

## HOW TO REDESIGN A COLLEGE COURSE USING NCAT'S METHODOLOGY

### VII. How to Assess Student Learning

The basic assessment question to be answered is the degree to which improved learning has been achieved as a result of the course redesign. Answering this question requires comparisons between the student learning outcomes associated with a given course delivered in its traditional form and in its redesigned form. There are two steps to achieve this goal: (1) establish the method of obtaining data, and (2) choose the measurement method.

#### **Q: How and when do you obtain the data?**

A: There are various ways to acquire the data.

- During the Pilot Term

This comparison can be accomplished in one of two ways:

#### 1. Parallel Sections (Traditional and Redesign)

Run parallel sections of the course in traditional and redesigned formats and look at whether there are any differences in outcomes—a classic "quasi-experiment."

#### 2. Baseline Before (Traditional) and After (Redesign)

Establish baseline information about student learning outcomes from an offering of the traditional format before the redesign begins and compare the outcomes achieved in a subsequent (after) offering of the course in its redesigned format.

*Note:* The number of students assessed should include at least 100 from the traditional format and 100 from the redesigned format.

- During the First Term of Full Implementation

Because there will not be an opportunity to run parallel sections once the redesign reaches full implementation, use baseline data from (a) an offering of the traditional format before the redesign began, or (b) the parallel sections of the course offered in the traditional format during the pilot phase.

The keys to validity in all cases are (a) to use the same measures and procedures to collect data in both kinds of sections and (b) to ensure as fully as possible that any differences in the student populations taking each section are minimized (or at least documented so that they can be taken into account).

## Q: What measures should you use?

A: The degree to which students have actually mastered course content appropriately is, of course, the bottom line. Therefore, some kind of credible assessment of student learning is critical to the redesign project.

Following are descriptions of four measures that may be used.

- *Comparisons of Common Final Exams.* One approach is to use common final examinations to compare student learning outcomes across traditional and redesigned sections. This approach may include sub-scores or similar indicators of performance in particular content areas as well as simply an overall final score or grade. (*Note:* If a grade is used, there must be assurance that the basis on which it was awarded is the same under both conditions—e.g., not curved or otherwise adjusted.)

### Examples

*Parallel Sections.* "During the pilot phase, students will register for either the traditional course or the redesigned course. Student learning will be assessed mostly through examination developed by departmental faculty. Four objectively scored exams will be developed and used commonly in both the traditional and redesigned sections of the course. The exams will assess both knowledge of content and critical-thinking skills to determine how well students meet the six general learning objectives of the course. Student performance on each learning outcome measure will be compared to determine whether students in the redesigned course are performing differently than students in the traditional course."

*Before and After.* "The specifics of the assessment plan are sound, resting largely on direct comparisons of student exam performance on common instruments in traditional and redesigned sections. Faculty have developed a set of common, objective questions that measure the understanding of key concepts. This examination has been administered across all sections of the course for the past five years. Results obtained from the traditional offering of the course will be compared with those from the redesigned version."

- *Comparisons of Common Content Items Selected from Exams.* If a common exam cannot be or has not been given, an equally good approach is to embed common questions or items in the examinations or assignments administered in the redesigned and traditional delivery formats. This design allows common baselines to be established. For multiple-choice examinations, a minimum of 20 such questions should be included. For other kinds of questions, at least two or three complex problems should be included.

### Examples

*Parallel Sections.* "The primary technique to be used in assessing content is common-item testing for comparing learning outcomes in the redesigned and traditional formats. Direct comparisons of learning outcomes will be obtained from 15 common complex problems embedded into course assessments: 5 early in the semester, 5 at mid-semester and 5 in the final examination in both the traditional and redesigned courses."

*Before and After.* "The assessment plan will address the need to accommodate a total redesign. The plan calls for a before/after approach using 30 exam questions from the

previously delivered traditionally-configured course and embedding them in exams in the redesigned course to provide benchmarks for comparison."

- *Comparisons of Pre- and Post-tests.* A third approach is to administer pre- and post-tests to assess student learning gains within the course in both the traditional and redesigned sections and to compare the results. By using this method, both post-test results and value-added analyses can be compared across sections.

### Examples

*Parallel Sections.* "The most important student outcome, content knowledge, will be measured in both redesigned and traditional courses. To assess learning and retention, students will take: a pre-test during the first week of the term and a post-test at the end of the term. The faculty, working with the evaluation team, will design and validate content-specific examinations that are common across traditional and redesigned courses. The instruments will cover a range of behaviors from recall of knowledge to higher-order thinking skills. The examinations will be content-validated through the curriculum design and course objectives."

*Before and After.* "Student learning in the redesigned environment will be measured against learning in the traditional course through standard pre- and post-tests. The college has been collecting data from students taking this course, using pre- and post-tests to assess student learning gains within the course. Because the same tests are administered in all semesters, they can be used to compare students in the redesigned course with students who have taken the course for a number of years, forming a baseline about learning outcomes in the traditional course. Thus, the college can compare the learning gains of students in the newly redesigned learning environment with the baseline measures already collected from students taking the current version of the course."

- *Comparisons of Student Work Using Common Rubrics.* Naturally occurring samples of student work (e.g., papers, lab assignments, problems) can be collected and their outcomes compared—a valid and useful approach if the assignments producing the work to be examined really are quite similar. Faculty must have agreed in advance on how student performance is to be judged and on the standards for scoring or grading (a clear set of criteria or rubrics to grade assignments). Faculty members should practice applying the criteria in advance of the actual scoring process so as to familiarize themselves with them and to align their standards. Ideally, some form of assessment of interrater agreement should be undertaken.

### Examples

*Parallel Sections.* "Students complete four in-class impromptu writing assignments. A standard set of topics will be established for the traditional and redesigned sections. A standardized method of evaluating the impromptu essays has already been established and will be used in grading each assignment. The essays are graded by using a six-point scale. The reliability measure for this grading scale has been established at 0.92. Additionally, each paper is read by at least two readers. The grading rubric will be applied to the four standard writing assignment prompts administered in parallel in simultaneously offered redesigned and traditional course sections."

*Before and After.* “The assessment plan is quite sophisticated, involving both before/after comparisons of student mastery of statistics concepts in the traditional course and the redesigned course. The design itself involves direct comparisons of performance on common assignments and problem sets using detailed scoring guides (many of which were piloted and tested previously and are thus of proven utility). Because the department has already established and benchmarked learning outcomes for statistics concepts in considerable detail, and uses common exercises to operationalize these concepts, the basis of comparison is clear.”

**Q: Should the assessments be different from those used in the course?**

A: We strongly recommend that you avoid creating add-on assessments for regular course assignments such as specially constructed pre- and post-tests. These measures can raise significant problems with student motivation. It is easier to match and compare regular course assignments.

**Q: How can we be sure that the students in parallel sections are equivalent if they have not been randomly assigned?**

A: If parallel sections are formed based on student choice, it would be a good idea to consider whether differences in the characteristics of students taking the course in the two formats might be responsible for differences in results. Final learning outcomes could be regressed on the following: status (full-time versus part-time), high-school percentile rank, total SAT score, race, gender, whether the student was taught by a full-time or part-time faculty member, and whether the student was a beginning freshman.

**Q: Are there other comparisons that would be useful to the redesign effort?**

A: In addition to choosing one of the three measures described earlier, the redesign team may want to conduct other comparisons between the traditional and redesigned formats such as:

- Performance in follow-on courses
- Attitude toward subject matter
- Deep vs. superficial learning
- Increases in the number of majors in the discipline
- Student interest in pursuing further coursework in the discipline
- Differences in performance among student subpopulations
- Student satisfaction measures