

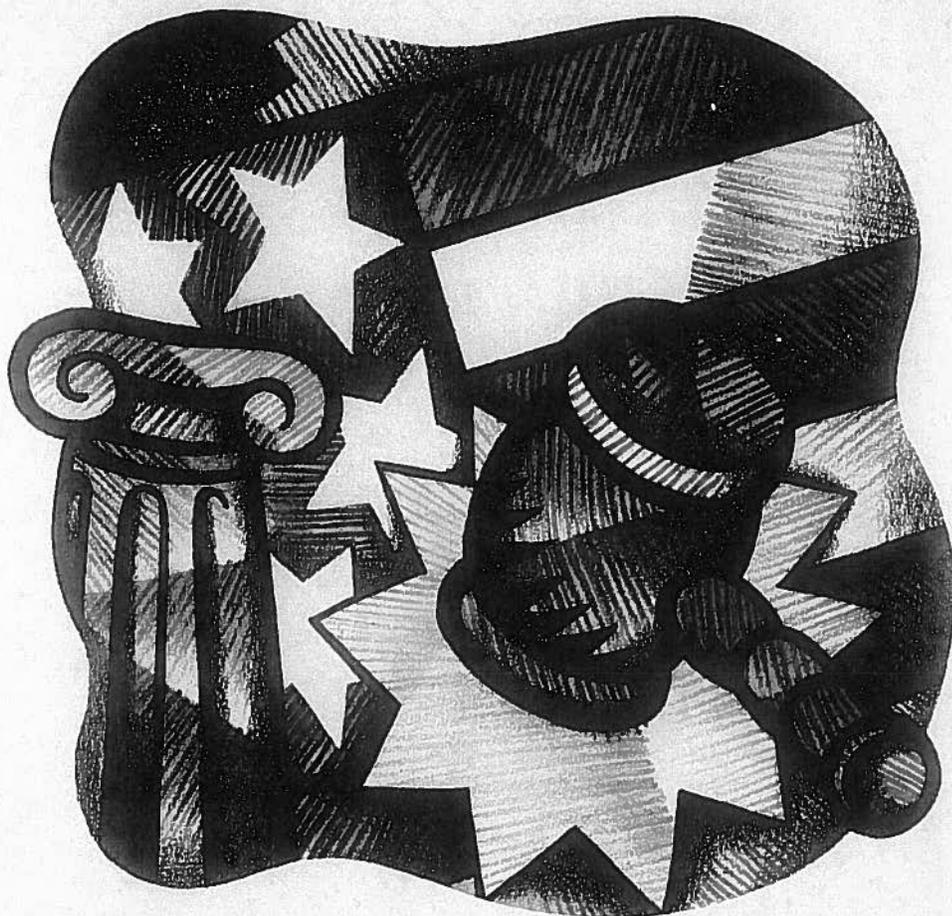
# Will the Courts Save Teachers?

*When seeking protection from patently unfair evaluation decisions, don't look to the courts.*

**W. James Popham  
and Marguerita DeSander**

In the next few years, thousands of teachers across the United States will lose their jobs. It won't be because of shrinking fiscal resources or because these teachers voluntarily chose other careers. Instead, the cause of these massive dismissals will be the markedly toughened teacher evaluation systems recently established in all but a few states.

It's difficult to overestimate the potential demoralization stemming from such impending dismissals. The relaxed conversations that typically take place in a school's faculty lounge are being replaced by tense speculations about fellow teachers' evaluation status. Will this year's friends be at the school next year, or will they be scrambling for another job—or even another profession? And the adverse impact on teacher morale is particularly potent when teachers perceive that the evaluation proce-



dures used to judge them are unfair.

Many teachers may hope that any teacher unfairly fired because of a flawed teacher evaluation process would be protected by the courts. After all, a “wrongful termination” is, by definition, wrongful. And aren't our courts supposed to rectify wrongs? Let's see.

## **Pressure for Tougher Teacher Evaluation**

Before addressing the issue of whether the courts will protect teachers from unsound high-stakes teacher evaluations, it's useful to understand why a host of stringent teacher evaluation procedures have recently emerged. Two federal initiatives were the



Two federal initiatives were the catalysts for the United States' current preoccupation with teacher evaluation.

### **Invalid Judgments Based on Student Growth**

As noted previously, federal officials have urged that student growth, typically signified by changes in students' test scores, should serve as a significant determinant when appraising teachers. In many states, students' "growth" will chiefly be judged by comparing their end-of-school-year performance on tests the previous year with their end-of-year performance on the tests this year.

This would be a sensible way of ascertaining how well a teacher had stimulated students' growth during the school year—that is, if there were evidence showing that improvements in students' scores on those tests actually indicated teachers' effectiveness. But there isn't.

In fact, in most instances, there is no evidence whatsoever that changes in students' state test scores are attributable to their teacher's competence. Such tests may actually be *instructionally insensitive*—that is, they may be unable to distinguish between well-taught and poorly taught students. Perhaps scores on a state's tests are more heavily influenced by the socioeconomic levels of the students assigned to a teacher than by the teacher's instructional skill.

If students' scores on instructionally insensitive tests constitute the "significant" evidence by which a teacher will be evaluated, can you see how inaccurate—and unfair—such a teacher appraisal system is apt to be? Reliance on tests not validated for the purpose of gauging teachers' instructional skills represents a major flaw in any teacher evaluation program.

### **Unreliable Classroom Observations**

A second error seen in many states' new teacher evaluation systems is an unwarranted, almost unthinking reliance on the evidence supplied by classroom observations. We sometimes see considerable weight being

catalysts for the United States' current preoccupation with teacher evaluation.

In 2009, the federal Race to the Top (RTTT) program promised whopping financial grants to states that were willing to undertake aggressive school reforms, including more rigorous teacher evaluation procedures. The federal guidelines said that teacher evaluation must be based on multiple sources of evidence but must include student test-score growth as a significant factor. Although Race to the Top was peppered with language calling for these new evaluation systems to improve teachers' skills, one clear mission of the recommended procedures was to remove ineffective tenured and untenured teachers.

Given the size of potential Race to the Top grants, officials in many cash-strapped states scurried to submit RTTT proposals. And to increase the likelihood of receiving one of these hefty grants, a number of state legislatures actually enacted laws mandating teacher evaluation procedures that coincided with the federal recommendations. Indeed, we now find almost half of U.S. states requiring that fully 50 percent of a teacher's evaluation be derived from test scores (Hull, 2013).

Two years later, in 2011, the U.S. Department of Education announced a second federal initiative. The Elementary and Secondary Education Act (ESEA) Flexibility Program offered

waivers from the sanctions that states would otherwise face under the ESEA's 2002 reauthorization, No Child Left Behind (NCLB). As with Race to the Top, the ESEA Flexibility Program encouraged states that sought waivers to promise energetic school reforms, including tough teacher evaluation programs closely tied to personnel decisions. And as with Race to the Top, the lure of NCLB waivers enticed many state authorities not only to apply for such waivers, but also to establish conditions that would increase their application's chances to be approved.

### **Major Flaws**

Spurred by the financial incentives of Race to the Top and the failure-avoidance options of the ESEA Flexibility Program, education officials in all but a few states have set out to beef up their teacher evaluation programs in the past few years. In general, the individuals who were crafting these new evaluation programs tried to do a solid job. Yet, candidly, most states' education authorities possess scant experience in creating teacher appraisal systems. As a consequence, in many states we've seen substantial amounts of invention as new teacher evaluation programs have sprung into life. Some of those inventions are good ones; some aren't. Let's look at two common problems that diminish the accuracy of many teacher evaluations.

given to classroom observation data, even though serious weaknesses may exist in the observation system being employed as well as in the way classroom observers actually collect evidence.

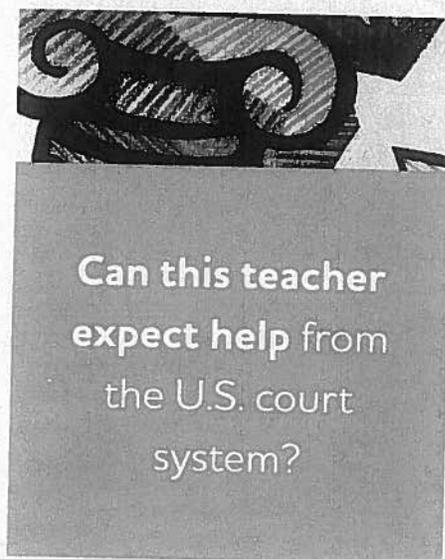
To illustrate, several of the observation systems now widely employed in U.S. teacher evaluation programs were patently designed for *formative* evaluation (that is, to help teachers improve) rather than for *summative* evaluation (that is, to supply the evidence necessary for removing ineffective teachers). See, for example, the original observation protocols of Danielson (1996) and Marzano (2007). And although a number of these observation frameworks are based on solid research showing positive relationships between students' achievement and teachers' use of particular instructional tactics, those relationships represent what's *likely* to occur, not what's *certain* to occur. Thus, a particular teacher might depart substantially from the positive instructional strategies identified in a given observation framework, yet still promote gobs of student growth. In addition, the number of classroom observations required by some states is too tiny, and the caliber of observer training is low.

In short, evidence of teacher quality collected by classroom observers may be quite wonderful, or it might be truly tawdry. To assign great evaluative weight to classroom observation evidence without verifying the quality of that evidence represents a serious flaw.

### Teacher Voices Are Missing

Given the difficulty of designing appraisal systems that judge teacher effectiveness accurately, as well as the remarkably high stakes, it would seem logical that teachers themselves should be involved in shaping such systems. Yet the vast majority of U.S. teachers appear to display little interest

in the nuts and bolts of the systems that could cost them a job or a career. Perhaps one reason that teachers in the firing line rarely seek to jump into the teacher evaluation fray is that they believe they will be unable to influence the nature of evaluation procedures carved out at higher levels of the policymaking pyramid. But another reason for such a lack of interest may be teachers' belief that, even if their state's teacher-appraisal program contains serious shortcomings, unfairly dismissed teachers will at least be protected by the courts. Let's consider that possibility.



### What to Expect from the Courts

Suppose teacher X has been unfairly dismissed because a flawed teacher evaluation system inaccurately rated her performance as marginal or substandard. Can this teacher expect help from the U.S. court system? The answer is a resounding no. U.S. courts have historically refused to substitute their judgment for that of a school board even if a termination is based on just a scintilla of evidence.

The only exception occurs when the fired teacher is a member of a class (for instance, a designated racial group) that's protected by the Fourteenth Amendment's equal protection

clause and Title VII of the Civil Rights Act of 1964 (Alexander & Alexander, 2012). However, the burden of proof is on the terminated teacher to demonstrate that his or her termination was based on being a member of that class. This burden of proof is substantial and difficult to satisfy.

Both federal and state courts have categorically declined to rule on the appropriateness of a teacher evaluation system or the evidence-collection procedures incorporated into that system. The role of the courts has always been to review the record to simply determine whether (1) laws, policies, and procedures established by the state and local authorities were followed, and (2) teachers were given due process. Even the most inadequate—even indefensible—evaluation system will avoid the rigors of court scrutiny as long as its procedures are applied in a consistent manner to all teachers affected.

Here's a case in point. In 2011, the Florida legislature introduced high-stakes teacher evaluation procedures through SB 736, a bill aligned with the state's successful Race to the Top grant proposal. The legislation required that 50 percent of a teacher's evaluation ratings—which would be tied to salaries, tenure, and other employment decisions—must be based on student performance growth on the state assessments (Education PreK–12 Committee, Florida Senate, 2011). But many Florida teachers do not teach in the grades (4–10) and subject areas (mathematics and reading) in which the state tests are administered. So, rather than delay implementation until assessments could be developed for every subject and grade level, the legislation allowed districts to evaluate teachers using standardized test scores for *students they did not teach* and for *subjects they did not teach*.

The Florida Education Association and the National Education Association initiated a lawsuit challenging

the legislation and asserting that basing teachers' evaluation ratings on the test scores of students or subjects they did not teach violated the U.S. Constitution's equal protection and due process clauses (Alexander & Alexander, 2012; Osbourne & Russo, 2011). One of the plaintiffs in the suit was 1st grade teacher Kim Cook. Because her school only went through 2nd grade, Cook was labeled "unsatisfactory" on the basis of the test scores of 4th and 5th grade students in an entirely different school—despite being her school's 2012 Teacher of the Year (Downey, 2013).

The courts offered little relief. In April 2013, the Circuit Court dismissed key claims of the lawsuit (Florida Education Association, 2013), a ruling that was upheld by another Circuit Court judge in June 2013 (Call, 2013). Although the Florida legislature eventually addressed this particular concern by passing a law requiring that a teacher's evaluation be based on the test scores of students whom the teacher actually taught, the education associations vowed to continue to challenge other problematic provisions of SB 736 (O'Connor, 2013). Clearly, educators throughout the United States will be watching with interest the results of this potentially precedent-setting challenge to the merits of teacher evaluation legislation.

### No Judicial Safety Net: Now What?

If teachers who have been fired because of flawed teacher appraisal programs, even seriously flawed ones, cannot count on the courts to protect them, what can teachers do to deal with this career-threatening issue? The answer is straightforward, if not easy. Teachers need to dig into the viscera of the particular teacher evaluation program most directly affecting them, learning enough about that program to identify its potential strengths and

weaknesses. If teachers spot significant weaknesses in a teacher appraisal program, they should pressure state, district, and school officials to remedy such deficiencies.

For example, by working with their local or state professional associations, small groups of teachers can point out to appropriate officials how to avoid the kinds of errors that are apt to reduce evaluation accuracy.



Evidence of teacher quality collected by classroom observers may be quite wonderful, or it might be truly tawdry.

Sensible suggestions, even those proffered by a solo teacher, can often alter an ill-conceived evaluative enterprise. School districts developing and implementing new high-stakes evaluation systems are ethically bound to use best practice—which includes giving teachers a voice in the process.

If we can't count on the courts to save the day, then teachers and schools must do their own day-saving—and quickly. ■

### References

- Alexander, K., & Alexander, D. M. (2012). *American public school law* (8th ed.). Belmont, CA: Wadsworth Cengage Learning.
- Call, J. (2013, June 25). Judge denies

request to rehear challenge to SB 736. *Florida Current*. Retrieved from [www.thefloridacurrent.com/article.cfm?id=33394405](http://www.thefloridacurrent.com/article.cfm?id=33394405).

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: ASCD.

Downey, M. (2013, April 16). Florida teachers file lawsuit today to stop evaluations that rely on test scores [blog post]. Retrieved from *Get Schooled* at <http://blogs.ajc.com/get-schooled-blog/2013/04/16/florida-teachers-file-lawsuit-today-to-stop-evaluations-that-rely-on-test-scores>

Education PreK–12 Committee, Florida Senate. (2011). S/CS/SB 736—Educational personnel. Retrieved from [www.flsenate.gov/Committees/BillSummaries/2011/html/0736ED](http://www.flsenate.gov/Committees/BillSummaries/2011/html/0736ED)

Florida Education Association. (2013, May 3). Teachers' lawsuit against merit pay law is dismissed [blog post]. Retrieved from *Palm Beach County Classroom Teachers Association* at <http://palmbeachcta.net/cta-blog/2013/05/03/teachers-lawsuit-against-merit-pay-law-is-dismissed>

Hull, J. (2013). *Trends in teacher evaluation: How states are measuring teacher performance*. Alexandria, VA: Center for Public Education, National School Boards Association.

Marzano, R. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA: ASCD.

O'Connor, J. (2013, May 15). What the Florida teacher evaluation lawsuit could mean for other states. Retrieved from *StateImpact Florida* at <http://stateimpact.npr.org/florida/2013/05/15/what-the-florida-teacher-evaluation-lawsuit-could-mean-for-other-states>

Osbourne, A. G., & Russo, C. J. (2011). *The legal rights and responsibilities of teachers: Issues of employment and instruction*. Thousand Oaks, CA: Corwin.

**W. James Popham** (popham@ucla.edu) is emeritus professor at the University of California—Los Angeles. His most recent book is *Evaluating America's Teachers: Mission Possible?* (Corwin, 2013).

**Marguerita DeSander** (desander@gwu.edu), an assistant professor at George Washington University in Washington, D.C., was formerly a practicing attorney specializing in employment-related law.