

America's "Failing" Schools: How Parents and Teachers Can Cope With No Child Left Behind by W. James Popham. New York: Routledge Falmer, 2004. 147 pp. ISBN 0-415-94947-5.

Underlying the No Child Left Behind (NCLB) act is the notion that a student's test performance provides an indication of the quality of their educational experiences. In his address to J.E.B. Stuart High School in Falls Church, Virginia on January 12, 2005, President Bush declared, "Listen, I've heard every excuse in the book not to test. My answer is, how do you know if a child is learning if you don't test." In *America's "Failing" Schools: How Parents and Teachers Can Cope with No Child Left Behind*, W. James Popham addresses this and other aspects of NCLB in a no-nonsense, uncluttered and direct style which is appealing to his target audience—parents and teachers. Through his use of catchy chapter titles, such as "Measuring Temperature with a Tablespoon," and exclusion of academic or technical terminology, Popham has the power and potential to engage the general public in these important educational issues. *America's "Failing" Schools* is just one of many books in Popham's assessment literacy campaign that provides a context for the prevalence of test-based accountability systems.

Popham acknowledges common criticisms of NCLB, such as the act's lack of funding and the intense focus on test performance, but he pushes the reader to focus on the real issue of trying to evaluate school quality. He takes great care not to place the blame on any particular group or political party, but instead attempts to empower parents and teachers to participate in a discussion of the merits of NCLB by providing them with information about indicators of educational quality.

In part because Popham's target audience is parents and teachers, this book may leave researchers and policymakers wanting more. For example, Popham glosses over the intricacies of measuring student learning and school quality. Although many people have a general understanding of terms such as "learning" and "quality," the nuanced nature of these concepts are lost in the text. After reading this book, one continues to wonder what it means to attend a school that is of high quality. Does this mean that the school effectively promotes student learning? And what does student learning look like? Can it be measured accurately on a large scale?

Popham sets forth a theory that instruction and testing are related, and that creating and implementing instructionally sensitive tests is crucial in test-based accountability systems. According to Popham, performance on traditional tests is "actually an indicator of the kinds of students who attend a school rather than how well a school's students have been instructed" (pp. 62-63). This leads Popham to discuss the importance of instructionally sensitive tests such as those

recommended in 2001 by the Commission on Instructionally Supportive Assessment, a group formed with the support of several associations¹. As explained on these organizations' Web sites, instructionally sensitive assessments can determine whether students received high- or low-quality instruction, as differences in test performance reflect differences in instruction. In other words, Popham believes that from test scores one would be able to judge whether and how well students were instructed on the tested material. He assumes, however, that students receive instruction on test material and that performance is responsive to such instruction. This link between performance and instruction is questionable, however, and has not been clearly established (see, for example, Anastasi, 1981).

As Popham points out, test performance is just a sample of what a student can do and "if folks really care about what happens to kids in classrooms, on-the-surface assessment doesn't cut it" (p. 81). Popham accurately conveys to his readers that performance on a single test is merely an estimate from which abstractions about students' skills or understanding of material are derived. Performance is also related to a variety of factors that have nothing to do with student learning or ability, such as being tired on the day of the test or familiarity with test questions or format. Despite this knowledge, readers will continue to question whether it is possible to separate the effect of a teacher's instruction from the influence of prior instruction and experience. Popham does not provide sound answers to this question, and he gives short shrift to the cumulative effects of prior instruction or experience on test performance. Thus, like current accountability measures, it is unclear if instructionally sensitive tests can separate the effects of previous instruction from current classroom instruction.

Although he advocates the use of instructionally sensitive tests, Popham also believes that there is a need to discuss what it means to evaluate quality education. This lack of discussion is an underlying problem with NCLB. Popham argues that improving the evaluation of quality education involves more than merely creating better tests, and provides suggestions that include using instructionally supportive assessments and student work samples, measuring student affect, and incorporating other non-test academic indicators.

Popham's ideas are laudable and have garnered much attention, but like measures of learning, they are difficult to define, measure, and relate to indicators of quality. For example, Popham proposes that non-partisan judges "blind-score" student essays and science laboratory reports to produce "credible evidence that some first rate pretest-to-posttest instruction has taken place" (p. 99). Yet while Popham suggests that the implementation and scoring of student work is fairly straightforward, these issues are still of great concern to the educational community. One such concern is that while standardized work samples might lack authenticity, the lack of guidelines or standardization jeopardizes the

evaluation because it is difficult to talk about quality without a common language or set of guidelines. Even if teachers were required to create a standard sample of student work, implementation across teachers will differ, which calls into question just how much of a student's work is actually a reflection of instruction. Thus, creating a link between instructional quality and student work is not as easy or clear as Popham would lead readers to believe (Atkin, Black, & Coffey, 2001; Ruiz-Primo & Li, 2004).

In *America's "Failing" Schools*, Popham successfully opens the discussion about evaluating quality education to parents and teachers by outlining the concepts of student assessment without being overly academic or technical. There are pros and cons to this assessment literacy campaign, however. While it has the potential to engage a larger audience, glossing over too many of the technical details may lead to inaccurate assumptions, including the notion that tests are the only way to measure student learning. Regardless of the type of test used, there will always be questions about the degree to which it measures student learning or school quality. Popham is successful in getting all of us, including teachers and parents, to continue asking questions and engaging in these important discussions about educational quality. As with most reform suggestions, however, the devil is in the details; elaborating on these details would go a long way toward helping parents and teachers understand the complexities of measuring educational quality.

Notes

¹ The Commission on Instructionally Supportive Assessment was formed with the support of the American Association of School Administrators, the National Association of Elementary School Principals, the National Association of Secondary School Principals, the National Education Association, and the National Middle School Association.

References

- Anastasi, A. (1981). Diverse effects of training on tests of academic intelligence. In B. F. Green (Ed.), *Issues in testing—coaching, disclosure, and ethnic bias*. *New Directions for Testing and Measurement*, no. 11 (pp. 5-20). San Francisco: Jossey-Bass.
- Atkin, J. M., Black, P., & Coffey, J. (Eds.). (2001). *Classroom assessment and the National Science Education standards*. Washington, DC: National Academy Press and the Committee on Classroom Assessment and the

National Science Education Standards, Center for Education, National Research Council.

- Bush, G. W. (2005, January 12). President discusses No Child Left Behind and high school initiatives. Retrieved May 19, 2005 from:
<http://www.whitehouse.gov/news/releases/2005/01/20050112-5.html>.
- Ruiz-Primo, M. A., & Li, L. (2004). On the use of students' science notebooks as an assessment tool. *Studies in Educational Evaluation*, 30(1), 61-85.
- Wilson, M. (Ed.). (2004). *Toward coherence between classroom assessment and accountability*. Chicago: University of Chicago Press.

Reviewer

Marsha Ing is a doctoral student in the Social Research Methodology division of UCLA's Graduate School of Education and Information Studies. Her research interests focus on methods of measuring student learning ranging from large-scale surveys to teacher assessment practices.