When Classroom Assessment was used in an introductory survey course, students learned more and their instructor received higher ratings.

Classroom Research in Psychology: Assessment Techniques to Enhance Teaching and Learning

Charles J. Walker

Frank Sinatra has been quoted as saying that the difference between a professional and an amateur is that an amateur cannot sing to empty tables. My interpretation of this quotation is that to be a professional singer, one must be more influenced by the music than by the audience. Professionalism in college teaching often seems similarly defined. The quality and accuracy of a lecture's content allegedly reveal more about the professionalism of an instructor than does his or her effect on or responsiveness to students. The singer metaphor seems particularly relevant to large introductory courses. The professor is at the microphone, the students are seated at a distance; the most natural thing for any instructor to do under these circumstances is to "sing."

The "teacher as singer" metaphor apparently works well for some faculty. These instructors consistently give inspired lectures. As a consequence, they receive high ratings from peers and students alike, and their students probably do learn a lot. However, after teaching Psychology 101 for several years, I began to realize that "singing" was not effective for me and that, in terms of my goals as a teacher, I had more in common with Mitch Miller than with Frank Sinatra. I wanted to sing along with, not to, the audience. Unlike my other courses, introductory psychology had, over time, become less fulfilling to teach. Negative indicators such as decreased class attendance, lower participation levels, and most significant of all, lower student test scores, suggested that I had problems to solve in Psychology 101. Relying on my skills as a behavioral scientist, I decided to
initiate a program of Classroom Research to identify those problems and search for solutions consistent with my teaching style.

One obvious difference between introductory courses and other courses is class size. In large introductory classes, it is more difficult to continuously monitor student performance (Johnson, 1988). In larger classes, it can be nearly impossible to learn students’ names, let alone the details of what, how, and why they are learning (Walker, 1980). In this Classroom Research effort, I chose to use Classroom Assessment Techniques (CATs) to monitor student performance and to collect other forms of student feedback. What I heard was not always music to my ears, but the information students shared did help me discover simple, efficient ways to improve learning.

Research Goal and Questions

While most college courses stress the learning of course content (Cross, 1988), this is particularly true of introductory courses. Introductory courses present instructors with the nearly impossible task of teaching the basic subject matter, history, and methods of an entire discipline. Students are faced with a correspondingly monumental learning task. With the demands of this sort of course in mind, I chose to focus on assessing and improving my achievement of one fundamental goal: increasing the number of my students who master the subject matter of psychology. I was also interested in discovering answers to questions such as “Which topics are more difficult to learn?” and “What effect does my teaching have on student evaluations?”

Planning Classroom Research

When student learning is assessed only by so-called objective tests—such as multiple-choice tests—the feedback to students is summative and evaluative, not specific and formative. By this I mean that the primary purpose of objective tests is to evaluate for the purpose of assigning a grade, while the primary purpose of Classroom Research is to bring to light specific adjustments needed to improve teaching and learning.

My essential goal was to increase the number of students who mastered the content of Psychology 101. I believe that for students who do not receive specific, formative feedback, an introductory course of this type can become a kind of “dilute lit mansion” of facts and theories. Therefore, for the benefit of my students and myself, I made a commitment to add as many “lights” to this “mansion” as my time and energy would allow. I planned to do this by using assessment techniques at least twice in each of the course’s five units, and more often in those units where the test scores of previous students had been consistently low.

I was concerned that the assessment techniques not be “plug-ins,” “add-ons,” or mere data-generating devices. I wanted to ensure that these techniques would encourage the development of metacognitive skills—the awareness and management of one’s thought processes—in students and that they would assist their learning. Decisions about which techniques to use and how to use them had to be guided by practical conceptions of student learning.

Evaluating Assessment Techniques
Within the MORE Model

Many theories and models in psychology and other disciplines have key concepts that can be incorporated into assessment techniques. Doing so is efficient because, in the process of assessing, concepts are presented that both inform and help organize memory. Pintrich (1988) and Weinstein and Mayer (1986) suggest useful models of student learning. From these sources, I abstracted application guidelines summarized by the acronym MORE (motivation, organization, rehearsal, and elaboration). I made sure that each CAT used conformed to the MORE guidelines. A summary of these guidelines follows.

Motivation. Good assessment techniques should enhance and stimulate student motivation. They can foreshadow the content of tests and hence capitalize on existing strong motives such as the desire to get good grades. But they can also tap sources of motivation such as social affirmation, which may result when CATs are shared or done in groups rather than alone. Or they can satisfy the higher cognitive needs of individuals by appealing to curiosity, allowing creativity, or reinforcing competencies that are part of a student’s sense of self-worth.

Organization. Assessment techniques should help students organize learning tasks. Specifically, useful CATs should help students better organize information, thereby enhancing its learning, efficient storage, and ready access.

Rehearsal. Practice, known as “rehearsal” in psychology, is an important step in reinforcing learning. A good assessment technique should give students an opportunity to practice what they have recently learned. This can be accomplished by applying a CAT after learning has occurred or by repeating a technique used earlier in learning.

Elaboration. The ability to generalize and make connections, referred to as “elaboration,” is the process of connecting new knowledge to previously learned knowledge. Effective assessment techniques should encourage students to actively connect new knowledge or concepts with previously learned concepts and knowledge stored in long-term memory. Elaboration is more likely to occur when students are asked to rephrase ideas in their own words or to illustrate new concepts with examples drawn from their own personal experiences.
Choosing and Applying Assessment Techniques

Because my essential goal was to increase the number of students who mastered the content of Psychology 101, I decided to use CATs in ways that promised to increase the amount of students’ readily accessible knowledge. CATs were used before lectures to encourage the active organization of knowledge and at the end of lectures to encourage the elaboration and rehearsal of knowledge. Finally, to arouse student motivation, I invented techniques that were relevant, challenging, and enjoyable.

Like instructors in other introductory behavioral science courses, I discovered that CATs such as the Memory Matrix (Cross and Angelo, 1988, p. 26) and Defining Feature Matrix (p. 44) were very useful. I also found adaptations of CATs like Focused Listing, Documented Problem-Set Solutions, Concept Maps, and Process Self-Analysis to be helpful (Cross and Angelo, 1988, pp. 22, 38, 68, and 116, respectively). Of all the assessment techniques I used, the two that produced the most dramatic increases in student learning were the Memory Matrix and the Concept Map. In each instance I adapted the technique, using MORE as an instructional design sequence to strengthen its capacity to encourage learning as it fulfilled an assessment function. The following section provides detailed examples of how I applied these two techniques.

An Application of the Memory Matrix Technique. Areas of knowledge within disciplines vary in conceptual elegance. Older areas of inquiry are frequently associated with concise, simple conceptualizations. The Memory Matrix illustrated below can be readily used as both a teaching technique for these elegant areas of content and as an assessment tool. For example, in psychology, I found the Memory Matrix very applicable for both teaching and assessing students’ learning of the basic principles of operant conditioning.

As described by Cross and Angelo (1988), the Memory Matrix technique is a two-dimensional matrix—a rectangle divided into rows and columns in which the row and column headings are given but the cells are empty—that students use to organize information into categories. I used the following eight-step “recipe” to present the Memory Matrix technique:

Detailed Instructions for Using a Memory Matrix

Step 1: Motivation. Arouse student motivation by giving real-life examples of operant phenomena—in this case, provocative instances of the contingencies of positive reinforcement, negative reinforcement, punishment, and penalty. Give titles or nicknames to each of your examples; use these titles to draw attention to the target behavior to be modified.

Step 2: Organization. Supply a memory organization device by drawing a 2 × 2 matrix of the operant principles, like that given below, on the chalkboard.

<table>
<thead>
<tr>
<th>Contingency Manipulation</th>
<th>Target Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increases</td>
<td>Decreases</td>
</tr>
</tbody>
</table>

When you present stimulus:
When you remove stimulus:

Step 3: Rehearsal. Ask students to copy the matrix reproduced above four times: twice on a sheet of paper you provide and twice again in their notebooks. To save time, some instructors may prefer to pass out copies of the matrix. However, having students draw their own may be worth the extra time because it causes them to be more actively involved in learning.

Step 4: Elaboration. To begin the elaboration process, ask students to write the titles or nicknames of the examples you gave them in the appropriate cells of the first copy of the matrix.

Step 5. Extend and clarify the concepts by providing the technical definitions of the operant principles. That is, define positive reinforcement, negative reinforcement, punishment, and penalty. Then ask them to write these terms in the appropriate cells of the second copy of the matrix.

Step 6. Next, have students compare the two models to identify any inconsistencies. That is, they should check to see if the example of positive reinforcement in the first matrix is in the same cell that positive reinforcement is in in the second matrix. This step will serve as a rehearsal event for some learners and a learning check for others. Once they have determined that the two matrices “make sense,” collect the sheets. These will serve the instructor as an assessment of his or her teaching effectiveness.

Step 7. On the chalkboard, reveal how you thought the exercise should be done. Then ask students for examples of other ways of doing it.

Step 8. After agreement has been achieved, ask students to write in the names of the types of contingencies on the third copy of the matrix, one of two copied in their notes. Finally, ask them to supply their own examples of these contingencies in the cells of the fourth matrix. This allows students to again rehearse and elaborate the concepts that they have just acquired.

An Application of the Concept Map Technique. Concept Maps are student drawings or graphic illustrations of the ideas, concepts, or theories they have learned (see Cross and Angelo, 1988; Novak and Gowin, 1984). I found that Concept Maps help introductory students learn facts and principles that are nested within complex theoretical structures. As Cross and
that the two schools were similar, or to draw larger buildings or more buildings if they believe a given school was more influential than others.

Step 3. To maintain motivation and encourage more elaboration, invite the students to add features and create places where their schools' protagonists and followers would enjoy living. This instruction opens the doors to student creativity, yet elicits responses that are not too difficult to interpret. Students might add clever features, such as a Gestalt Circle, a Wundt Introspection Center, a Thorndike Puzzle Palace, a Freud Phallic Fountain, or, one of my favorites, a William James Stream of Consciousness.

Step 4. Collect their maps with the promise that they will be returned the next class. During your analysis of their feedback, write supportive and corrective comments on each student's map. Hand the maps back for inclusion in the students' notes. Share photocopies or chalkboard drawings of good and poor maps and explain your judgments, then invite students' reactions.

Feedback for Learners

Because it is difficult to give attention to individuals in large classes, I decided to give each student performance feedback whenever feasible. In doing this I hoped to become more familiar with my students as unique learners and people. However, I did not always ask them to sign their CAT responses, and sometimes, when little time was available, I reviewed only half of the students' replies. I found data based on random samples of 25 to 50 percent of the class to be quite reliable for assessing the whole class's understanding.

After I had gathered the students' responses, to gain the benefits of immediate feedback I wrote an example on the chalkboard of the kind of reply I had expected. I then opened up a short discussion by asking, "Did anyone see another way of doing this?" At the beginning of the next class, I returned the individual response sheets with two kinds of comments on them. Students got feedback on content—that is, on their grasp of facts and concepts—and feedback on learning processes—that is, on their study tactics and learning strategies. For instance, with the CAT on operant conditioning, when a learner confused negative reinforcement with punishment, I would comment on how the two contingencies are, in fact, both similar and different, suggesting ways the student might distinguish between the two. Several Classroom Researchers in psychology have suggested very effective ways to teach students to distinguish punishment from negative reinforcement using a matrix (Flora and Pavlik, 1990).

If I detected problems in learning skills such as note taking, listening,
or reading, I suggested practical ways to improve these skills. Or, if a learner gave evidence of problems with basic learning strategies, I suggested ways to change their learning habits. For example, in the case of elaboration problems, I simply urged students to try generating several examples of each concept from their own experience.

Feedback for the Instructor

As I reviewed the students’ replies, I kept a record of the number of them who made errors on key points. When more than a fifth of the students had trouble understanding a concept, I considered this sufficient evidence to warrant a change in my teaching. The first action I took was to ask them to clarify their replies. Then if a new way to communicate the lesson occurred to me at the moment, I immediately retaught the material. If not, I took the problem home and returned with a new approach to the material the next time. My overriding concern was to increase the amount of content shared and learned, not merely “covered.”

Some student feedback had implications for my role as an instructor in domains other than classroom teaching performance. For example, we identified conceptual errors in the text or conflicts between text and lecture materials. Other Classroom Assessment feedback suggested that I could improve their learning by simply reducing the amount of material covered in my lectures. And, as I will discuss in more detail later, student feedback had implications for evaluation procedures, student-teacher role relations, and classroom norms.

Project Evaluation

In evaluating this project, both learner outcome and instructor outcome measures need to be taken into account.

Learner Outcome Measures. The principal goal of the project was to increase the number of students who mastered the knowledge and methods of psychology. To assess how well this goal was achieved, I compared the test performance of students taught without CATs against the performance of those taught with CATs. Specifically, I subjected the total test scores of students taught in 1988–89 and 1989–90 to tests of statistical significance. Because the course content, texts, and examinations were similar across these two academic years, I felt a comparison could be useful.

Assessment, Feedback, and Practice Result in Better Learning. The results were provocative. On average, students taught with CATs scored significantly higher on tests—at the .001 level of statistical significance—than students who learned without the benefit of CATs. The total test means of these two groups were 86.3 percent and 76.5 percent, respectively. In addition, while only 18 percent of the 1988–89 students received A or B grades on tests, 59 percent of the 1989–90 students received A or B test grades. Students taught with CATs scored higher on each of the course’s five units than did students taught without CATs. I observed these positive effects of Classroom Assessment on learning with students who took the course in the fall of 1989 and again with a second group of students who took the course in the spring of 1990.

Consideration of Other Variables. Two confounding variables—retesting and student entry characteristics—must be considered before these results can be confidently interpreted. The 1989–90 students were given the opportunity to take “recovery quizzes,” and the points they earned by mere guessing on these quizzes may have artificially inflated their test scores. After each unit exam, students were given the option of answering five new questions on the material that unit tests had diagnosed as being the most difficult. Recovery quiz points were retroactively added to students’ test totals. Overall, approximately half of the 9.8 percent improvement in mean test score totals of the CAT-instructed group can be attributed to the effect of recovery quizzes.

While improved student performance might have been the result of mere guessing on the recovery quizzes, not new learning, the data suggest that this was not the case. Even assuming that students earned one out of a possible five points per quiz by chance, the mean difference between the 1988–89 and 1989–90 grades was still statistically significant at the .05 level. It appears that assessment, additional feedback, and additional practice resulted in better learning. Hardly a surprising outcome, perhaps, but a very heartening one.

Finally, if students taught with CATs had been more academically able than students taught without CATs, the test performance difference could be accounted for by differences in student entry characteristics. However, the converse was observed. When the cumulative grade-point averages (GPAs) of the two groups were compared, the 1988–89 students had a slightly higher mean GPA (2.79) than the 1989–90 students (2.71).

Instructor Outcome Measures. Other researchers have found that there is a strong positive correlation between students’ learning and their summary ratings of a course and its instructor (Howard, 1984). Specifically, students tend to rate courses and instructors more highly when they believe they have learned a lot and/or when they have received high grades. What I observed in my student ratings is in accord with previous research. Students rated both the course and my performance as an instructor highly. All of the instructor and course ratings exceeded 4.0 on a Likert 5-point scale. And 5 was the most commonly given rating on all but one of the scales. Only the prompt “Outside of class, the instructor provided assistance on an individual basis when needed” received a modal rating of 4.

The ratings of the 1989–90 students were high, but were they higher than those of students who learned without CATs? Unfortunately, reliable
data were not available to answer this general question. The rating forms were changed after the 1989 academic year. However, a direct comparison was possible for one unchanged item. On the item "Overall, as it was presented by the instructor, a student can learn a lot from this course," the 1989–90 students gave higher ratings than the 1988–89 students; the means of these two groups were 4.58 and 4.00, respectively. This difference was statistically significant at the .001 level.

These results, taken together with other, less easily quantifiable feedback from students, such as written comments, thank-you notes, letters sent to the dean of arts and sciences, and increased class attendance, strongly indicate that, in general, the 1989–90 students learned more and were more satisfied in the course than their predecessors. In particular, students expressed satisfaction with my use of CATs.

Discussion and Implications

The results of this Classroom Research project support the conclusion that using CATs can increase student learning. Most important, I found that when assessment techniques are used as part of a program of Classroom Research, they stimulate change in the classroom environment and guide a process of course development. Because assessment techniques involve students in their own learning and invite them to share control of the classroom, classroom norms and student-teacher role relations change.

Changing Student-Teacher Role Relations. My role and the role of students began to change when they suggested that the assessment process be expanded to include the unit examinations. They asked that the unit tests be used to detect their learning problems, as well as to evaluate their performance and to assign grades. Specifically, they recommended that I give them the opportunity to be retested on the subjects they had not learned well; they wanted to prepare for and take what they called "recovery quizzes." As has been observed by other instructors (Davidson, House, and Boyd, 1984), I discovered that retesting students improved their learning.

Adopting the Role of Learning Coach. As has often been noted, two essential steps in good course planning are formulating course goals and teaching objectives, and sharing this information with students to guide their learning. In compliance with these two principles, I had identified my course goals and teaching objectives and had included a general description of Classroom Assessment procedures in a detailed, thorough course syllabus for Psychology 101. Moreover, I had informed students several times during the first week of class that assessment techniques would be used to teach key topics and that there would be a correlation between the CATs used in class and examination questions. "When you do an assessment technique, you'll help me discover better ways to teach, and at the same time you'll learn the kind of information you should know on unit examinations," I had explained.

This set the stage for shared control over learning. Later, as a result of complying with the students' request to give them recovery quizzes, several other productive changes occurred. For example, during testing I adopted the role of a learning coach, rather than that of academic police officer. This cooperative spirit continued into the test review sessions. Instead of engaging in confrontation, students and I found ourselves cooperating, trying to identify the most troublesome topics and exploring ways to understand and teach that which had not yet been learned or taught.

Acting on Student Advice. Perhaps the single most useful piece of advice I received from students was to reduce the amount of material covered in lectures. I needed their advice in all but one of the course’s five units, the one on social psychology—the unit I am best prepared to teach. Because social psychology is my area of expertise, I felt I should emphasize this topic and “cover more material.” Apparently I “covered” the students, not social psychology. This was the only unit in which I observed little or no difference in student learning between the classes taught with and without assessment techniques. This result does not support the oft-made assumption that doing disciplinary research is the best way to prepare for teaching. As other Classroom Researchers have discovered, I found that there was no direct relationship between my research expertise in the discipline and my teaching effectiveness (Feldman, 1987).

In my case, this may have happened because I equated my ease in teaching the material with the ease of learning it. My overconfidence caused me to do less assessment in this unit, and when I did assess it, I used techniques that were very difficult for students. As a result, they gave me feedback that was confusing and of limited use for planning corrective changes in my teaching. Nonetheless, it is now clear that the unit on social psychology should be a focal point for my next Classroom Research project in Psychology 101.

Assessment Techniques Gave New Life to the Course. Students were not the only ones who benefited from my use of assessment techniques. I discovered new life in a course that had become very routine. Because my students expressed a desire for more active, participatory forms of learning, I changed many of my lectures to respond to their needs. For example, instead of simply lecturing on Freud's theory of personality structure, I had students role play the id, ego, and superego in a dating-game simulation. In the role play, three students portrayed the id, ego, and superego as bachelors or bachelorettes on the "Dating Game" television show. I used this demonstration as an assessment technique. During their role plays, I provided feedback on their faithfulness to Freudian theory and took mental notes on how my preparatory lecture had affected their learning. From this experience, I derived guidelines for converting many of my other lectures into assessment and active-learning opportunities.
Using CATs in a program of Classroom Research is a means to discover more effective teaching and learning methods. Because it advances instructors and students toward common goals, it is mutually rewarding. Most important, Classroom Research can dramatically improve student learning. And, as this teacher has discovered, it is far more fulfilling to sing along with full tables than to empty ones.

References


By carefully watching and listening to his students explain how they solved simple physics problems, this instructor shed light on how students “see” physics and how to help them see it more clearly.

Classroom Research in Physics:
Gaining Insights into Visualization and Problem Solving

David M. Nakaji

My interest in and experiences with Classroom Research have focused on discovering and understanding the cognitive processes underlying students’ visualization and problem solving in physics. Visualization, as I will use the term, is the formation of “mental pictures,” images generated and manipulated by the human mind. The link between visualization and problem solving has always fascinated me, and my experiences, both as student and as instructor, have led me to believe that visual images play a paramount role in producing meaningful solutions to problems.

Do students generate and access mental pictures when they go about attacking problems? If so, what kind of pictures are they? What do they look like? How are they being used and how frequently? How important are they? These were some of the questions that motivated me to seek an approach to inquiry that would at least partially, if not fully, disclose the answers. Classroom Research has proved to be such an approach for me.

My first exposure to the concept and methods of Classroom Research came when K. Patricia Cross and Thomas A. Angelo of the University of California, Berkeley, initiated a pilot program at Los Medanos College. That pilot program, co-sponsored by a grant from the Ford Foundation and Pew Charitable Trusts, was designed to train a small cadre of our instructors in Classroom Research. I volunteered to participate, welcoming an opportunity to finally engage in activities that promised to be directly related to my interests and concerns within the classroom.

With the benefit of hindsight, I can see that even my questions regard-