

# Using Pre- and Post-Testing to Assess a Criminal Justice Program Learning Outcome

BY DAVID A. MACKEY AND ROBERT E. FITZPATRICK



**P**rogram assessment, despite a surprising level of denial among some academics, is not a fad. While academia wallows in a plethora of literature promoting the value and importance of program assessment, anecdotal evidence supports a decades-long attitude—among faculty in particular—that assessment is at most a periodic concern receiving cursory attention during critical points in the regional accreditation cycle. Politicians and administrators may suddenly look to department chairs for some evidence of learning outcomes and value for tuition dollars. Deans and department chairs, then, sometimes desperately look to faculty for evidence of student learning outcomes. Committees suddenly arise to write reports supporting their departments’ imagined assessment gestalt. But despite sporadic publications in professional journals, do many campuses truly establish that *sine qua non* of “a culture of assessment” that accreditation and assessment literature repeatedly advocate as being of real value to administrators, faculty, and (most importantly) students?

Faculty must create a culture of assessment by providing evidence of both the level of student learning and students' satisfaction with the program. A culture of assessment implies continuous program improvement based on assessment results and corresponding curriculum adjustments based on the changing needs of the field. Evidence-based assessment permits faculty and administrators to abandon the unnecessarily defensive and all-too-common assumption that assessment need be a difficult, time-consuming, fault-finding exercise with no positive outcomes. In fact, many assessment options are neither overwhelming in scope nor onerous in application. Rather than highlighting departmental inadequacies, assessment may provide proof of success and frequently furnishes convincing justification for positive curriculum change and requests for additional resources.

This paper presents the assessment plan, results, and implications for one of a department's student learning objectives. The assessment of the first of six program objectives is part of a larger assessment plan for the department, but given space limitations, only the results of one objective are presented here. Goal one states that students will "demonstrate an understanding of criminal justice processes, terminology, and history" of the field. The point of our study, however, was to look at ways to assess student learning in terms of the value-added qualities of the Plymouth State University course The Criminal Justice System (CJ 1010). A central question to address was what students actually learned that they did not know before being exposed to the course. It is tempting to believe they start their academic careers as empty receptacles waiting to be filled with knowledge and experience, but where are the data to back up this belief? Criminal justice remains a popular field of study, and the pre-test/post-test approach helps discern how much discipline-specific knowledge students have when they begin their academic program. The real test of assessment, as detailed by Astin (1993), is therefore the measure of value added captured by the Inputs–Experience–Outputs (I-E-O) model of assessment. Astin advocates that all three elements—inputs, experience, and outputs—are essential for truly meaningful assessment.

### Literature Review

*Need for Assessment in Criminal Justice Programs.* In recent decades, the emphasis on assessment in higher education has moved away from indirect measures of educational outcomes toward more direct ones. In the past, student grade distributions,

---

A culture of assessment  
implies continuous  
program improvement  
based on assessment  
results and corresponding  
curriculum adjustments  
based on the changing  
needs of the field.

---

enrollment trends, time to degree, student diversity, institutional endowment, the number of faculty publications, and faculty-to-student ratios have all had their day in the sun as assessment indicators. Regarding the shifting sands of assessment, Myers (1994) observed that "[a]lthough most academics would agree that the quality of faculty research enhances both graduate and undergraduate programs, such productivity is not sufficient" (p. 31). As valuable as this information might be, it still does not answer the key assessment question. Nevertheless, the

criminal justice literature on the subject is far from abundant when compared with other disciplines.

One of the most persuasive articles on the value of program assessment is Finckenaue's (2005) paper "The Quest for Quality in Criminal Justice Education." Finckenaue presents an inspirational message projecting a promising future for a discipline that has had a dubious past. He quotes extensively from a 1985 *Boston Globe* exposé indicting a state funding program for criminal justice education known as the Police Career Incentive Pay Program—the so-called Quinn Bill. The Quinn Bill authorized a bonus for police officers who received a criminal justice degree from an approved program. The bonuses were 10 percent for an associate's degree, 20 percent for a bachelor's degree, and 25 percent for a master's or law degree. The ensuing investigation into the lack of rigor of a number of programs caused lasting embarrassment for the field and spurred political demands for accountability. Finckenaue, however, goes on to eloquently describe the good that came from the debacle, framing it as a wake-up call motivating the profession to undertake serious outcomes assessment, resulting in an increased impetus for certification in criminal justice promulgated by the Academy of Criminal Justice Sciences (ACJS). ACJS certification standards require that "[t]he program undergoes systematic evaluation of all program components and uses the results for program improvement" (Academy of Criminal Justice Sciences, 2005, p. 17). As is the case with regional accrediting bodies—in the authors' case, the New England Association of Schools and Colleges (New England Association of Schools and Colleges, 2011)—assessment requires supporting evidence to demonstrate student learning outcomes.

Departmental assessment plans should demonstrate a connectedness between the mission of the university or college and that of the department. Likewise, individual course objectives should

connect with the departmental mission and objectives. Bruns and Bruns (2009) state, “Essentially, all goals must stem from and reflect back to the mission statement” (p. 11). This is probably the most significant point when considering an assessment program or even a simple assessment project. The mission statement and department objectives, after all, may be the closest a department will come to a unified statement of purpose. It is a statement of the claims made for the expected outcomes of the program.

Once this connectedness is achieved, a department has flexibility in adopting a comprehensive strategy to assess various aspects of the program. For instance, empirical assessment generally falls into one of two types: indirect or direct measures. Indirect evidence is passive. It might include inferences derived from such data as grade point averages, focus group comments, surveys, or exit interviews, to name just a few of the possibilities. In most cases, these data may give no indication of what students are learning; such an inference would require direct evidence. Direct evidence includes standardized or locally developed tests, portfolios, or course-embedded assessment. One standardized test for criminal justice is the DSST exam (DSST, 2010).

**Criterion versus Benchmarking.** Interpreting assessment test results depends largely on the nature of the exam. For instance, a criterion-referenced test measures performance against an established baseline. The legal blood alcohol level is an example. Simply put, a criterion-referenced test is a test you either pass or fail. Some criterion-referenced tests have definitive results (such as blood alcohol content), while others might indicate a range (such as the academic grades A through F), but at some point the criterion-referenced test taker either passes or fails.

Alternatively, benchmarking is the major alternative to criterion-referenced tests and one that seems to be frequently misapplied in the popular culture—either purposely to achieve a political end or through ignorance of the procedure. Benchmarking is sometimes called peer referenced or norm referenced. When benchmarking, we are comparing the results from one test taker (or a group of test takers) to another. You cannot fail a benchmark; you can merely note how one score compares with another.

Astin’s (1993) seminal work is often seen as the key work for understanding value-added assessment and the I-E-O model. In addition, Peat and Moriarty’s (2009) discipline-focused, com-

prehensive work is the best choice specific to criminal justice. Peat and Moriarty present thoroughly developed chapters dealing with all relevant forms of assessment. The book also includes case studies and appendices of rubrics, forms, checklists, and resources to jump-start a fledgling assessment program.

Using pre-test and post-test scores at our institution would allow us first to see if students had achieved the objectives defined by our department learning outcome, and second to compare the results of successive cohorts. Assessment allows us to use the results by analyzing scores on subsets of questions to consider course design or even curriculum changes and to see if the changes are having the desired effect, as measured by improved test scores in future years. The benchmark (even though it is an internal benchmark) looked like a good starting point for part of a more comprehensive departmental assessment program. Having decided on direct benchmark assessment, we reviewed the criminal justice literature for examples of best practice and the experiences of others in the field.

#### Method

Plymouth State University enrolls about 4,200 undergraduates. The undergraduate criminal justice major began in the fall of 2003 with an entering cohort of 45 students. In only seven years, the program had grown by a factor of seven, employing five full-time and numerous part-time faculty serving more than 320 majors. The rapid growth of the program is probably not that different from the experience of the other nearly 100 U.S. public institutions with a comparable enrollment offering four-year programs in criminal justice (CollegeSource, 2012) What definitive evidence could we give to distinguish our program from any other—or, at the least, what valid claims could we make about the outcomes of our program?

Once our departmental mission and program objectives were developed, a comprehensive assessment plan was formulated. To measure the first program objective, we developed an assessment instrument, one that would help answer the assessment question, what evidence do we have that we are delivering the program we have advertised? Our assessment project was based on the mission and goals of the department and the application of Astin’s (1993) value-added I-E-O assessment model. Rather than outlining the entire assessment plan, this study reports results for the first of the criminal justice department’s six goals: “Demon-

---

The objective is to provide students with the foundation of a disciplinary level of understanding of the field that will prepare them for subsequent courses in the major.

---

strate an understanding of criminal justice processes, terminology, and history.” The Criminal Justice System, CJ 1010, is a required survey course for the bachelor’s degree in criminal justice and a prerequisite for many of the upper-division courses. It is expected that students who pass CJ 1010 will have knowledge of the criminal justice process to include the dual hierarchy of criminal justice; the formal stages of the criminal justice system; terminology; and knowledge of the history of police, courts, and corrections. The objective is to provide students with the foundation of a disciplinary level of understanding of the field that will prepare them for subsequent courses in the major.

Progress toward the goal was, and continues to be, measured by the administration of a pre-test and a post-test. The pre-test is administered during the June orientation to incoming students who have declared criminal justice as their major. The instructions for the instrument explain that the test is used for program evaluation purposes only and that the results will not affect students’ academic standing. The post-test is administered as a part of the CJ 1010 course during the final week of classes. Some instructors have elected to incorporate the post-test into their final exam, while others have used it as a review for the final exam. The exam consists of 40 objective items selected from both previous semesters’ CJ 1010 exams and based on key terms contained within the five leading introductory criminal justice textbooks. We noted whether the terms and concepts in the testing instrument were covered in the margin callouts, chapter key terms, or glossary of the textbooks. We have opted to begin assessment of the first learning objective with pre- and post-tests for the introductory course in the major. This approach allows the department greater control over the impact of transfer students and retention, which would impact the interpretation of results of a four-year follow-up for a post-test.

Six iterations of the exam were used for this analysis (spring 2006, fall 2006, fall 2007, fall 2008, fall 2009, and fall 2010). In situations where the standard post-test was not administered, such as when an instructor forgot to administer the post-test exam or a student was absent during the class period when the exam was administered, unadjusted mean scores (i.e., unscaled) for course exams were used as the post-test scores, since these were comparable to the format and origination of the standardized exam.

**Results**

A total of 462 students took the pre-test exam during the June orientations. The scores ranged from 10 to 90. The student with a score of 90 for the pre-test is a good illustration of the importance of the I-E-O model. Clearly, this student was not an empty receptacle upon his or her entrance into the program. The average score was a 40.8 (standard deviation 10.4). The scores on the post-test ranged from 18 to 100. Some first-year students declared criminal justice as a major after the June orientations. Some incoming students missed the June orientation sessions. Likewise, some students declared other majors after taking the pre-test as a declared criminal justice student at the June orientations. A total of 403 students took the post-test as enrolled students in CJ 1010. The mean score on the post-test was 64.2 (standard deviation 17.6). A total of 310 students took the pre-test as an incoming student at the June orientations and had a post-test based either on a true post-test or had the average of their exams in the CJ 1010 course. There was a statistically significant mean difference of 23.4 points (see Table 1) between the pre-test scores and the post-test scores. Just over 60 percent of the post-test scores were based on the standardized 40 question objective exam rather than the average of test scores in the course. The mean score for students who took the true post-test was a 57.8. The mean was 74 for the students for whom we had to use their average exam score in CJ 1010 as their post-test score.

Based on an analysis of the results and the measurement process for goal one, there was a statistically significant difference in the mean scores for students completing the post-test at the

**Table 1. Difference of means for pre- and post-test for matched students only**

Test	N	Mean	Standard Deviation	Standard Error	Difference	Significance
Pre-Test	310	41.3	9.9	.56	22.4	.00*
Post-Test	310	63.7	16.9	.96		

**Table 2. Difference of means for pre- and post-test content-knowledge exam**

Test	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval for Mean
Pre-Test (all scores)	462	40.8	10.4	.49	39.8–41.8
Post-Test (all scores)	403	64.2	17.6	.87	62.5–65.9
True Post-Test	243	57.8	14.49	.93	55.97–59.61
Average of Exams	160	74.0	17.34	1.37	71.27–76.65

conclusion of CJ 1010. Students' post-test scores were much better when their test score averages were used instead of the true post-test instrument, although both remained higher at a statistically significant level than the pre-test (see Table 2). We analyzed separately mean scores for all post-tests, true post-tests, and post-test scores based on an average of the students' exams in the course. Statistically significant differences in mean scores were identified for all three post-test scores. The mean for the true post-test was 17 points higher than the pre-test, while it was 33 points higher between the pre-test and the post-test score based on an average of exam scores. One likely explanation is that students are highly motivated to do well on exams but less so on an assessment instrument that does not affect their grade. One possible approach may be to integrate the post-test into the final exam and extract the scores for those items.

### Discussion

We would like to stress that this analysis reports progress toward only one of the department's six objectives. There are several recommendations for action. Although there is a statistically significant difference in the mean scores between the pre-test and post-test, the mean for the post-test is not great. One could interpret this score as a D on a traditional A–F scale. The department seeks to increase mean post-test scores to a mean of 70 to 75 using the standardized exam not only as a goal for our first departmental objective, but also to better prepare students for subsequent courses in the major. For continuity in the evaluation process, the department assigns one faculty member to monitor progress for each goal. Data have been collected for the last six years of incoming students. The instrument has been refined three times (changing a total of five questions) due to an examination of problematic items not covered in significant depth in every section of the CJ 1010 course. Revisions were done at the conclusion of the semester's post-test to avoid an *instrumentation* bias.

We have also recently added a variable to the database that indicates which instructor a particular student had for CJ 1010 (all instructors are full-time members of the department). At some point in the future, we will have the ability to compare scores on the post-test by instructor. There are many implications of such a difference if detected, although not all should be construed in a negative light. For some it may involve a reinforcement of the key terms and concepts for the instructor, while for others it may be recognized that teaching CJ 1010 may not be one of their strengths. In addition, we added an item to the pre-

test instrument to identify transfer students who already have taken the equivalent of CJ 1010 in order to allow comparisons between transfer students and first-year, first-semester students. This addition was done at the last iteration of the exam, and currently fewer than 10 students who have received transfer credit for CJ 1010 have taken the exam. In the future, the Department of Criminal Justice seeks to administer the content-area exam in CJ 4800 (Criminal Justice Seminar, our capstone course) as another data source for measuring progress toward goal one. The content-knowledge-based exam would be used to estimate the extent of knowledge and terminology retained by students during their course of study in the program. Comparisons would be possible between transfer students and students completing the bulk of their coursework in the program.

### References

- Academy of Criminal Justice Sciences. (2005). ACJS certification standards. Retrieved from [http://www.acjs.org/pubs/167\\_667\\_3517.cfm](http://www.acjs.org/pubs/167_667_3517.cfm).
- Astin, A. W. (1993). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. Phoenix, AZ: Oryx Press.
- Bruns, D., & Bruns, J. (2009). Decoding the complexities of assessment strategies in criminal justice education. *Academy of Criminal Justice Sciences Assessment Forum*. Retrieved from <http://www.acjs.org/pubs/uploads/ACJSAssessmentForumBruns9-2009.pdf>.
- CollegeSource, Inc. (2012). *CollegeSource online* [Data file]. Retrieved from <http://www.collegesource.org/home.asp>.
- DSST. (2010). Fact sheet: Criminal justice. Retrieved from [http://www.getcollegecredit.com/images/uploads/documents/Criminal\\_Justice.pdf](http://www.getcollegecredit.com/images/uploads/documents/Criminal_Justice.pdf).
- Finckenaue, J. (2005). The quest for quality in criminal justice education. *Justice Quarterly*, 22(4), 413–426.
- Myers, L. B. (1994). The evaluation of criminal justice programs: Assessment of evolving standards in context. *Journal of Criminal Justice Education*, 5(1), 31–48.
- New England Association of Schools and Colleges. (2011). Standards for accreditation. Retrieved from [http://cihe.neasc.org/standards\\_policies/standards/standards\\_html\\_version](http://cihe.neasc.org/standards_policies/standards/standards_html_version).
- Peat, B., & Moriarty, L. J. (2009). *Assessing criminal justice/criminology education: A resource handbook for educators and administrators*. Durham, NC: Carolina Academic Press.

*David A. Mackey is a Professor in the Criminal Justice Department at Plymouth State University in Plymouth, NH. Robert E. Fitzpatrick is the Emerging Technologies Librarian in Lamson Learning Commons at Plymouth State University in Plymouth, NH.*