PROJECT MUSE®

High-Stakes Testing and the Default Philosophy of Education

Michael Gerard Gunzenhauser

Theory Into Practice, Volume 42, Number 1, Winter 2003, pp. 51-58
(Article)

Published by Ohio State University College of Education

➔ For additional information about this article

<https://muse.jhu.edu/article/41412>

High-Stakes Testing and the Default Philosophy of Education

In this article, the author argues that high-stakes testing may lead to a default philosophy of education that holds in high regard a narrow bundle of knowledge and skills. A default philosophy is defined as the vision of education that results from a lack of reflective, engaged dialogue among educators and school communities about their goals and practices. In the context of high-stakes testing, one predominant default philosophy results from an inordinate focus on the tests themselves. As has been shown in research studies throughout the United States, this creates a context in which conversations about the meaning and value of education cannot take place without performance on standardized tests taking center stage. Within this context, dialogue is more difficult, and other possible philosophies of education become harder to articulate and implement. When the default philosophy of education dominates in a school, school district, or state, the possibilities for improving education reform and innovation are limited.

THE RUSH TO ACCOUNTABILITY in the form of high-stakes testing has had multiple effects on public schools. In this article I draw atten-

tion to one impact that has been rarely addressed—the effect that high-stakes testing has had on the philosophies of education in operation in schools. By philosophy of education, I mean a vision for the purpose and value of education.

The impact of high-stakes testing is rather substantial, I will suggest, because high-stakes testing brings with it a “default” philosophy of education. In the current context of schooling, the default philosophy is one that places inordinate value on the scores achieved on high-stakes tests, rather than on the achievement that the scores are meant to represent. Because of the power of this default philosophy, teachers in the current climate may find themselves doing things that fall short of their visions of themselves as educators, such as spending extra time drilling students on practice tests, de-emphasizing or eliminating untested subject matter, or teaching to the test (Jones et al., 1999).

It is not just a matter of teachers being able (or not being able) to avoid changing their practices. I aim to show that accountability systems in the form of high-stakes testing create philosophical conflicts and effectively limit the possibilities for schools. In this article, I explore the philosophical foundations of high-stakes testing and show how a default philosophy of education may take precedence over the substantial work that teachers and school communities need to do daily to create their own philosophies of education. I argue that

Michael G. Gunzenhauser is an assistant professor of educational studies at Oklahoma State University.

dialogue among teachers and other members of the school community about the purpose and value of education is essential for creating the public education that communities want and need. I conclude by identifying implications of the default philosophy of education and suggest alternatives for how school communities may maintain dialogue as we all work within the context of high-stakes testing.

What is a Philosophy of Education?

First, a clear definition of a philosophy of education is needed. For the purposes of this article, a philosophy of education is defined as a set of ideas and commitments about the purpose and value of education that guides our practice and helps us make choices. In other words, a philosophy of education addresses *why* we educate so that we make better choices about *who*, *what*, *where*, *when*, and *how* we educate. A philosophy of education provides answers to fundamental questions about the role of education in a society. As Postman (1985) suggests:

The question is not, Does or doesn't public schooling create a public? The question is, What kind of public does it create? A conglomerate of self-indulgent consumers? Angry, soulless, directionless masses? Indifferent, confused citizens? Or a public imbued with confidence, a sense of purpose, a respect for learning, and tolerance? (p. 18)

Also needed is a notion of where a philosophy of education comes from and how it gets put into place in a school. Taking a democratic perspective, a philosophy of education arises from ongoing conversations with others—our students, our colleagues, the public, and the authors we read—that cause us to reflect on our own ideas and consider multiple perspectives. A philosophy is never complete; it is constantly enriched by reflection and ongoing conversation. Each conversation provides the terms for future discussions.

These conversations are fundamental to public education, which can be understood as a series of sites where many people come together to put into place their philosophies of education. Teachers, administrators, parents, students, and other members of the communities in which schools are located work together to negotiate their multiple philosophies and form a school community. Within

school communities, dialogues about philosophies of education are fundamental to the educational enterprise. A clear philosophy of education, an answer to the question of why we educate, provides the grounding for more particular decisions. As Postman (1985) articulates:

The answer to this question has nothing whatever to do with computers, with testing, with teacher accountability, with class size, and with the other details of managing schools. The right answer depends on two things, and two things alone: the existence of shared narratives and the capacity of such narratives to provide an inspired reason for schooling. (p. 18)

Along with Eisner (2002), I would argue that the lack of attention to these fundamental questions leads to aimless pursuits of school reform. In the rush to reform schools, too little emphasis is placed on the underlying philosophy of education that is being served. And also as Eisner advocates, the purpose and value of education should be the subject of community discussions:

We are not clear about what we are after. Aside from literacy and numeracy, what do we want to achieve? What are our aims? What is important? What kind of educational culture do we want our children to experience? In short, what kind of schools do we need? (p. 577)

As suggested by both Postman (1985) and Eisner (2002), rich dialogical responses to these questions should lead to richer and more purposeful educational experiences.

A Default Philosophy of Education Underlying High-Stakes Testing

A “default” philosophy of education is a term used to describe what results from a lack of reflective, engaged dialogue by educators and school communities about their goals and practices. The default philosophy underlying high-stakes testing is a philosophy of education in which tests designed to be part of a system of accountability drive the curriculum, limit instructional innovation, and keep educators from establishing their own priorities and visions. The philosophical underpinnings of high-stakes testing help to explain the conflicts that result.

While high-stakes testing itself is a varied phenomenon, for the purposes of this article, “high-stakes testing” refers to the use of standardized testing measures as criteria for determining the

quality of schools, promotion of children to the next grade, high school graduation, teacher bonuses, or the governance of a school. High-stakes testing is a method associated with the school accountability movement, which in turn is connected with the standards movement, a related development that has brought together various people who wish to maintain high standards for school curricula and high expectations for the performance of all students. The No Child Left Behind (NCLB) Act of 2001 expands the role of high-stakes testing by legislating their incorporation in states' school accountability programs.

The phenomenon of high-stakes testing, and the standards movement from which it springs, emerges from a particular philosophy—a behaviorist, positivist philosophy that places great emphasis on what can be measured quantitatively (Crotty, 1998). This is a philosophy that not only has implications for education but builds from a philosophy of reality and the ability of science to perceive that reality. This view no longer informs the work of today's psychometricians (i.e., those skilled in the administration and interpretation of objective psychological tests) who, since the influence of Karl Popper and other philosophers of science, have taken a more probabilistic and fallibilistic view of knowledge. Psychometricians and other statisticians believe that what we know scientifically is only known with a certain probability. For them, all knowledge is built on foundations but is fallible; it is our best approximation of the truth until we are proven false (Crotty, 1998; Phillips & Burbules, 2000; Siegel, 1997).

Psychometricians apply this distinction between foundations and fallibility in their use of standardized test measures. They are always careful to make a distinction between the measurement and what it is supposed to measure. This distinction is echoed in teachers' consternation when they know a student knows something but is unable to get the test question correct. That is because a measurement (e.g., a test score) is an approximation of something else, and test scores are typically approximations of student achievement, or, in other words, students' knowledge of subject matter.

Measuring student achievement is much more complex than, for example, taking measurements

in chemistry. While 0 °C is a reliable temperature at which we may expect water to freeze, a certain cut-off score for a graduation exam is much less reliable as an indicator of student achievement. Psychometricians work diligently to make sure their measurements become better and more useful in their ability to measure student achievement, but all of their measurements are fallible. Because tests are always *measures* of student ability, if we are to be consistent with their philosophical foundations, they must always be used with caution and a healthy dose of respect for their fallibility (Mertens, 1998; Phillips & Burbules, 2000).

Furthermore, psychometricians are careful to point out that not all forms of achievement are able to be measured. Nevertheless, standardized tests are considered to be effective components of standards reform because they can be used as a measure of students' achievement of standards. Policy makers and other advocates of high-stakes testing may claim that high-stakes testing is "scientific" or "objective," but the underlying philosophy should lead us to be more cautious about the test results. The philosophy itself borrows rhetorical grounding from positivist science, giving rhetorical weight to the perception that raising the stakes of tests is appropriate because the tests are administered to all students equally and without bias.

That defense, however, suggests a faith in measurement that even psychometricians and statisticians do not have (Linn, 2000). Measurement specialists advise caution in the use of standardized measures to make individual decisions about things such as diagnoses and grade promotion. Instead, a test should be one among many tools that schools use to evaluate students.

Those who advocate for the use of assessments warn against the inappropriate use of high-stakes tests. Linn (2000) has stated that "assessment systems that are useful monitors lose much of their dependability and credibility for that purpose when high stakes are attached to them" (p. 14). One problem associated with the use of test scores to judge the performance of schools is that scores tend to fluctuate from year to year for reasons unrelated to student achievement (Bracey, 2001; Linn & Haug, 2002). Statisticians can deal with those fluctuations by lengthening the accountability cycle in

order to avoid undue emphasis on test score changes from one year to the next.

From a scientific standpoint, high-stakes tests cannot do all that policy makers want them to do. Because of the high stakes attached to the tests, policy has had the unintended effect of encouraging a default philosophy of education: a vision of education that values highly what can be measured, and more problematically, it values most highly the measurement itself. Eisner (2001) calls this confusion of the role of tests "rationalization."

This does not mean that schools are philosophically doomed. However, when this default philosophy of education dominates, other possible philosophies of education are more difficult to articulate and implement. That is probably the most unfortunate aspect of high-stakes testing, because conversations in communities about the meaning and value of education cannot take place without performance on standardized tests taking center stage.

Paradoxically, some of the most difficult philosophies to articulate and implement are those consistent with much of the stated rationale for increased accountability, notably the documented gaps in achievement between students of different economic, ethnic, and racial backgrounds. Based on the text of the NCLB Act, the rationale for high-stakes testing seems to be high standards for all children, accountability, parental involvement, and particular attention to populations of students that schools have neglected. The language in Title I of the Act is explicit on these points. The second and third goals for Title I of the act are these:

2. Meeting the educational needs of low-achieving children in our Nation's highest-poverty schools, limited English proficient children, migratory children, children with disabilities, Indian children, neglected or delinquent children, and young children in need of reading assistance;
3. Closing the achievement gap between high- and low-performing children, especially the achievement gaps between minority and nonminority students, and between disadvantaged children and their more advantaged peers.

These are essentially the same rationales shared by other advocates of the standards movement in education who do not advocate high-stakes

testing (Thompson, 2001). For many years, various educators and researchers have advocated high standards, pushing for stronger curricula and the use of assessment. Part of the push for assessment is to identify problem areas—aspects of the curriculum that need strengthening, students that need additional attention, and schools that need resources (King & Mathers, 1997; McGehee & Griffith, 2001). The intention is to use the resources of educational research to help schools and districts improve and to insist that all students have access to high quality curricula (Bracey, 2001; Jorgenson & Vanosdall, 2002). Accountability programs that include high-stakes testing support the high standards movement but place greater significance on the use of standardized tests to enforce standards.

In other words, there are other ways people have advocated for achieving the second and third goals of the NCLB Act, while at the same time maintaining high standards. In most of the literature on high standards, the underlying vision is equity, both in terms of opportunity and outcomes (Thompson, 2001). For advocates of inquiry learning in science, for example, the emphasis is on raising expectations across the board for students' achievement in science, with inquiry-based science seen as promoting a higher standard of achievement (Jorgenson & Vanosdall, 2002). At the same time, this push for better science instruction requires that assessments match up with the goals of inquiry-based science.

The concern for equity can be a significant part of a philosophy of education, but there is a prior philosophical question that needs to be asked: What kind of education would we like everyone to have an equal opportunity to achieve? To this question, advocates of high standards have multiple responses, tied together by things such as higher-order concepts, creativity, and the ability to apply knowledge to new settings. The notion of high standards itself does not necessitate certain views of education, just those that share with it a concern for equity. This changes, however, with the addition of high-stakes testing to high standards.

High-Stakes Testing and the Default Philosophy in Practice

In practice, there is strong evidence that the high-stakes testing aspect of accountability programs

has taken on a dominant position, with the result that a default philosophy of education has been inordinately influential. First, it should be noted that some positive effects of high-stakes testing have been reported. In such cases, the high profile of tests has led educators to devote greater attention to populations of students whose achievement may have been previously neglected. Certainly Kozol's (1991) work suggests that neglect is a significant, ongoing problem in the public education of economically and socially vulnerable children in the United States.

Researchers in Texas studied four school districts that, under the pressure of high-stakes testing associated with that state's accountability program, raised the test scores of minority and second-language students and narrowed the gap between White and non-White achievement (Skrla & Scheurich, 2001). Faced with the public availability of test scores of not only the entire district's population of students but also a breakdown of scores by racial and socioeconomic groups, educators in these districts addressed the achievement gap explicitly. The researchers found evidence that the presence of the high-stakes component forced the schools to change their deficit-thinking model of non-White students. These districts were difficult for the researchers to identify, however, suggesting that the effects are not widespread (Skrla & Scheurich, 2001).

In a study of schools in North Carolina implementing an arts-based program at the same time that the state instituted its accountability program with high stakes for schools labeled *low performing*, researchers identified an elementary school principal who appreciated the attention her school received after it was labeled *low performing* (Gunzenhauser & Noblit, 2001). Being named *low performing* after the first year of the state's accountability program enabled her to receive additional resources from the district superintendent for remedial instruction—resources she believed would not otherwise have been available to her. She had argued for many years that the high poverty rate at her school (and the large percentage of newly trained teachers) led to inequities in funding in this district. Presumably, the labeling of schools as *low performing* made it politically possible for neglected schools to obtain a greater share of resources from the school board.

At the same time, however, this North Carolina school experienced some of the detrimental effects of high-stakes testing. In order to devote more time for test preparation, the principal adopted a policy teachers called "90/90/60," in which instruction each day encompassed 90 minutes of reading, 90 minutes of math, and 60 minutes of writing, the only three areas of the curriculum tested in the state's elementary schools (Gunzenhauser & Noblit, 2001). Other subjects, including science, social studies, physical education, and the arts (this school had full-time specialists to teach visual art, music, dance, and drama), vied for the remaining time in the school day. Test scores improved, and the school eventually lost its *low performing* label, but large portions of the state curriculum were left untaught.

This phenomenon of curriculum "narrowing" (King & Mathers, 1997) is reflected in case studies of other schools participating in the same arts-based program, even though the arts program encouraged schools to teach the state's full curriculum (Gordon, 2002; Groves, 2002; Gunzenhauser & Noblit, 2001; Murillo & Flores, 2002). According to a survey of teachers in North Carolina, curriculum narrowing was prevalent; the survey revealed a widespread decline in time devoted to the untested areas of science and social studies after the state's high-stakes accountability program was instituted (Jones et al., 1999).

Standards-based reforms are also supposed to help teachers identify the most important goals in their curricula, and testing can do that, but only if test scores are reported in relation to goals and if structures are in place to use those scores to improve curricula. In their study, Elmore and Fuhrman (2001) found that in order to use assessments effectively, schools need internal accountability first, and by that they mean developing answers to fundamental aspects of accountability: "what they expect of students academically, what constitutes good instructional practice, who is responsible for student learning, and how individual students and teachers account for their work and learning" (p. 69). Elmore and Fuhrman found that in most schools under the gun of high-stakes testing, teachers are working harder, spending more time, and exerting more effort preparing students for testing. However,

schools are not fundamentally improving what they are doing, Elmore and Fuhrman contend; instead they are devoting inordinate time with concern about students' scores and not enough to students' learning.

Instruction has changed as well, with schools devoting time to test preparation strategies as opposed to content, and teachers are displeased (McNeil, 2000a; Patterson, 2002). McNeil (2000a, 2000b) has noted the increased attention to test preparation materials and pep rallies as the stakes of tests have gotten higher, crowding out other instructional objectives. Boston high school principal Linda Nathan (2002) questions the spending of curriculum dollars on test-preparation materials that only help students become better test takers and make little contribution to their overall education. Nathan argues that she would rather spend the money on reducing class sizes, an expenditure that fits more appropriately into her philosophy of education. Because the default philosophy of high-stakes testing is so strong, however, Nathan is unable to ignore the need to devote resources to preparing students for the graduation exam. She decries the effects on the individual experiences of students, whose only reason for dropping out of high school is the high-stakes graduation test itself. Not only does the default philosophy lead to conflict, but it leads to practices that defeat the goals of equity that are the foundation of the high-standards movement. Beyond Nathan's experience, there is concern that on a national level, high-stakes graduation tests are leading to higher dropout rates among the lowest-achieving students and having no quantifiable effect on other measurements of achievement (Jacob, 2001).

As several studies have suggested, teachers are critical of the effects of high-stakes testing on their work. I would argue that this is indicative of teachers having to engage in instructional practices that are inconsistent with their philosophies of education (Groves, 2002; Jones et al., 1999; Murillo & Flores, 2002). I propose that without significant conversations that not only address our ideals for education but also the political context in which we work as educators, our work will perpetuate a default philosophy of education that nearly all educators will find troubling.

What Educators Can Do in the Current Context

It is not likely to be constructive to implore teachers and school communities to develop and adopt philosophies of education that ignore the context and pressures of high-stakes testing. Certainly the most significant changes need to come at the state and national levels, with further study and consideration of the role and effects of the high stakes associated with accountability policies (Gunzenhauser & Gerstl-Pepin, 2002; King & Mathers, 1997; Sirotnik, 2002). Most needed is a redefinition of public school accountability in which whole communities, including lawmakers and psychometricians, share in the responsibility for high standards and equity in education (Sirotnik, 2002).

Dialogue is needed at every level and needs to be informed by multiple philosophies of education. In individual districts and school communities, the following approaches that emerge from the literature may be helpful for schools to maintain dialogue and take greater control of their educational destinies.

Maintain dialogue in schools

Above all, despite pressures from high-stakes testing programs, it is crucial that teachers and school communities remain engaged in dialogue about the value and purposes of education that provide the goals for public education. Individual teachers, school communities, and school districts should develop rich philosophies of education that are based not only on reflections on practice but the careful consideration of alternative visions. Principals are key figures in maintaining this dialogue, because in many settings, the principal has responsibility for facilitating or directing dialogue, particularly in relation to test preparation (Gordon, 2002; Patterson, 2002). Principals should put themselves in the position of protecting students from the stresses of high-stakes tests and discourage instructional methods that raise scores without serving the school's philosophy of creating educated persons.

Expand internal accountability

According to Elmore and Fuhrman (2001), internal accountability should precede external accountability. Schools can make use of multiple

forms of assessment that reflect a broad curriculum, including but not limited to standardized tests. These expanded systems, more consistent with schools' philosophies, may provide alternative forms of accountability, particularly if used to communicate with parents and communities about students' progress.

Engage high standards

If schools can indeed agree that standards-based reform fits with their philosophies of education, then the most appropriate way to deal with high-stakes testing is to engage not with the tests but with the high standards. This will involve schools connecting with the curriculum that the district or state is expecting them to teach. As happened in the arts-based program in North Carolina discussed earlier, schools were encouraged to focus on the state curriculum, a holistic curriculum, and teach it in its entirety (Cozart & Gerstl-Pepin, 2002). To do this, these schools needed to become familiar with the entire scope and sequence of the state curriculum so they could more effectively integrate the arts across the curriculum. Those who did so successfully were also effective in warding off the default philosophy associated with the high-stakes testing program (Corbett, Wilson, Noblit, & McKinney, 2001).

Connect to higher-order concepts

Preparing for individual standardized tests can easily lead to ignoring the higher-order concepts on which comprehensive, holistic curricula are based (Cozart & Gerstl-Pepin, 2002). Teachers need to have an understanding of the higher-order concepts associated with their curricula and push for changes in tests if the tests do not assess students' acquisition of those higher-order concepts. Part of this process is that teachers need to articulate how the major concepts and themes across curricula and grade level effectively serve their overall vision of an educated person. Teachers may ask themselves, for instance, how the district's social studies curriculum contributes to the development of civilly minded and informed students.

Spend money to support goals, not scores

One troubling phenomenon is the extent to which schools are using precious resources on test

preparation materials (McNeil, 2000a, 2000b; Nathan, 2002). Measures to raise scores such as pep rallies, cram sessions, and after-school drilling may result in higher scores in the short run but are unlikely to serve the longer-term goals of learning. Schools should minimize these purchases and instead prioritize spending on goals that are essential to their philosophies of education.

Develop alternative accountability goals

Schools and school districts should develop plans for addressing disparities in equity of opportunity and achievement that are more meaningful than high-stakes testing measures. As mentioned previously, the North Carolina principal who received additional resources when her school was named *low performing* is part of a district that funneled resources for remediation into its low performing schools, which enrolled higher percentages of non-White and low-income children (Gunzenhauser & Noblit, 2001).

Take community responsibility for all schools

Competition associated with high-stakes testing makes working in schools labeled *low performing* difficult (Murillo & Flores, 2002). If one of the main goals of accountability is to locate schools that need additional resources (King & Mathers, 1997), then school districts (or groups of districts in interdependent metropolitan areas) need to take responsibility for identifying those schools and distributing resources across schools where they are most needed. Research studies are emerging now that show just how much even parents of school children base their perceptions of schools not on informed perspectives but by conjecture and rumor (Holme, 2002). Public support for accountability programs arises partly from the public's lack of knowledge about what schools in their own communities are doing and the goals they serve. Teachers need to open and maintain dialogue with their communities about the direction of schools.

Conclusion

Based on the current research about high-stakes testing, it has become very difficult for educators to discuss the purpose and value of education. As I have argued, this narrowing of the conversation requires attention at the national and state level to lower the

stakes of testing, because it puts in place a default philosophy of education that is indefensible. At the school and school district level, it is essential that schools maintain dialogue about the purpose and value of schooling among the various members of the school community and not allow themselves to settle for the default philosophy of education associated with high-stakes testing.

References

- Bracey, G.W. (2001). The 11th Bracey report on the condition of public education. *Phi Delta Kappan*, 83(2), 157-169.
- Corbett, D., Wilson, B., Noblit, G., & McKinney, M. (2001). *The arts, identity, and comprehensive education reform: A final report from the evaluation of the A+ Schools Program*. Winston-Salem, NC: Thomas S. Kenan Institute for the Arts.
- Cozart, S.C., & Gerstl-Pepin, C. (2002). Teacher morale gets a boost. *Journal of Staff Development*, 23(2), 29-31.
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. London: Sage.
- Eisner, E.W. (2001). What does it mean to say a school is doing well? *Phi Delta Kappan*, 82(5), 367-372.
- Eisner, E.W. (2002). The kind of schools we need. *Phi Delta Kappan*, 83(8), 576-583.
- Elmore, R.F., & Fuhrman, S.H. (2001). Holding schools accountable: Is it working? *Phi Delta Kappan*, 83(1), 67-72.
- Gordon, J. (2002). From Broadway to the ABCs: Making meaning of arts reform in the age of accountability. *Educational Foundations*, 16(2), 33-53.
- Groves, P. (2002). "Doesn't it feel morbid here?" High-stakes testing and the widening of the equity gap. *Educational Foundations*, 16(2), 15-31.
- Gunzenhauser, M.G., & Gerstl-Pepin, C.I. (2002). Guest editors' introduction: The shifting context of accountability in North Carolina and the implications for arts-based reform. *Educational Foundations*, 16(2), 3-14.
- Gunzenhauser, M.G., & Noblit, G.W. (2001). *Reforming with the arts: Creativity in A+ classrooms and schools*. Winston-Salem, NC: Thomas S. Kenan Institute for the Arts.
- Holme, J.J. (2002). Buying homes, buying schools: School choice and the social construction of school quality. *Harvard Educational Review*, 72(2), 177-205.
- Jacob, B.A. (2001). Getting tough? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis*, 23(2), 99-121.
- Jones, G.M., Jones, B.D., Hardin, B., Chapman, L., Yarborough, T., & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81(3), 199-203.
- Jorgenson, O., & Vanosdall, R. (2002). The death of science? *Phi Delta Kappan*, 83(8), 601-605.
- King, R.A., & Mathers, J.K. (1997). Improving schools through performance-based accountability and financial rewards. *Journal of Education Finance*, 23(2), 147-176.
- Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Crown Publishers.
- Linn, R.L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Linn, R.L., & Haug, C. (2002). Stability of school-building accountability scores and gains. *Educational Evaluation and Policy Analysis*, 24(1), 29-36.
- McGehee, J.J., & Griffith, L.K. (2001). Large-scale assessments combined with curriculum alignment: Agents of change. *Theory Into Practice*, 40(2), 137-144.
- McNeil, L.M. (2000a). Creating new inequalities: Contradictions of reform. *Phi Delta Kappan*, 81(1), 728-734.
- McNeil, L.M. (2000b). Sameness, bureaucracy, and the myth of educational equity: The TAAS system of testing in Texas public schools. *Hispanic Journal of Behavioral Sciences*, 22(4), 508-523.
- Mertens, D. (1998). *Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches*. Thousand Oaks, CA: Sage.
- Murillo, E.G., & Flores, S.Y. (2002). Reform by shame: Managing the stigma of labels in high-stakes testing. *Educational Foundations*, 16(2), 93-108.
- Nathan, L. (2002). The human face of the high-stakes testing story. *Phi Delta Kappan*, 83(8), 595-600.
- Patterson, J.A. (2002). Exploring reform as symbolism and expression of belief. *Educational Foundations*, 16(2), 55-75.
- Phillips, D.C., & Burbules, N. (2000). *Postpositivism and educational research*. Lanham, MD: Rowman & Littlefield.
- Postman, N. (1995). *The end of education: Redefining the value of school*. New York: Vintage Books.
- Siegel, H. (1997). *Rationality redeemed? Further dialogues on an educational ideal*. New York: Routledge.
- Sirotnik, K.A. (2002). Promoting responsible accountability in schools and education. *Phi Delta Kappan*, 83(9), 645-673.
- Skrila, L., & Scheurich, J.J. (2001). Displacing deficit thinking in school district leadership. *Education and Urban Society*, 33(3), 235-259.
- Thompson, S. (2001). The authentic standards movement and its evil twin. *Phi Delta Kappan*, 82(5), 358-362.