How’re We Doin’? The Power of Assessment: One Simple Assessment Tool Can Produce Multiple Applications

Charles R. Pastors

The Motivating Question

Northeastern Illinois University [NEIU] is a public institution on the Northwest side of Chicago, serving over 10,000 students, many of whom are returning adults or first in their family to attend college. On its website, NEIU describes itself as one of the most diverse universities in the nation and as the only four-year public Hispanic-Serving Institution in the Midwest. Almost 20 years ago, as chair of the political science department, I was interested in trying to find out “how we were doing” in our basic American National Government course (PSCI 216). With almost 20 sections taught each term, it was a major time and resource commitment of the department. It served to satisfy a university general education requirement (and, as such, was probably the only political science course most students would experience during their undergraduate program), and as an entry level prerequisite to the political science major. It was taught by all levels of faculty—junior/senior tenure track faculty and adjuncts—and at the central campus and at satellite locations.

I was motivated by a desire to test two very common faculty and administrative perceptions—that tenure track and “senior” faculty were “better” instructors than adjuncts; and that students at the satellite campuses (largely non-majority populations) were “not prepared” for regular college course-level offerings, and, thus, had to be treated with lower expectations for success and lower standards for achievement. Using the standard end-of-course student evaluation commonly administered in every class wasn’t coming close to addressing my question.

The Assessment Design

Returning from a Chautauqua Short Course taught by Dr. Ruth Day of Duke University on “Cognition and Learning,” I felt I had been handed a conceptual tool to investigate both areas—Day’s model of the growth from amateur to expert (that is, the difference between a beginning piano student and Ramsey Lewis or Artur Rubinstein). While the amateur/expert development continuum
might seem similar to Bloom’s Taxonomy, there are differences in how individuals move upward along the scale.

I devised a pilot pre/post assessment instrument to be administered at the immediate start of the first class of course 216 and at as near the end of the course as possible. On the pre-course instrument, students were asked what subjects they thought they would be studying in the upcoming course. (“No looking in the textbook for chapter headings!”) On the post-course questionnaire, they were asked about the subjects they had studied in the course being completed. (Again, “No peeking at class notes or other materials.”) On both instruments twenty spaces were provided for students to list the subject names down the left side of the response page. Students then were asked to take the subjects they had listed and group or cluster any they thought were somewhat related. They were to write them together on the right side of the page and circle them, or otherwise indicate that they had been grouped. Analogies to these groupings were given as those in a grocery store--with produce, dairy, and deli, --or in a hardware store—with plumbing, painting, gardening, and electrical. About five minutes were allowed for each task.

It was strongly emphasized to participating students that we wanted to know “how we were doing” as their instructors, so we could find out how we could do better and improve for our future students. Thus, we asked them to put their student IDs on their response sheets, so we could see how well we did with each of them as individuals. Further, we strongly emphasized that their responses would remain totally confidential. As suggested by Jillian Kinzie (2006), students are more likely to participate when they understand the purpose of the survey and how the results will be used. The more tailored the invitation to participate, the more likely students are to accept.

I had anticipated that, while students might be somewhat or moderately motivated to help us, they would be highly motivated to help other students like themselves, who would be taking future classes with us. Over several sections during several terms, over 95% of students asked to participate took the exercise seriously and accurately identified themselves on the response sheets—allowing pre/post individual comparisons. (The instrument, and the few sample responses mentioned below are available on my website at www.questforstudentsuccess.com.)
I also had several expectations:

1. Some students would have some experience with or awareness of how the political process of government in the United States works. For those students, the length of items would be longer on both the pre-course and post-course responses. They would be more likely to identify at least rudimentary groupings of items—such as campaigns, elections, voting—before taking the class. On the post responses, their item lists would remain long, with somewhat more complex examples included, and their groupings would be a bit more involved and coherent—such as primary/general elections, voter turnout, electoral college. These student would receive an “A” or “B” in the course, perhaps without expending too much effort.

2. Other students would have less experience/awareness of American politics. Their initial topics list would be short and relatively simple—such as government, voting, taxes—with little or no grouping of topics. By the end of the course, their topic list would grow, with some increased complexity, and with at least basic groupings of individual items occurring. The divergence of the pre/post responses were expected to correlate with student course grades—students “learning more” would (generally) be rewarded with higher course grades.

For the instrument design to be completed, the subjects students in the pilot project would list in their responses would have to be given some sort of complexity/sophistication score. If on a scale of 1-5, “voting” was a 1, and “electoral college” was a 3, 4, or 5, what would be the score for “congressional district,” “one-person-one-vote,” or “primaries”? After the pilot run of this assessment, if the expected general response patterns appeared, I planned to ask two or three department members who regularly taught sections of 216 to join me in scoring the various terms students in the pilot had provided. The student pre/post “item scores” then could be use as a marker for “learning” during the course. I hoped (and expected) that “learning” and course grades would correlate, especially for students in group (2) above. I also hoped that “learning” would be comparable across sections—whether led by “senior,” “tenure track,” or “adjunct” faculty, and whether on the main or satellite campuses.

Pilot Assessment Findings

The pilot instrument was administered in several 216 sections over three terms to almost 200 students:
1. Over 90% of students with a high number [15-20] “expect to study” items in their pre-course response received a B or an A in the course.

2. In student group (2), above, lower to higher number of items listed in the “expect to study”/”have studied” strongly followed the pattern of grades received. The greater the difference in the length of the pre/post response lists, the higher the grade earned in the course.

3. The increase in the ability to identify and group related items and the complexity of the groupings also matched grading patterns.

4. Findings (1), (2) and (3) were consistent across sections—indeedendent of instructor status or campus location.

Five examples from one section of 216 are summarized in Table 1 (See www.questforstudentsuccess.com for the actual responses.

### Figure 1

<table>
<thead>
<tr>
<th>Items in Pre List</th>
<th>Items in Post List</th>
<th>Change in Item Number</th>
<th>Course Grade</th>
<th>Brief summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sam</td>
<td>6</td>
<td>8</td>
<td>2</td>
<td>D</td>
</tr>
<tr>
<td>Angelo</td>
<td>5</td>
<td>15</td>
<td>10</td>
<td>C</td>
</tr>
<tr>
<td>Juanita</td>
<td>14</td>
<td>20</td>
<td>6</td>
<td>B</td>
</tr>
<tr>
<td>Sally</td>
<td>14</td>
<td>20</td>
<td>6</td>
<td>A</td>
</tr>
<tr>
<td>Juan</td>
<td>8</td>
<td>15</td>
<td>7</td>
<td>A</td>
</tr>
<tr>
<td>[class averages]</td>
<td>10.8</td>
<td>18.2</td>
<td>7.4 (70% change)</td>
<td></td>
</tr>
</tbody>
</table>

Analysis, Limitations, Suggested Final Steps, and General Application

The pilot results supported the claim that:

1. Students in 216 were moving up the early stages from amateur to expert—
2. For the students in group (2), above, student “learning” was being reflected accurately in the grades received. It was “knowledge added” that was being measured and rewarded in the grading process, rather than some exterior standard of “What you need to know [What score is needed] to get a ‘B’.” Students were “learning,” at least the basics of the course subject.
3. (1) and (2) were occurring in 216 sections across variations in the status of the instructor, or the physical class location.
Since this was a pilot project, no statistical analysis was performed on the student responses. With a scale scoring of student responses, and with a larger number of students responding, it would be possible to report findings with more certainty. For example, a student learning differential might be discovered between classes with new or experienced faculty, or where faculty had “lowered” expectations for the ability or preparedness of some student groups.

I left the department to become acting dean of the college, and retired two years later. The department did not follow up on this pilot. However, this assessment case study continues to provide a model of how a pre/post in-class assessment activity, consuming only a total of 20 minutes of class time, can address multiple topics of interest to any department or institution:

- Is student learning taking place in our classes?
- Is learning taking place under all instructors, and in all locations?
- If learning is uneven, what might be associated with any variations?

The greatest asset of the assessment model used here is that it can be applied in any discipline and at any level. Additionally, it is cheap, relatively non-intrusive to the instructional classroom setting, and provides multiple insights into class, course, and program performance. The role of the department, the department chair, or the dean becomes that of demonstrating the political will needed to “close the assessment loop” by doing something about “how we’re doin’” by “doin’ som’thn” where improvement is warranted. And that, when it comes to any assessment, is usually where the shortcoming lies.

Reference
*Assessment Update* 18 (2): 4-6.

Charles R. Pastors is professor emeritus at Northeastern Illinois University and principal at Quality Undergraduate Education for Student Transformation (QUEST) in Oak Park, Illinois.