Boot school

Congratulations to Oakland mayor Jerry Brown on his plan to open a military academy (see "A Few Good Schools," Summer 2001). Chicago's experience with military academies has been overwhelmingly positive. I hope Oakland's is equally successful.

In 1999 Chicago opened its first public military high school, the Chicago Military Academy at Bronzeville, in a historic African-American neighborhood. Last year the city began converting Carver High School on the Far South Side into its second military school. Both schools are part of the Chicago Public Schools system, not charter schools like Mayor Brown's academy.

We started these academies because of the success of our Junior Reserve Officers Training Corps (JROTC) program, the nation's largest. JROTC provides students with the order and discipline that is too often lacking at home. It teaches them time management, responsibility, goal setting, and teamwork, and it builds leadership and self-confidence.

Not surprisingly, the high-school graduation rate for JROTC students in the Chicago Public Schools is 20 percent greater than the citywide average. It's a little early to measure the success of our academies, but the first class at Bronzeville scored 40 percent better than the citywide average in reading and 30 percent better than the average in math. Perhaps a clearer sign of success is that 1,300 students applied for 110 openings in Bronzeville's next entering class.

Though a military academy isn't for everyone, for some it is just what they need in order to make something of their lives.

Mayor Richard M. Daley
Chicago, Illinois

Do teachers matter?

Let me respond to a few points in Michael Podgursky's review of my report "How Teaching Matters: Bringing the Classroom Back into Discussions of Teacher Quality" (See "Flunking ETS, Check the Facts, Summer 2001). First, the notion that the study is tilted in favor of measures of classroom practice and against measures of socioeconomic status (SES) is incorrect. Adding up the effects for all the classroom-practice variables is standard procedure. The fact that very few teachers engage in all of the effective practices does not invalidate the procedure; it merely suggests that there is room for teachers to improve. Nor does this procedure give classroom practices an advantage over socioeconomic status. The socioeconomic variable was also created by adding together six items—in this case, before the models were estimated. The socioeconomic measures are certainly not as rich as would be desirable, but neither are the measures of classroom practice.

Second, the use of data from the 8th grade National Assessment of Educational Progress (NAEP) in this study is appropriate. It is certainly true that data from just one year cannot be used to establish a causal relationship, as I note in the report. Cross-sectional data are properly used, as they were in this study, to confirm the findings of more robustly designed small-scale studies. Without such validation, it is difficult to know if the findings of small-scale studies will hold true for different students in different schools. The advantage of longitudinal over cross-sectional studies also should not be oversold. Although the outcome in a cross-sectional study is a student's test score, and the outcome in a longitudinal study is improvement in a student's test score, neither case demonstrates that the teacher caused the outcome. Demonstrating a causal hypothesis requires an experimental design.

Harold Wenglinsky
Educational Testing Service
Princeton, New Jersey

Michael Podgursky's latest target in his ongoing war on the National Board for Professional Teaching Standards is a study authored by my colleagues and me at the University of North Carolina, Greensboro (see "Defrocking the National Board," Check the Facts, Summer 2001). Here I'll answer just a few of his concerns (an extended reply is available at www.educationnext.org).

Podgursky questions the ways in which we measured student achievement. The measures used in the study were 1) writing assignments in response to prompts designed by experienced teachers, and 2) assessments of the depth of student understanding of concepts targeted in instruction. Podgursky claims that without statistical controls for students' background characteristics, these data are suspect.

Were standardized, multiple-choice tests used as the measure of student achievement, an adjustment for students' socioeconomic status would certainly have been appropriate. Performance on standardized multiple-choice tests is affected by many factors not under teachers' immediate control. In this study, however, we were interested in the students' depth of understanding of concepts from a unit designed by each teacher for the students in her class. Whether socioeconomic differences affect student achievement under these circumstances depends primarily on the quality of the observation protocols, the quality of the training the observers and assessors received, and their skill in...
Podgursky asserts that neither this study nor any other "has ever shown that National Board–certified teachers are better than other teachers at raising student achievement." 

The basis for this assertion is Podgursky's belief that the only way to measure the performance of teachers is to examine the performance of their students on standardized multiple-choice tests. We believe that setting worthwhile instructional goals is a crucial aspect of accomplished teaching, and the extent to which students have met the goals set by their teachers is a rational and utterly defensible measure of student achievement.

The National Board uses highly trained teachers in the relevant disciplines to evaluate the submissions of teachers seeking advanced certification. Podgursky questions the National Board's use of teachers' peers in the evaluation of their work, though this is common practice at the college level, and calls for evaluation by principals and parents instead. Principals can evaluate some aspects of a teacher's performance (attendance, classroom management) but are in no position to evaluate other critical teaching attributes, such as in-depth subject-matter knowledge and the ability to present content in developmentally appropriate ways that engage students. Most parents know even less than principals about what goes on in actual classrooms or how teaching practices should be evaluated.

Podgursky asserts that we chose a particular sample of teachers in order to stack the deck in favor of positive findings. In fact, we chose a sample with teachers who were close to the certification score, and clearly below the certification score, as well as teachers who were clearly above the certification score. Our argument can be made for random sampling of teachers since it facilitates generalization, but it is highly unlikely that random sampling would have resulted in a materially different outcome, as the mean scores for the relevant populations of certified and non-certified teachers were quite similar to those of the study sample.

The National Board's system for identifying accomplished teachers is the most comprehensive assessment of actual teaching practice yet devised. Teachers must provide evidence of professional accomplishments and the involvement of community resources and students' families in the educational process. They must demonstrate their knowledge of their subject matter and their ability to select developmentally appropriate curricular materials to teach that content. These are rigorous, research-tested evaluation criteria that, as our study shows, are clearly identifying teachers of high professional caliber.

LLOYD BOND
University of North Carolina, Greensboro
Greensboro, North Carolina

Michael Podgursky replies: I agree that the cross-section correlations in Harold Wenglinsky's study do not support causal interpretations. However, his report is replete with prescriptive statements and policy recommendations. These recommendations are based not on experimental research or a larger body of studies, but on the findings in this study.

I explained carefully why the adding-up exercise that forms the basis for his conclusion that "teachers matter most" is flawed. Wenglinsky implicitly concedes this point and now states that the result of his exercise "merely suggests that there is room for teachers to improve." He too is a causal interpretation, but it is at least more cautious and defensible than the original.

Two basic design flaws exist in the study conducted by Lloyd Bond et al. As Bond concedes, the study over-sampled high-scoring certified and low-scoring noncertified teachers, thus exaggerating the measured differences in quality between the two groups. Bond states that it is "highly unlikely" that he would have obtained different results had he sampled randomly. This is entirely speculative—the data presented on mean test scores are irrelevant on this point. The authors' claim that "board certified" teachers "significantly" outscored their noncertified counterparts on 11 of 13 dimensions of good teaching is based on a flawed statistical test.

Second, the researchers failed to control for previous test scores or socioeconomic differences between the students of the certified and non-certified teachers. As I noted in my review, the children taught by the noncertified teachers were disproportionately low income. Bond does not dispute this, but makes the extraordinary claim (without evidence) that the methods used by his trained observers make it unnecessary to control for students' backgrounds or previous achievement. A rigorous study design would have compared certified with noncertified teachers who worked with similar student populations, or it would have collected extensive control data on students' backgrounds. The researchers did neither.

The "highly trained teachers" the National Board uses as scorers are moonlighting teachers who receive two to four days of training. They are paid $125 a day, and most have not passed the certification assessments they are grading. In fact, peer review of teaching in higher education is entirely localized and bears no resemblance to the centralized and very costly process of the National Board. There is no "national "certification" of experienced college teachers.

Although several hundred million dollars have been invested in board certification and bonuses to date, no one has yet undertaken a rigorous study of whether the students of board-certified teachers learn more and whether the board-certiﬁcation process is a cost-efﬁcient way to identify superior teachers.