience design. As students approached graduation, a team of faculty and professionals would be convened to evaluate the portfolios using a formal rubric.

A portfolio-based learning outcomes assessment system would impose significant constraints on faculty who teach the classes in which the components of the portfolio were required. Given our previous example, whoever taught the foundational knowledge class would be required to include the leisure theory essay as an assignment. Perhaps a budget assignment would also be required in a management/administration class, and maybe the class on recreation experience design would require a risk management plan. We viewed such approaches as inappropriate intrusions on academic freedom of faculty. We embrace the value of allowing each faculty member who teaches a class the opportunity to design instructional approaches that she or he believes will best facilitate student engagement and learning. We also recognize that different faculty members may teach a given class during different semesters. Each should be free to craft her or his own assignments. Our product specification about intrusiveness was directed at preserving that freedom to the greatest extent possible.

Specification 4: Meet COAPRT, SACS, and Institutional Expectations for learning outcomes. The fourth criterion was that the learning outcomes must meet the requirements of COAPRT and our university's regional accrediting body, the Southern Association for Colleges and Schools (SACS). At the most fundamental level, acceptable learning outcomes are required. Learning outcomes can be written at a range of specificity. Suen (1990), for example, distinguishes between "molar" and "molecular" behavioral learning outcomes. A molecular behavior is an "individually recognized specific behavior for which further breakdown into components becomes meaningless" (p. 181). In contrast, molar behaviors are "categories of behaviors with meaningful component behaviors." Given that neither SACS nor COAPRT show a preference for either level of specificity, we followed the institutional practice of Texas A&M University, which is highly molar. To illustrate, all undergraduate students who graduate from Texas A&M University (including, but not limited to those who major in Recreation, Park and Tourism Sciences) are expected to achieve the following (molar) learning outcome:

Communicate effectively, including the ability to:

- Demonstrate effective oral communication skills (which could include the use of languages such as American Sign Language for those who do not communicate orally)
- Demonstrate effective writing skills
- Demonstrate effective nonverbal communication skills (which could include appropriate use of performance, design or representations such as maps, tables and graphs)
- Listen actively and critically
- Present work effectively to a range of audiences
- Effectively communicate original and creative ideas

Walvoord (2004, p. 123) presented molar-level learning outcomes in her highly influential book, Assessment Clear and Simple. One of her examples of learning outcomes of a theology department, for example, was

Theology undergraduate majors will

- Demonstrate awareness of the religious nature of the human person
- View theology as a science
- Conduct theological inquiry
- Describe and analyze the central areas and key issues of theology, especially in the Roman Catholic tradition

Our learning outcomes are presented in Table 3. Note that these are molar outcomes. Outcomes 2, 3, and 4 correspond directly to the three COAPRT accreditation standards in the 7.0 series. These standards require programs to produce learning outcomes in a) foundational knowledge, b) recreation and tourism experience planning and delivery, c) management of organizations that provide park, recreation, and tourism services, and d) a professional internship.

Specification 5: Consequential evidence of validity: Use the tests only for the purpose for which they are designed. During our design process, a member of our faculty pointed out a serious potential ethical concern: "How do I (as a teacher of a class in which learning outcomes assessment will be conducted) know that you will not use results of this assessment to evaluate me as a teacher?" That individual's concern illustrates the general principle advanced by Messick (1989) in a seminal paper on validity. Messick distinguished between evidential and consequential facets of validity of inferences that evaluators can make from test scores. Evidential validity corresponds to the descriptions that are pervasive throughout behavioral science research methods texts. It is the collection of evidence that a score derived from a measurement tool can reasonably be assumed to indicate the quantity of the phenomenon being measured. Included in consideration of the evidential basis of validity are content-related evidence of validity (content relevance and representativeness of a test) and criterion-related evidence of validity (predictive and concurrent evidence of validity and convergent and divergent evidence of validity).

While the evidential basis of validity has to do with empiricism and theory, the consequential basis of validity has to do with ethical uses of test scores. The general principle is that test scores should only be used for the purpose for which the test was intended. In the case of our learning outcomes assessment project, measures that we create are to be used to make judgments about the extent to which students are, in fact, achieving the learning outcomes that we have established. To use scores from a given test for another purpose, including performance evaluation of an individual faculty member, would be an ethical violation. To document our commitment to appropriate use of assessment measures, we include the following statement in our assessment plan:

> The educational principle of consequential evidence of validity establishes that validity of inferences made from test scores is inextricably bound to the intended use of a test. As such, it is vital to establish that tests used in this assessment process are to be used for the sole purpose of assessment of student learning outcomes. Although results of this process may imply that particular educational processes should be evaluated to sustain or improve learning, use of test scores for reporting in instructor performance evaluations or other unrelated or tangentially related purposes would comprise unethical use of the assessment process.
>
> Further, although results may or may not be part of the grading structure for individual classes, test scores will be considered information protected by the Family Educational Right to Privacy Act. As such, scores of individual students will not be published in any way, and no data will be provided that would enable identification of scores of any individual student. Students are, of course, free to use scores from tests that they complete in any way that they choose.

Table 3

Learning Outcomes for the Recreation, Park, and Tourism Sciences Major

- **Undergraduate Learning Outcome 1:** RPTS graduates will have reasoning, communication, diversity, and analytic skills appropriate to a strong general education.
  - **Reasoning Skills**
  - **Communication Skills**
  - **Diversity Skills**
  - **Analytic Skills**

- **Undergraduate Learning Outcome 2:** RPTS graduates will have acquired professional experience through the application of recreation, park, and tourism principles, theories and analytical methods by successfully completing professional internships.

- **Undergraduate Learning Outcome 3:** RPTS graduates will demonstrate knowledge of the scope of the profession, professional practice, and the history, scientific, and philosophical foundations of recreation, park, and tourism industries.
  - **History of parks, recreation, and tourism**
  - **Social science foundations of parks, recreation, and tourism (economics, psychology, sociology, geography)**
  - **Scope and nature of the park, recreation, and tourism industries**

- **Undergraduate Learning Outcome 4:** RPTS graduates will be able to design, plan, and implement recreation and tourism experiences among a diversity of clientele, settings, cultures, and contexts.
  - **Event Administration**
  - **Event Design**
  - **Event Marketing**
  - **Event Operations**
  - **Event Risk**

- **Undergraduate Learning Outcome 5:** RPTS graduates will be able to understand and apply profession-related principles and practices of management and administration.
  - **Management history**
  - **Management functions (planning, organizing, leading, directing, controlling)**
  - **Financial management for park, recreation, and tourism managers**
  - **Park, recreation, and tourism services marketing**
  - **Strategic management**
  - **Supervisory-level management**

Consensus on the set of specifications (Table 2) signaled the conclusion of the "measure" phase.

Analyze Phase

The analyze phase of DMADV involves evaluation of optional designs for the innovation. In our application, the task was to choose learning outcomes assessment tools that were optimal, given the specifications that we identified in the "measure" phase of our project. To proceed, we developed a list of possible direct and indirect measures, organized those into a matrix (an "X, Y chart") and then evaluated each measure according to each criterion (Table 4). An extensive list of examples of direct and indirect measures is provided by the WEAVEonline system (http://www.weaveonline.com/). We anticipated that the final set of 2013 COAPRT accreditation standards would require at least two direct measures. Based on our evaluation (Table 4), we chose to design our program around three major assessment tools and processes: pre/post tests (a direct measure), employability ratings by professionals who supervise our interns (a direct measure), and a senior exit survey (an indirect measure).

The pre/post tests were to be administered at the beginning and end of targeted classes that address the respective learning outcomes. We also recognized that, in the interest of the criteria of efficiency and minimal intrusiveness, the necessity for conducting the pretests would vanish as we collected sufficient numbers to achieve a strong estimate of the population pretest mean. We could administer the tests only at post-test, and compare the means against the stable population estimates. The employability ratings would be judgments by the professionals who supervise our interns. Those professionals are simply asked to rate whether or not they would hire each intern if funds were available for a position. The indirect measure of learning outcomes would be conducted through a survey of our graduating students, a practice that was already in place.

Performance specifications must also be established for each measure. For the pre/post tests, the criterion we chose was a 20% increase of the post-test mean, as compared to the pre-test mean (or the estimate of the population mean). That level of performance recognizes that many students will arrive at a new class with some degree of knowledge about the content that they have acquired in other classes, through personal research, or through professional experience. The performance standard for the employability measure (the ratings by professionals who supervise our interns) was that 90% would indicate that they would hire the intern if funds were available. For our student survey, we decided that an average score of 4 or higher on a 5-point scale would signal achievement of the learning outcomes measured through that instrument. The selection of these instruments and performance standards signaled the completion of the "analyze" phase of our development process.

Design Phase

The design phase of our project involved conceptualizing learning outcomes as a process, constructing the tests and exit questionnaire, and writing an assessment plan. The result of conceptualizing learning outcomes as a process is depicted in Figure 1. The process involves students taking sequences of classes and learning content related to each of the COAPRT standards. Students progress from foundational knowledge through recreation experience design, and management and marketing. Learning outcomes assessments are conducted at the completion of targeted classes, using the testlets. The process culminates

Table 4

"Analyze" Phase Worksheet

Direct Measures	Measure Selected?	Efficient	Commensurate benefits	Minimally Invasive	Compelling	Comment
Testlets, Pre-Post	Yes	H	H	H	H	Pretests can be discontinued after population mean is estimated, greater efficiency
Employability/Intern Evaluation	Yes	H	H	H	H	Routinely collected, a compelling goal of professional preparation programs
Portfolio Evaluation	No	L	H	L	H	Very labor intensive to evaluate, also may present threat to academic freedom
Capstone Assignment	No	L	M	L	H	Requires assignment that engages all Learning Outcomes
Comprehensive Exam	No	L	H	M	H	Could be administered as part of required internship experience
Standardized Test Result	No	L	L	M	L	CPRP Exam possible, but does not yield scores per learning outcome
Writing Exam Result	No	L	H	L	H	Must be evaluated by multiple faculty and topic must cover all learning outcomes
Class-imbedded assignments	No	L	L	L	L	Content representativeness is a major concern
Indirect Measures						
Exit survey	Yes	H	H	H	H	Surveys are widely accepted ways of assessing opinions
Alumni survey	No	L	H	L	L	A given alumnus may have little awareness of current curriculum and instruction
Curriculum review results	No	H	H	L	H	Periodic, e.g., every 7 years
Employer survey results	No	L	H	H	H	A given employer will have no exposure to most of the students
Focus Group Results	No	L	H	L	H	Very costly to administer, code and analyze data
Graduate School Acceptance Ra	No	L	L	H	L	Provides no evidence of attainment of learning outcomes
Transfer acceptance rates	No	L	L	H	L	Provides no evidence of attainment of learning outcomes
Advisory board evaluation	No	L	H	H	L	Advisory board will have no experience with the majority of the students
Placement data	No	L	H	H	H	Very difficult to collect and monitor placement
Benchmarking	No	L	H	H	H	Unlikely to be able to share identical outcomes data across institutions

H=High, M=Medium, L=Low performance with respect to the extent to which the criterion will meet customer/stakeholder needs.

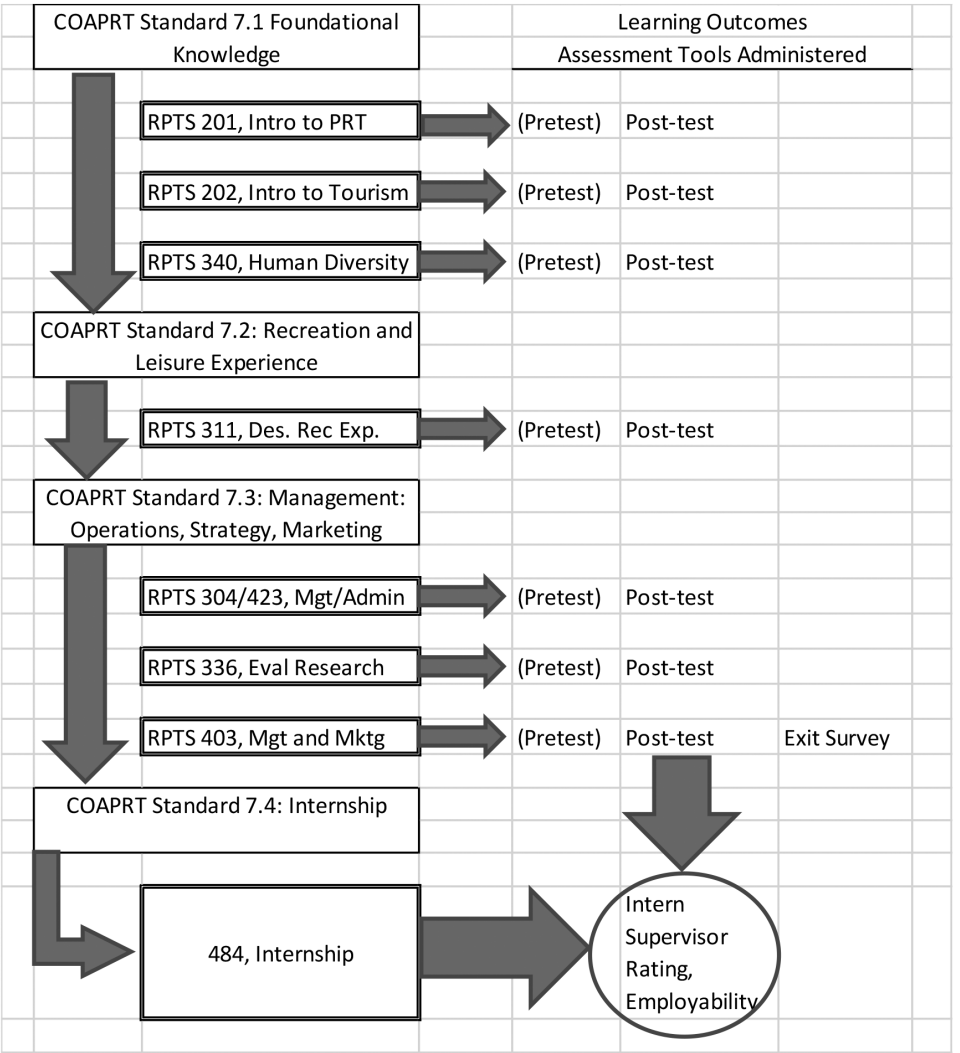

Figure 1. Design Phase: Learning Outcomes Assessment Process

in actual professional practice, via a 400-clock hour capstone internship experience. At the end of the internship, professionals who supervise the students submit the "process outcome" rating of student attainment of the learning outcomes, which is the employability rating of the interns. Also at or near the conclusion of the process, students use the senior survey to report their subjective opinions of the extent to which they have succeeded in achieving the learning outcomes.

The questionnaires and test that are used in the process were constructed using standard procedures for measurement tool construction. Virtually all university faculty in parks, recreation, and tourism are prepared as social scientists and thus have a degree of familiarity

with these procedures. For those who would like an excellent applied "refresher," I would suggest *Scale Development* by DeVellis (2012) and *Summated Rating Scale Construction* by Spector (1992). A large number of more sophisticated sources describe psychometric theory. The classics include books by Nunnally and Bernstein (1994) and Cronbach (1990). In addition to questionnaires and tests, COAPRT standards required that we design and develop a written learning outcomes assessment plan. Our assessment plan is updated each year and is organized into six sections:

I. Introduction
II. Ethical Use of Test Scores
III. Undergraduate Programs
IV. Graduate Programs
V. Assessment Calendar 201x

We also prepare a succinct written summary of assessment results each year. Those results are presented to our full faculty and to committees charged with continuous quality improvement of our undergraduate and graduate programs for action, as indicated and appropriate.

Verify Phase

The DMADV verify phase involves implementation of the program and evaluation of its effectiveness and efficiency. That process is ongoing. One indicator of the success of the program was successful reaccreditation by COAPRT. In the next year, our assessment program was judged to be in compliance with University (and SACS) expectations. During Academic Year 2013, our assessment program was highlighted as an outstanding strength in the final summary of our seven-year Academic Program Review. Our assessment program was, in fact, described as "among the best we have seen" in the final document forwarded to the State of Texas Higher Education Coordinating Board.

Yet, challenges remain. Our learning outcomes are stated in much more general terms than those that are being crafted by our colleagues at other COAPRT-accredited institutions. We embrace our approach as part of our commitment to minimal intrusiveness on the freedom of faculty to structure classes in ways they believe will best meet students' needs. Perhaps, though, COAPRT will ask us to be more prescriptive at our next reaccreditation review. Like all principles set forth by an accrediting body, as a department we would have to assess whether more stringent requirements meet our obligations to balance the requirements of the assessment process against its ultimate value.

Another challenge is ensuring quality of the testlets we use for the pre-post assessment. Several semesters of data collection have produced sufficient numbers of responses to provide a defensible estimate of the population mean for the pretest. As such, we have improved efficiency by discontinuing the practice of administering pretests. We simply compare post-test means with the pretest population estimates. But, issues remain. Testlets were not developed using "item maps" to ensure balanced coverage of content, and items were critiqued only by the individuals who constructed the testlet and by one member of the development team. Further, we did not formally define our learning outcomes to indicate the nature of the learning. Should we require recognition, use, or such discovery-level cognitive operations as analysis, synthesis, and evaluation (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956; Reigeluth, Merrill, & Bunderson, 1978)? We need to be more

specific in our definitions. We are in the process of evaluating our testlets and making improvements. When we have completed those improvements, it will be necessary to re-estimate population pretest means or to otherwise choose performance specifications. In addition, work is needed on our senior survey items to improve content-related evidence of validity by ensuring the relevance and representativeness of our test items.

We will also need to attend to the matter of "closing the loop." Closing the loop means using our assessment program to inform curriculum quality improvement. Interesting tensions exist here. Accreditation pressures from COAPRT and regional accrediting bodies indicate that performance specifications should be high enough to detect the need for curriculum improvements, yet when programs are evaluated, including through accreditation, meeting the learning outcome goals is a standard by which the efficacy of the program is judged. Thus, programs must design their assessment programs to both detect problems and to demonstrate success. Further, we will need to avoid the error of making constant revisions due to what may be random variation ("chance cause variation") in test score means for a specific academic term. To maintain efficiency, we must distinguish between special cause variation, which we can control, and chance cause variation (random variation), over which we have limited ability to control. As noted previously, these quality improvement actions add significant costs over and above the substantial enormous costs of the learning outcomes assessment process itself. We continue to be concerned about cost efficiency as we continue to invest time and resources in our learning outcomes assessment.

As is true of any case study, the circumstances at Texas A&M and in our department are unique. Procedures described in the case may not be appropriate for other programs, and perhaps the actions and approach we took would prove fully unsuccessful elsewhere. However, we did find the DMADV process to be helpful in creating our learning outcomes assessment program. Perhaps readers will find elements of the story helpful in designing or optimizing their learning outcomes assessment programs.

References

Bloom, B., Englehart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1 cognitive domain.* New York: McKay.

Churchill, G., & Iacobucci, D. (2010). *Marketing research: Methodological foundations.* Mason, OH: South-Western CENGAGE Learning.

Cronbach, L. (1990). *Essentials of psychological testing.* New York: Harper & Rowe.

DeVellis, R. (2012). *Scale development: Theory and applications.* Los Angeles: Sage.

Edwards, W., & Newman, J. (1982). *Multiattribute evaluation.* Los Angeles: Sage.

Evans, J., & Lindsay, W. (2005). *An introduction to Six Sigma and process improvement.* Mason, OH: Thomson South-Western.

Evans, J., & Lindsay, W. (2008). *Managing for quality and performance excellence* (7th ed.). Mason, OH: Thomson.

Hutchings, P. (2010, April). Opening doors to faculty involvement in assessment. National Institute for Learning Outcomes Assessment, Occasional Paper #4. Retrieved from http://www.learningoutcomesassessment.org/documents/PatHutchings.pdf

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement.* (pp. 13–103). New York: Macmillan.

Nunnally, J., & Bernstein, I. (1994). *Psychometric theory.* New York: McGraw-Hill.

Pyzdek, T. (2003). *The Six Sigma handbook*. New York: McGraw-Hill.

Reigeluth, C., Merrill, M., & Bunderson, V. (1978). The structure of subject matter content and its instructional design implications. *Instructional Science, 7,* 107–127.

Spector, P. (1992). *Summated rating scale construction*. Los Angeles: Sage.

Suen, H. (1990). *Principles of test theories.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Taylor, G. (2009). *Lean Six Sigma service excellence: a guide to green belt certification and bottom line improvement.* Fort Lauderdale, FL: J. Ross Publishing, Inc.

Thatte, D. (2012, January 23). DMADV Case study: Performance management system redesign. Retrieved from http://www.isixsigma.com/methodology/design-for-six-sigma-dfss-methodology/dmadv-case-study-performance-management-system-redesign/

The Pennsylvania State University Office of Planning and Institutional Assessment. (2008). DMADV: An approach for developing new initiatives. Innovation Insights. Innovation Insight Series Number 20. Retrieved from http://www.psu.edu/president/pia/innovation/

Walvoord, B. (2004). *Assessment clear and simple: a practical guide for institutions, departments, and general education*. San Francisco: Josey-Bass.