

Journal of Teacher Education

<http://jte.sagepub.com>

Teaching under High-Stakes Testing: Dilemmas and Decisions of a Teacher Educator

Rosemary E. Sutton

Journal of Teacher Education 2004; 55; 463

DOI: 10.1177/0022487104270186

The online version of this article can be found at:
<http://jte.sagepub.com/cgi/content/abstract/55/5/463>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

On behalf of:



American Association of Colleges for Teacher Education (AACTE)

Additional services and information for *Journal of Teacher Education* can be found at:

Email Alerts: <http://jte.sagepub.com/cgi/alerts>

Subscriptions: <http://jte.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

TEACHING UNDER HIGH-STAKES TESTING DILEMMAS AND DECISIONS OF A TEACHER EDUCATOR

Rosemary E. Sutton
Cleveland State University

In this article, the author reports on the curricular changes and testing focus that have occurred in her teaching since PRAXIS II: Principles of Learning and Teaching (PLT) tests were mandated in Ohio, where she teaches at an urban, open admissions, commuter university. Classroom dilemmas are analyzed as they relate to changes in assessments, curriculum content, and instructional strategies when teaching educational psychology to preservice teachers. The benefits and costs of the increased importance of educational psychology in the teacher education program and the advantages and disadvantages of trying to ensure that the students do well on someone else's examination are also discussed. The author's experiences are compared with published reports on teachers' reactions to high-stakes testing, and the author considers how the context in which she teaches and her personal teaching biography influenced her decisions.

Keywords: *high-stakes tests; PRAXIS II; educational psychology; teacher education; urban teachers*

During the 1990s, the use of standardized achievement tests for high-stakes certification of beginning teachers increased rapidly. In 1998, the reauthorization of Title II of the Higher Education Act (P.L. 105-244) mandated that institutions report pass rates of teacher education program completers on teacher tests. The majority of states use the PRAXIS I and II series of tests developed by Educational Testing Service (2003a).

The increased use of high-stakes tests for beginning teachers as well as K-12 students is part of widespread educational reform taking place in the United States. This reform has occurred among grand narratives of national decline in economic competitiveness, safety, and educational achievement (Cochran-Smith & Dudley-Marling, 2001). Reversing the decline, it is claimed, requires an imposition of new standards for instruction, curriculum,

teacher training, and teacher licensure (Yinger, 1999).

Some research indicates that the standards imposed through high-stakes testing of K-12 students narrows curriculum to test-driven content and basic skills (e.g., Haney, 2000; Smith, 1991). Teachers can become "testing coaches" (Sacks, 1999) less likely to use innovative instructional practices such as cooperative learning, whole language, and higher order thinking activities, and they can become angry and fearful at the perceived loss of control (Cimbricz, 2002). However, other research suggests that although the tests may influence what teachers teach, they do not influence how they teach (Firestone, Mayrowetz, & Fairman, 1998). In addition, the grade level, subject matter, status, experience, and local context all influence teachers' reactions to high-stakes state testing (Cimbricz, 2002).

Research on the impact of high-stakes tests for beginning teachers on teacher education faculty is limited. Cochran-Smith and Dudley-Marling (2001) found no curricular or program changes in five institutions they studied when high-stakes tests for beginning teachers were introduced in Massachusetts in 1988. However, resources shifted to discussions about the tests and what to do about them. In addition, the high failure rate and the secrecy surrounding the content of the test contributed to morale problems of faculty and administrators as well as the grand narrative of what was wrong about teaching and teacher education. Ludlow, Shirley, and Rosca (2002) classified institutional investment in preparation for the teacher tests in Massachusetts as heavy at besieged institutions with very low pass rates, moderate at institutions with borderline pass rates, and minimal at institutions with high pass rates.

This article reports the curricular changes and testing focus that have occurred in my teaching at Cleveland State University since PRAXIS II: Principles of Learning and Teaching (PLT) tests were mandated. Reports of my experiences are augmented by the notes I routinely make for class preparation, changes in my syllabi and assessments over time, students' end-of-semester course evaluations, and theory and research in motivation. Three classroom dilemmas faced in teaching educational psychology to preservice teachers are analyzed: how much to change assessments, curriculum content, and instructional strategies. Also discussed are the benefits and costs of the increased importance of educational psychology in the teacher educational program and the advantages and disadvantages of trying to ensure that students do well on an external examination. In the final section, institutional responses to the implementation of the PRAXIS II tests are outlined. I begin by describing the context of these changes: the high-stakes test, the institution in which I work, and my own teaching experiences.

CONTEXT

In spring 2003, 16 states and the Department of Defense Dependents Schools required that

teacher-education candidates pass PRAXIS II: PLT to obtain a license (Educational Testing Service [ETS], 2003a). At that time, the PRAXIS II: PLT tests consisted of three case studies each followed by seven multiple-choice and two constructed response questions. Twenty-four additional multiple-choice questions were included. The time limit was 2 hours, and there were three versions of the test: Grades K-6, 5-9, and 7-12 (ETS, 2001). States determined their own minimum passing scores and, in Ohio, the cut scores were among the highest in the nation: 168 for both K-6 and 5-9 tests and 165 for 7-12.

Most of the content in this test is that traditionally addressed in educational psychology courses (Sudzina, 2001). For example, topics include motivation, cognitive strategies, instructional methods, child development, assessment, and evaluation. I had taught educational psychology to preservice teachers for more than 20 years so this high-stakes test raised crucial issues about the content, assessments, and teaching strategies used in the course.

At Cleveland State University, all undergraduate students must take an educational psychology course, and five or six sections are offered every semester. Two psychology courses (Introduction and Child or Adolescent Psychology) and one education course (Introduction to the Profession) are prerequisites for this course, but during the period I am writing about, the prerequisites for this course were not systematically reinforced because of software problems in the registration system. Students are encouraged to take their education courses in a specific sequence but class schedules depend on work commitments and on what courses are open during registration. Thus, students in one section of an educational psychology class have a wide variety of prior experiences, vary in class standing from sophomore to postgraduate, seek licensure in a range of areas, and concurrently take a mixture of classes.

Cleveland State University is an open admissions, commuter, and urban university enrolling more than 14,000 students. Each year, more students transfer from community colleges or other universities than enter as freshman and

about an equal number of students enroll part time versus full time. Currently, 56% of the students are women, 22% are minority, the average age of undergraduates is 27 years (Cleveland State University, 2003), and 95% of students live off campus (USnews.com, 2004).

Recent data indicate that Cleveland State students are more economically disadvantaged than those at other urban universities and so work more hours, have more family commitments, take longer to graduate, and spend more time commuting and less time on studies outside of class (Indiana University, 2002). Some students take one or two semesters off during their teacher education program to earn enough money to pay debts and save for future tuition. Even after 20 years of teaching in this institution, I am in awe of the sacrifices and dedication of the students. Gaining a teaching license can be a step into the middle class, a way of retooling skills in an economy that is no longer industrial, an economic solution to single parenthood, or the fulfillment of a lifelong dream to help children. I feel an enormous responsibility to ensure the sacrifices of the students are worth it; that is, I do everything I can to help them learn, develop, and achieve their goals. One of these goals has now become passing PRAXIS II: PLT.

Many of the students are not good standardized test takers. College admissions tests are not required and during 1999-2002 the percentage of Cleveland State University students retaking the K-12 PRAXIS II: PLT tests because they failed the first time was nearly twice as high as the state average.¹ Equity concerns associated with standardized tests frequently dominate faculty and student discussions. Recent national data indicate that students of color fail the PRAXIS II tests at a higher rate than do White students. For example, although 83% of White students exceeded Ohio's qualifying scores on the K-6 PLT, only 51% of African American students, 65% of Hispanic students, and 64% of Asian American students did (Sutton, 2003b). These differential pass rates are likely to further exacerbate the shortage of teachers of color in local urban school districts

that physically surround the university and where many of the graduates teach.

The implementation of PRAXIS II: PLT also raises many concerns about lack of validity evidence. What evidence is there that this test predicts successful teaching? What are the consequences of widespread implementation of such tests? Do they lead faculty and students to focus on the wrong aspects of teaching? I have written about equity issues in the past (e.g., Sutton, 1991, 1996, 1997; Sutton & Fleming, 1994) and this influenced how I framed the dilemmas and made decisions. For example, knowing that level of inequality often varies by institution contributed to my belief that teaching does make a difference.

According to Zancanella (1992), the biography of the teachers' past experiences is an important force in the ways they respond to state-mandated tests. I grew up and taught mathematics in New Zealand more than 25 years ago when passing national high school curriculum examinations was necessary for success in high school. I learned to teach under the demands of a national curriculum and did not think it an imposition because I then believed such curricula arose from a consensus of educators and I had the freedom to decide how I taught the material. These formative career experiences mean I am not automatically opposed to the concept of national curriculum, but more recent educational experiences have contributed to serious concerns about how curriculum decisions are made.

DILEMMAS AND DECISIONS

How Much Should Course Assessments Mirror PRAXIS II: PLT?

Because the initial pass rates on the PRAXIS II tests were disappointing, the Dean's Office in the College of Education and Human Services encouraged the faculty to take the tests. Taking the PRAXIS II: PLT, 5-9, test in January 2000 was a sobering experience and I concluded, similar to Sudzina (2001), that many undergraduates without experience of similar assessments

would find it difficult to do well. The test is long, reading the case studies and answering the multiple-choice and constructed response questions left little time to spare. I was unsure of the correct answer on at least five of the multiple-choice items and I could not determine what was wanted on all of the constructed response items. My problems in determining the right answers not only related to the lack of context in the multiple-choice items (Darling-Hammond, 2001) but also that I believed the research evidence associated with some questions was ambiguous so there was no clear right answer.

Taking the test raised important issues about the assessments in the educational psychology course. I had required a variety of in-class and out-of-class writing exercises, based on observations students made, summarizing and applying research and theory, analyzing cases, and exploring teaching dilemmas. Students were encouraged to revise and resubmit their written work. In grading the assignments associated with case studies and dilemmas, I focused on supporting arguments rather than right answers. After taking the PRAXIS II: PLT test, I was concerned that my tolerance for a wide variety of answers, which I believed helped students learn to clarify their thinking, may inadvertently mislead students into believing such tolerance existed on PRAXIS II: PLT when it did not.

I decided to reduce the number of writing assignments and to add "case study quizzes." Students now read cases from *Case Studies: Applying Educational Psychology* (Jackson & Ormrod, 1998) and take quizzes comprising up to five multiple-choice and two constructed response items. The items are more detailed than those in PRAXIS II: PLT but they do help students learn to read the cases carefully and apply educational psychology concepts related to them. In a typical semester, students now read at least 12 case studies. Very early in the semester, students discuss the open-ended questions at the end of the case studies in groups so that they learn to read the cases carefully and base their discussions on the facts in the case and their relationship to educational

psychology concepts rather than some gut-level reaction such as, "I think the teacher was great." By the 4th week, students in small groups answer several multiple-choice questions and open-ended questions on the case study. By midsemester, students begin to take case study quizzes at the end of the coverage of each chapter. My use of case study quizzes is consistent with research in Arizona indicating that state-mandated tests encouraged elementary school teachers to use instructional methods and material that resembled the tests (cited in Cimbricz, 2002).

Students do get more skilled at reading cases and answering questions associated with the cases. The frequent chapter assessments also encourage students to keep up with their reading and discourage them from missing class (Thompson, 2002). Most of the case study quizzes are open book and some students report that the quizzes are learning experiences because they are able to reread specific passages in the text in order to answer questions. Students who have recently taken PRAXIS II: PLT report that their experience in reading case studies, answering the questions based on them, and discussing the answers and test-taking strategies when the quizzes are returned was helpful. Students also report that reading the case studies carefully to prepare for the quizzes helps them think about practice.

However, the addition of the case study quizzes means that students do less writing and the course is a designated university writing across the curriculum course. Students do the minimum amount of writing required for such courses (2,000 words with opportunities to revise) but no more. The Educational Psychology faculty have tried to have the writing across the curriculum designation moved to another course but have been unsuccessful. Assessments reflect teaching priorities, and these have changed with the emphasis on PRAXIS II: PLT.

How much teaching priorities should have changed is at the heart of the dilemma. I believe that my choice to increase my emphasis on PRAXIS II: PLT-type assessments and decrease the amount of writing when other instructors

made different choices is related to my experiences of assessment as a cultural genre (Greenfield, 1997). My New Zealand education involved very few multiple-choice tests and my first experiences with high-stakes, multiple-choice tests felt very foreign. I did not know how to study or specific test-taking strategies, such as reading the answers before the question, yet my scores on one such test, the Graduate Record Examination (GRE), determined whether I was accepted into a U.S. graduate program. Tests based on case studies are a particular genre of assessment that requires specific strategies and skills, and many students need practice in this kind of assessment.

Another problem with the use of case-study quizzes is there are now few whole-class, open-ended discussions associated with the case studies. The use of case studies quizzes is unlikely to result in some of the reported benefits of case study use, such as framing problems and reflecting on teachers' work from multiple perspectives (Darling-Hammond, 2002). For example, the multiple-choice items make it difficult to stress the importance of multiple viewpoints in cases, and students' writing about the cases is limited to their answers to constructed response items on the quizzes. Each semester, some students say they would like more general discussion about the case studies, but I have not been able to determine how to fit that into the course.

I also have changed my end-of-semester assessment. I added a closed-book final exam, which mimics the format of PRAXIS II: PLT; that is, it includes case studies in the same format as the case study quizzes and some additional multiple-choice items. Students study for this exam in traditional ways, rereading and summarizing the designated chapters and studying their notes to master the content. Before this final exam was added, students completed an end-of-semester reflection paper based on the question, "What two of three main ideas or themes do you think you will take with you from this course?" The reflection paper focused on the integration and synthesis of learning; the current final examination does not. The concep-

tual model for the College of Education and Human Services is "Teacher as a Responsive, Reflective Professional: A Partner in Learning," but now the students in my class do less written reflection. This is an undesirable outcome but the priorities of the teacher education program are complex because high pass rates on the PRAXIS II: PLT scores are crucial to the institution. To some extent, teaching priorities and assessments are a zero-sum system—adding priorities and assessments means reducing emphasis in other areas.

How Much Should Course Content Be Influenced by PRAXIS II: PLT?

I was astonished that there was a question on Maslow's theory of needs on one PRAXIS II: PLT test. Maslow's theory is old, with little current empirical support (Ormrod, 2003), and there is a wealth of more recent relevant theory and high quality research on motivation. Although there is evidence that many K-12 teachers narrow the curriculum once high-stakes tests are introduced (Cimbricz, 2002), I have broadened the content coverage within the Educational Psychology course to include topics such as Maslow.

Since the imposition of PRAXIS II: PLT, I struggle more with the importance of covering all the content, even content that does not seem important for practice. Overwhelmed students ask which of a plethora of information in the textbook is central and I find it more difficult to answer than I used to. Educational Testing Service's test preparation guidelines do not help prioritize some of the content. For example, under the topic "Student Development and the Learning Process" is the subtopic "Examples of Important Theorists" (ETS, 2003b). What does *important* really mean if Maslow was included on a recent test?

Broadening the content coverage creates the danger that the curriculum becomes "a mile wide and an inch deep"—a description often used to describe the K-12 math and science curriculum in the United States (McKnight & Schmidt, 1998). Surface knowledge of educa-

tional psychology principles is unlikely to help preservice teachers make appropriate classroom applications. For example, preservice teachers may conclude that praise is important for students' motivation rather than the crucial nuances that some types of praise are motivationally facilitating, whereas other types are debilitating (Dweck, 2000).

Content coverage in educational psychology is further complicated by the tradition of presenting a variety of theoretical perspectives to describe key concepts and educational applications rather than teaching big ideas that are common across varying perspectives (Ormrod, 1998). Students are expected to grapple with theoretical distinctions that are crucially important for researchers but may not be for beginning teachers. For example, the principle "learning is enhanced when learners engage in self-evaluation" is derived from behaviorists' programmed instruction, the information-processing concept of comprehension monitoring, and social cognitive theory of self-regulation (Ormrod, 1998, p. 11). Preservice teachers often learn about this principle as a small part of three theoretical approaches each with its own specific terminology making application to classroom practice difficult.

Following the PRAXIS II: PLT content guidelines too closely creates other problems. For example, child development is given very little attention (ETS, 2003b), yet many educational psychologists believe that understanding child development is crucial for effective teaching. Adapting course content too closely to the test content also means that the test developers control the curriculum rather than educational institutions and faculty members.

The content changes that I have made leave less time to spontaneously "adapt, create and diverge"—a finding reported by Smith (1991, p. 10) in studies of teachers in Arizona with high-stakes testing. It also allows less time to focus in areas where one's knowledge and passion is particularly high. Obviously, too much attention to one's own passions in a core class can create problems for a coherent program, but determining the appropriate balance, which has implications for academic freedom and is

always complex in core classes, is exacerbated under high-stakes testing.

What Are Effective Teaching Strategies?

The pressure I felt to help my students master the range of content raised another dilemma: Could I help prepare students to pass this exam while not compromising my beliefs about the importance of constructivist ideas in learning? This is the same challenge that many K-12 teachers, especially urban teachers, face with state-wide, mandated, standardized tests. Before 1988, I had included a number of commonly recommended K-12 practices in teaching the educational psychology classes because I believed it was important to model these practices, discuss how they relate to educational psychology concepts, and describe how they are used in local schools. After the introduction of PRAXIS II: PLT, I read more methods books aimed at K-12 teachers and talked to several K-12 master teachers for more ideas. As a result, I now include more of Kagan's (1994) cooperative learning structures that are useful for mastery (e.g., trade-a-problem, rotating review) as well as continue to use the more common techniques (e.g., think-pair-share, jig-saw, numbered heads). Although college students will tolerate a 2-hour class session of "lecture and discussion," elementary and middle school students typically will not, so it is important to consistently model a greater diversity of instructional strategies.

My goal is to model how a class can focus, in part, on mastery and preparing for a test, but also be enjoyable, motivating, and contain a wide diversity of teaching and assessment strategies. I make this goal explicit by briefly reflecting on each teaching strategy after its use in class. Reflection is a component of the teacher education model and it helps students understand that some of their peers prefer and believe they learn more from lectures, some from cooperative learning structures, and others from whole-class discussions. Of course, a variety of preferred learning styles is prevalent in K-12 students as well as preservice teachers.

This approach has advantages. Social cognitive theory stresses the importance of modeling as a form of learning (e.g., Bandura, 1963; Pintrich & Schunk, 2002), and preservice teachers report that this modeling not only helps them learn educational psychology content but also to think about their future teaching. On the other hand, my focus on students' mastering the content and applying it has diminished the time spent on some forms of higher order thinking. The students do less brainstorming and are less likely to spend time discussing difficult dilemmas (e.g., teaching students with nonstandard dialects, or what it means to be a moral teacher). Current students sometimes complain about the lack of unstructured, whole-class discussions on topics that arise from the readings. The structured nature of the class means there are fewer teachable moments and the instructional pace is sometimes too fast. I am constantly aware that if I do not help prepare students for PRAXIS II: PLT, some of my students, no matter their sophisticated analyses of dilemmas or how beneficial open-ended discussions are, will never be able to teach in Ohio.

The reduced time spent on higher order thinking (apart from application) is consistent with research on the impact on high-stakes testing on K-12, but my increased use of cooperative learning structures and other nontraditional mastery techniques is not (Cimbricz, 2002). Teaching under high-stakes testing has become more complex, not less, in contrast to Smith's (1991) predictions for K-12 teachers. I believe that my decision to experiment with a greater variety of instructional techniques stems from my own experiences as an educator and my framing the dilemma from the perspective of a classroom teacher. This perspective is consistent with my background in educational psychology and teaching rather than policy or social foundations. I did not focus on PRAXIS II: PLT as an illustration of the grand narrative of decline in education, although I do talk about this with students.

Part of the dilemma in deciding what teaching strategies to use is determining what constitutes effective teaching in this course. How

important are student evaluations, peer evaluations, or a coherent teaching philosophy? Should effectiveness be measured by students' knowledge of educational psychology concepts on various assessments (including PRAXIS II: PLT), their skills in applying educational psychology concepts in the classroom, or their appreciation of the value of these concepts? Empirical research on pedagogical course work and teacher effectiveness is not fine-grained enough to determine what kinds of course work contribute to teacher effectiveness (Allen, 2003) or the types of teaching strategies within courses that are most effective.

What Are the Costs of the Increased Importance of Educational Psychology?

Since the implementation of PRAXIS II: PLT, students seem to perceive that the educational psychology class is more important. Although I always communicated the importance of the class to students, they typically believed that the really important courses involved methods, field experiences, and student teaching. If students were taking one of these classes simultaneously with the educational psychology course, work in the educational psychology class often suffered. Students would say things such as, "I really like this class but was up to 2 a.m. preparing my lesson plans for my unit."

Now that passing PRAXIS II: PLT is a prerequisite for licensure, students take the class more seriously and some believe that because educational psychology content is on a mandated standardized test it must be important. I did not have the foresight to collect data on students' motivation in educational psychology before PRAXIS II: PLT was mandated, so I do not have "hard" data to document this belief, but modern expectancy-value motivation theory predicts that motivation depends on both expectancy-related and task-value beliefs (e.g., Eccles & Wigfield, 2002). An increase in the importance of the task may increase task-value beliefs and motivation and performance. According to this model, task-value beliefs include four components: utility value, attainment value, intrinsic value, and cost (Eccles &

Wigfield, 2002). The utility value of educational psychology, or how well the task relates to current or future goals, has obviously increased with the introduction of PRAXIS II: PLT. Although doing well in the educational psychology course is not the same as doing well on PRAXIS II: PLT, students believe they are related. It may be that the attainment value, or personal importance placed on doing well on the task, also has increased for some students. Tasks that allow individuals to demonstrate competence in a domain that is related to salient aspects of one's self schema (e.g., being a teacher) are related to attainment value, and educational psychology may have become more salient in the self-schema of becoming a teacher. It is less likely that intrinsic value or enjoyment in educational psychology classes has increased, unless the instructors have become more skilled in fostering intrinsic motivation.

Relative cost includes performance anxiety, fear of failure, or increased effort, and this may have increased for some students. Maintaining the right balance of helping students understand the format of PRAXIS II: PLT but not increasing their debilitating anxiety is difficult. Extreme cases of test anxiety in my classes do appear to have increased, although this may reflect the recent increased severity of college students' psychological problems (Benton, Robertson, Tseng, Newton, & Benton, 2003) or the unreliability of anecdotal evidence. Students are aware that failure on the PRAXIS II tests is possible and they develop strategies to help ensure success before student teaching. Increasingly, students plan to take PRAXIS II: PLT immediately after they have taken the educational psychology course so there is time to retake the test if necessary. They say it "only" costs about \$100 so it is worth the risk. Students increasingly talk about investing time studying for PRAXIS II: PLT (e.g., over a semester break), and this additional effort is gratifying for those of us who believe that understanding educational psychology principles is important for effective teaching. However, whereas educational psychology content has become privileged, social foundations content has not; it

does not appear on any PRAXIS II tests required in Ohio, so students do not spend extra time seeking to understand that content. What consequences might this have at the programmatic level for a teacher education program in an urban university?

The increased sense of importance of educational psychology also is evident in administrative decisions made at the department and college levels. Part-time instructors and beginning doctoral students are less likely to teach educational psychology than 5 years ago, and replacing educational psychology tenure track faculty who leave with tenure track faculty has become automatic. However, the College of Education and Human Services has limited resources and so has to rely on part-time instructors and doctoral students, and their use outside of educational psychology is another indication that the PRAXIS II: PLT test has increased the importance of educational psychology at the expense of other courses.

What Are the Consequences of Helping Students Do Well on Someone Else's Exam?

Jenks (1998) argues that one advantage of national exams is that students and teachers are on the same side: Both are trying to ensure that the student does as well as possible on someone else's exam. This is in contrast to a class in which norm-referenced grading occurs and students perceive the teacher as controlling the valued, but limited, resource of high grades. Can a national, external, curriculum-based examination such as PRAXIS II: PLT allow instructors to focus more on mastery goals than performance goals as the grades in the classroom are less important?

The distinction between mastery goals (also called learning or task goals) and performance goals (also called ego and ability goals) is central to motivational goal theory (Dweck, 2000; Pintrich & Schunk, 2002). When mastery goals are dominant, the reasons for learning focus on understanding, mastering the content, solving problems, and self-improvement. When performance goals are dominant, the reasons for learning focus on demonstrating superiority to

others, striving to be the best in the group, seeking public recognition for high performance, and avoiding looking dumb (Dweck, 2000; Pintrich & Schunk, 2002). According to goal theory, a focus on mastery goals should increase student persistence and intrinsic motivation, encourage effective learning strategies, and reduce self-handicapping strategies (e.g., Dweck, 2000; Midgley, Arunkumar, & Urdan, 1996). In college classrooms, mastery goals are positively associated with student engagement and an evaluation approach that is based on competence rather than norms but negatively associated with harsh evaluation and a focus on evaluation rather than learning (Church, Elliott, & Gable, 2001).

Bishop (1998, 1999) argues that grading on the curve gives students a personal interest in persuading each other not to study. A serious student makes it more difficult for others to get good grades. In contrast, when learning is assessed by an outside standard, students no longer have a personal interest in getting teachers off track or persuading each other to refrain from studying. I had not graded on the curve for 18 years, but the recent addition of PRAXIS II: PLT does seem to have shifted students' perceptions. When I first assigned the case study quizzes, I thought that I had added too many forms of assessment, but when I asked students on the course evaluation if there were too many case studies, they said, "no," even though in the last 8 weeks of class there were seven case study quizzes, a paper, and a final exam. Marsh (2001) has made the distinction between good (useful) and bad workload and it seemed from informal comments and course evaluations that the students evaluated the workload as good.

When I ask students late in the semester if they can determine my goals by what has gone on in class so far, they always include "helping us do well on the PRAXIS II exam." This contributes to positive relationships in the classroom, especially with the weaker students or students who have a poor history with standardized tests. I tell the students what I learned by taking PRAXIS II: PLT, inform them of the university resources to help them prepare for the test, and encourage students who have taken the

PRAXIS II: PLT to describe their experiences. Students interpret these behaviors, as well as the mini-case study quizzes (see above), as my caring about them. Research has documented the importance of teachers' caring at K-12 (Ladson-Billings, 1994; Wentzel, 1996) and college levels (Buskist & Saville, 2001). Teacher caring may help student attendance, motivation, achievement, and the classroom atmosphere. Recent students in my classes seem less likely to be angry when things do not go their way or they get a bad grade on a quiz.

However, there are some potential dangers in identifying too closely with students' immediate concerns. My ultimate goal is to help to develop effective K-12 teachers, and focusing too much on subgoals such as passing PRAXIS II: PLT can distort priorities. The life stories of many of the students are compelling, and it is difficult to decide how flexible to be in deadlines or assignments. How much do I focus on helping the preservice teachers who are struggling because of the complexity and difficulty of their lives rather than insisting that they meet the standards needed for beginning teachers who teach K-12 students? The addition of PRAXIS II: PLT has made this balance harder to maintain.

Identifying too closely with students also can make it difficult to identify with some aspects of the profession and to be open-minded when reading some of the research literature because it can seem distant, irrelevant, or idealistic. It is easy to react the way K-12 urban teachers often do in relation to research, that is, they (the authors) know little about the "real world" of an urban, commuter, open admissions university such as Cleveland State University.

INSTITUTIONAL RESPONSE

The implementation of PRAXIS II (PLT and content tests) tests raised important issues for the institution. Because the initial pass rates of the students on the PRAXIS II were disappointing, the College of Education and Human Services paid for many of the faculty in the Education and Arts and Science Colleges to take the PRAXIS II tests as well as for an on-campus Educational Testing Service workshop. The col-

lege continues to pay the test fee for any faculty members who wish to take the test. An associate dean spent days developing a test preparation workshop for students, writing a test preparation booklet, and reporting regularly on trends in the PRAXIS II tests. Faculty spent hours examining and discussing the PRAXIS II data and modifying their courses to align their curricula. A similar investment of time and resources was documented by Cochran-Smith and Dudley-Marling (2001) in their study of teacher testing in Massachusetts, and it resembles the heavy investment of activities of institutions described by Ludlow et al. (2002) as besieged.

The implementation of PRAXIS II: PLT has increased collaboration among those teaching education psychology. A common textbook is chosen by consensus each year and instructors meet regularly to share resources, discuss dilemmas, and share instructional strategies. Increased collaboration among K-12 teachers because of state-mandated testing has been reported by Grant (2000).

One way to increase pass rates of an exit test is to raise entrance requirements. A faculty committee proposed raising the minimum grade point average (GPA) to gain acceptance into the education program from 2.50 to 2.75. After considerable discussion, it was defeated for several reasons. First, the majority of faculty believed that the mission of the college and university was to accept students with varying competencies and to help them improve their skills during course work. A clear distinction was made between entrance and exit competencies. Second, analyses indicated that a significant proportion of African American students would be eliminated with such a proposal, contrary to the urban mission of the teacher education program. Third, there were no data indicating that raising the entrance GPA by .25 would significantly improve candidates' pass rates on PRAXIS II: PLT or their teaching effectiveness.

After the decision was made, analyses of the Cleveland State University PRAXIS II: PLT data revealed some surprisingly good news: In 2000-2002, overall pass rates were higher than the state average and the African American-White

and gender test score gaps were smaller than the state average (Sutton, 2003a). This was cited as one indicator of success by the recent National Council for Accreditation of Teacher Education (NCATE) board of examiners accreditation report. Maintaining test scores continues to be a major concern and the first 2003-2004 annual goal of the College of Education and Human Services is that "students will exceed 91% passage on PRAXIS II and maintain 100% passage on Praxis III" (Cleveland State University, 2004).

Institutional investment in helping students pass PRAXIS II: PLT is not unique to Cleveland State University. In 2003, a consortium of universities, 2-year colleges, and 4-year colleges in Northeast Ohio agreed on a common curriculum for a variety of preservice core classes. PRAXIS II: PLT dominated much of the discussions in the Educational Psychology Group and the content topics outlined in the ETS test preparation guidelines (2003b) were used as an organizer for the common syllabus guidelines (Northeast Ohio Regional Collaboration, 2003).

CONCLUSION

In this article, I have described my dilemmas as an experienced teacher educator in an open admissions, urban university when PRAXIS II: PLT was mandated. My teaching now requires more effort and skill than it did before the tests were required because of the greater complexity of the day-to-day decisions, the wider range of teaching activities used, and the increased complexity of the teaching dilemmas. This is in contrast to the contention that high-stakes, multiple-choice testing ultimately deskills teachers (Smith, 1991). Since the implementation of PRAXIS II: PLT, I have altered the assessments, content, and teaching methods in my course; this is consistent with the conclusion that "state-mandated [K-12] testing does matter and does influence what teachers say and do" (Cimbricz, 2002, p. 10). The amount of this influence no doubt depends on the type of educational institution. Instructors at institutions where students are selected on the basis of their high standardized test scores are less likely to find the dilemmas as complex, or to modify assessments

and teaching strategies as much as those of us who work in an open admissions context.

Cochran-Smith and Dudley-Marling (2001) reported that teacher education faculty and administrators were demoralized with the advent of the Massachusetts high-stakes teacher test. This is not my experience, perhaps because PRAXIS II: PLT was not introduced in a mean-spirited manner and its content is public.

My reaction also is related to my teaching biography and approach to teaching. I think about teaching courses as designing and doing complex jigsaw puzzles. The context, instructor, and curriculum influence the shapes of the pieces and the complexity and picture on the puzzle. Teaching a course for the first time is like trying to do a puzzle without knowing what the complete picture looks like, what pieces are needed, and how the pieces fit together. With experience, the picture becomes clearer and the pieces needed to form the puzzle become identified. Small modifications of the course require altering the shapes or replacing a few pieces to make them fit better—not changing the entire picture. At some point, putting the same pieces together to form the same puzzle is no longer challenging and the course needs breaking apart, that is, a new jigsaw needs to be designed. When PRAXIS II: PLT was implemented, I was ready to design a new jigsaw puzzle, so it was a timely catalyst to rethink my teaching.

During 2003, the format of the PRAXIS II: PLT tests was modified. An extra case study has been added and there are no multiple-choice items related to the case studies (ETS, 2003b). This created another question: Should I change my course assessments and instruction again to match more closely the new version of PRAXIS II: PLT? My immediate reaction this time was frustration. I had just developed a clear picture and determined where all the pieces of the puzzle went when the test format was changed. On reflection, I realized that even though I had no control over the changes in the PRAXIS II: PLT format, less emphasis on multiple-choice items is preferable and the increased emphasis on constructed-response questions will allow more writing in the course while still helping students prepare for PRAXIS II: PLT. Some pieces

of the puzzle will have to be modified but the final picture should be more aesthetically pleasing.

NOTE

1. Analyses of unpublished raw data obtained from Educational Testing Service in 2003.

REFERENCES

- Allen, M. (2003). *Eight questions on teacher preparation: What does research say?* Denver, CO: Educational Commission of the States. Retrieved January 4, 2004, from www.ecs.org
- Bandura, A. (1963). *Social learning and personality development*. New York: Holt, Rinehart & Winston.
- Benton, S. A., Robertson, J. M., Tseng, W.-C., Newton, F. B., & Benton, S. L. (2003). Changes in counseling center client problems across 13 years. *Professional Psychology: Research & Practice, 34*(1), 66-72.
- Bishop, J. H. (1998). The effect of curriculum-based external exit exam systems on student achievement. *Journal of Economic Education, 29*, 171-182.
- Bishop, J. H. (1999). Nerd harassment, incentives, school priorities, and learning. In S. E. Mayer & P. E. Peterson (Eds.), *Earning and learning: How school matters* (pp. 231-280). Washington, DC: Brookings Institution.
- Buskist, W., & Saville, B. K. (2001). Creating positive emotional contexts for enhancing teaching and learning. *American Psychological Society Observer, 14*(3), 12, 13, 19.
- Church, M. A., Elliott, A. J., & Gable, S. L. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology, 93*, 43-54.
- Cimbricz, S. (2002). State-mandated testing and teachers' beliefs and practices. *Educational Policy Analysis Archives, 10*(2). Retrieved June 17, 2003, from <http://epaa.asu.edu/epaa/v10n2/>
- Cleveland State University. (2003). *Institutional research and analysis 2002-2003 factbook*. Retrieved June 3, 2003, from <http://www.csuohio.edu/iraa/FactBook03/>
- Cleveland State University. (2004). *College of education, important documents*. Retrieved June 3, 2004, from <http://www.csuohio.edu/coehs/>
- Cochran-Smith, M., & Dudley-Marling, C. (2001). The flunk heard round the world. *Teaching Education, 12*(1), 49-63.
- Darling-Hammond, L. (2001). Teacher testing and the improvement of practice. *Teaching Education, 12*(1), 11-34.
- Darling-Hammond, L. (2002). Toward pedagogy of cases in teacher education. *Teaching Education, 13*(2), 125-135.
- Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109-132.

- Educational Testing Service (ETS). (2001). *Principles of learning and teaching: Tests at a glance*. Princeton, NJ: Author.
- Educational Testing Service (ETS). (2003a). *The PRAXIS series: Advancing professional excellence*. Retrieved March 9, 2003, from <http://www.ets.org/praxis/prxstate.html>
- Educational Testing Service (ETS). (2003b). *The PRAXIS series: Advancing professional excellence*. Retrieved June 23, 2003, from <http://www.ets.org/praxis/prxstate.html>
- Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95-113.
- Grant, S. G. (2000). Teachers and tests: Exploring teacher perceptions of changes in the New York State testing program. *Educational Policy Analysis Archives*, 8(14), 1-16. Retrieved June 25, 2003, from <http://epaa.asu.edu/epaa/v8n14.html>
- Greenfield, P. M. (1997). Why you can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*, 52(10), 1115-1124.
- Haney, W. (2000). The myth of the Texas miracle in education. *Educational Policy Analysis Archives*, 8(41). Retrieved June 25, 2003, from <http://epaa.asu.edu/epaa/v8n41/>
- Indiana University. (2002). *NSSE 2002 means summary report for Cleveland State University*. Bloomington: Indiana University Center for Postsecondary Research and Planning.
- Jackson, D. L., & Ormrod, J. E. (1998). *Case studies: Applying educational psychology*. Columbus, OH: Merrill.
- Jenks, C. (1998). Racial bias in testing. In C. Jenks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 55-85). Washington, DC: Brookings Institution.
- Kagan, S. (1994). *Cooperative learning*. San Clemente, CA: Kagan Cooperative Learning.
- Ladson-Billings, G. (1994). *The dreamkeepers: Successful teachers of African American children*. San Francisco: Jossey-Bass.
- Ludlow, L. H., Shirley, D., & Rosca, C. (2002). The case that won't go away: Besieged institutions and the Massachusetts teacher tests. *Educational Policy Analysis Archives*, 10(50). Retrieved June 15, 2003, from <http://epaa.asu.edu/epaa/v10n50/>
- Marsh, H. W. (2001). Distinguishing between good (useful) and bad workloads on students' evaluations of teaching. *American Educational Research Journal*, 38, 183-212.
- McKnight, C. C., & Schmidt, W. H. (1998). Facing facts in U.S. science and mathematics education: Where we stand, where we want to go. *Journal of Science Education and Technology*, 7(1), 57-76.
- Midgley, C., Arunkumar, R., & Urdan, T. C. (1996). "If I don't do well tomorrow, there's a reason": Predictors of adolescents' use of academic self-handicapping strategies. *Journal of Educational Psychology*, 88, 423-434.
- Northeast Ohio Regional Collaboration. (2003). *Instructors' guide: Educational psychology regional collaboration on teacher education*. Unpublished manuscript.
- Ormrod, J. E. (1998, April). *Teaching teachers: The problem with emphasizing "isms."* Paper presented at the annual Educational Research Association Conference, San Diego, CA.
- Ormrod, J. E. (2003). *Educational psychology: Developing learners* (4th ed.). Upper Saddle River, NJ: Prentice Hall.
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2nd ed.). Upper Saddle River, NJ: Merrill.
- Sacks, P. (1999). *Standardized minds: The high price of America's testing culture and what we can do to change it*. Cambridge, MA: Perseus Books.
- Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11.
- Sudzina, M. R. (2001, April). "Psyching out" the PRAXIS II Tests of Learning and Teaching: What our students need to know. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Sutton, R. E. (1991). Equity and computers in the schools: A decade of research. *Review of Educational Research*, 61, 475-504.
- Sutton, R. E. (1996). Advancing the knowledge base: Diversity and teacher education. Essay review of "Teaching diverse populations: Formulating a knowledge base." *Journal of Teacher Education*, 47(4), 310-314.
- Sutton, R. E. (1997). Equity and high stakes testing: Implications for computerized testing. *Equity and Excellence in Education*, 30(1), 5-15.
- Sutton, R. E. (2003a). *PRAXIS II PLT scores at Cleveland State University: Pass rates higher, racial and gender gaps smaller than statewide averages*. Unpublished manuscript.
- Sutton, R. E. (2003b, August). *Teaching educational psychology with the accountability of PRAXIS II: Possible benefits?* Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Sutton, R. E., & Fleming, E. S. (1994). Evaluation of EQUALS: An inservice program designed to promote excellence and equity in mathematics. *Equity and Excellence in Education*, 27, 26-31.
- Thompson, B. (2002, June 21). If I quiz them they will come. *Chronicle of Higher Education*, p. B5.
- USnews.com. (2004). *America's best colleges, 2004*. Retrieved January 3, 2004, from http://www.usnews.com/usnews/edu/college/rankings/rankindex_brief.php
- Wentzel, K. R. (1996). Social goals and social relationships as motivators of school adjustment. In J. Juvonen & K. R. Wentzel (Eds.), *Social motivation: Understanding children's school adjustment* (pp. 226-247). New York: Cambridge University Press.
- Yinger, R. J. (1999). The role of standards in teaching and teacher education. In G. A. Griffin (Ed.), *The education of teachers: Ninety-eighth yearbook of the National Society for the Study of Education* (pp. 85-113). Chicago: University of Chicago Press.

Zancanella, D. (1992). The influence of state-mandated testing on teachers of literature. *Education Evaluation and Policy Analysis*, 14(3), 283-295.

Rosemary E. Sutton is a professor of educational psychology and the Director of Assessment at Cleveland State

University. Her recent research has focused on two areas related to teachers' emotions: emotion regulation and day-to-day anger and frustration experienced while teaching.